

Sustainability and Maturation of School Turnaround: A Multiyear Evaluation of Tennessee’s Achievement School District and Local Innovation Zones

Lam D. Pham 

Vanderbilt University

Gary T. Henry

University of Delaware

Adam Kho

University of Southern California

Ron Zimmer

University of Kentucky

Recent evaluations of reforms to improve low-performing schools have almost exclusively focused on shorter term effects. In this study, we extend the literature by examining the sustainability and maturation of two turnaround models in Tennessee: the state-led Achievement School District (ASD) and district-led local Innovation Zones (iZones). Using difference-in-differences models, we find overall positive effects on student achievement in iZone schools and null effects in ASD schools. Additional findings suggest a linkage between staff turnover and the effectiveness of reforms. ASD schools experienced high staff turnover in every cohort, and iZone schools faced high turnover in its latest cohort, the only one with negative effects. We discuss how differences in the ASD and iZone interventions may help explain variation in the schools’ ability to recruit and retain effective teachers and principals.

Keywords: *school turnaround, school reform, teacher and principal turnover, teacher and principal effectiveness, long-term effects, difference-in-differences*

FEDERAL policymakers have invested substantial resources into turning around persistently low-performing schools, with the expectation that interventions would swiftly improve student achievement within 3 years (Herman et al., 2008). However, research evaluating school turnaround reports mixed results (Dee, 2012; Dragoset et al., 2017; Zimmer et al., 2017). Since most of the existing research evaluates turnaround in the first few years of reform, results may be mixed because some reforms have immediate effects while others need longer to take hold. Longer term reforms align with the literature on earlier comprehensive school reform models, which find that school improvement efforts need more than 3 years to produce positive results (Aladjem et al., 2010; Berends et al., 2002; Bloom et al., 2001; Borman et al., 2003; Gross et al., 2009; Stuit, 2010).¹ Moreover, even if reforms do produce positive results early

on, little is known about the sustainability of turnaround reforms. This article fills these gaps in the literature by examining both schools that have implemented turnaround reforms for multiple years and later cohorts that begin reforms after the turnaround models have matured.

Evidence from more mature turnaround initiatives is especially important under the current Every Student Succeeds Act (ESSA). Under ESSA, more states are designing reforms that will operate for longer. For example, North Carolina’s ESSA plan gives its lowest-performing schools up to 12 years of school improvement supports (North Carolina Department of Public Instruction, 2018). In this article, we inform the sustainability of longer term ESSA reform models using updated data from 6 years of reform in Tennessee. These data allow us to contribute new evidence from one of the few turnaround initiatives across the country



that have been in continuous operation for longer than 3 years (exceptions include Chin et al., 2019; Harris & Larsen, 2016, 2019).²

We build on a 3-year evaluation by Zimmer et al. (2017) of Tennessee's Achievement School District (ASD) and local Innovation Zones (iZones), both of which have operated multiple cohorts of schools. We further describe both models below, but note that the ASD relies on dramatic changes in school governance where chronically low-performing schools are removed from local districts and restarted, mainly by charter management organizations. In contrast, local iZones place low-performing schools into intradistrict networks that are managed by dedicated district staff. Although the ASD and iZones have different governance structures, both models replaced teachers and principals in the first year of turnaround, and after 3 years, Zimmer et al. (2017) found that iZone schools produced significant positive student achievement effects while ASD schools produced null effects.

With novel data from a longer panel, we examine whether iZone schools have sustained positive effects and whether the ASD has produced positive results given more time. We estimate the ASD and iZone effects using difference-in-differences (DID) models that compare ASD and iZone schools with similarly low-performing schools receiving no turnaround interventions. We refer to these low-performing, nonturnaround schools as comparison schools and describe them further below. Averaging across all years and cohorts, we find that iZone reforms raised test score gains by approximately 0.10 standard deviations (*SD*), which is similar in magnitude to the positive effects reported by Zimmer et al. (2017) and from reforms in Massachusetts (Papay & Hannon, 2018) and San Francisco (Sun et al., 2017). In contrast, ASD schools did not produce significant effects in any cohort, year, or subject.

In addition to student achievement, our study expands on Zimmer et al. (2017) by examining how ASD and iZone reforms affect student attendance and mobility, along with the turnover rate, effectiveness, and average experience of teachers and principals. Neither ASD nor iZone reforms significantly affected student attendance or mobility; however, both models increased teacher and principal turnover in Year 1, when staff replacement was intentional. After Year 1, ASD schools continued experiencing high teacher and principal turnover, with lower performing and less-experienced replacements. Because of high-staff turnover, ASD schools may have had trouble maintaining a consistent improvement strategy, which helps explain the null ASD results. In contrast, after Year 1, staff turnover in iZone schools returned to levels similar to those in comparison schools. Moreover, the earlier and most successful iZone cohorts recruited effective teachers and principals, while the most recent and least successful cohort was the only one to recruit less-effective teachers and less-experienced principals. Overall, we find a

consistent relationship between the successful recruitment and retention of effective educators and positive effects on student achievement. In addition to reporting results, we discuss below potential reasons why ASD and iZone schools experienced differences in staff turnover.

Like Zimmer et al. (2017), most existing studies of turnaround concentrate on the earlier years reform when interventions focus on addressing barriers to improvement. In this study, we instead examine later stages when reforms are focused on increasing capacity and sustaining improvements. In doing so, we address three questions:

1. To what extent have iZone schools been able to sustain positive effects?
2. Over a longer, 6-year period, do ASD schools produce positive effects on student achievement?
3. To what extent does either intervention affect student attendance and mobility along with the turnover rate, effectiveness, and experience of teachers and principals in turnaround schools?

School Turnaround in Tennessee

Tennessee's current turnaround approach originated from the state's 2010 First to the Top (FtT) legislation. The legislation required Tennessee's Department of Education (TDOE) to intervene in the state's lowest-performing 5% of schools, called priority schools. Using FtT, Tennessee applied for and won approximately US\$500 million of Race to the Top (RtT) funding to implement its proposed initiatives. Beginning in 2012–2013, priority schools were either placed into the ASD, joined an iZone, closed, or received no interventions.

Tennessee's boldest turnaround model is the ASD, a statewide school district staffed by full-time TDOE personnel and led by a superintendent who reports directly to Tennessee's Commissioner of Education. When chosen for the ASD, schools are removed from their local education agency (LEA) to be governed by TDOE, restarted under new management, and required to replace the principal and at least 50% of teachers. ASD schools were restarted under a charter management organization (CMO), except five that were directly managed by the ASD.³ The ASD relied on CMOs, because TDOE leaders believed freedom from bureaucratic oversight would give schools flexibility to focus on reforms, and previous research finds that some Tennessee charter schools out-performed the state average (Center for Research on Education Outcomes, 2019). Notably, ASD schools are not schools of choice but continue to enroll students from their local catchment area.

As ASD operations were beginning in 2012, TDOE also approved the creation of local iZones in Shelby County Schools (Memphis) and Metro-Nashville Public Schools. Local iZones differ from the ASD primarily in that iZone

TABLE 1
Number of Schools by Reform Approach and Cohort

| Cohort No. | First year of turnaround reforms | Non-iZone, non-ASD priority schools | ASD schools | iZone schools | Closed |
|------------|----------------------------------|-------------------------------------|-------------|-----------------|--------|
| Cohort 1 | 2012–2013 | 65 | 6 | 13 ^a | 1 |
| Cohort 2 | 2013–2014 | 45 | 6 | 11 | 3 |
| Cohort 3 | 2014–2015 | 59 ^b | 5 | 4 | 8 |
| Cohort 4 | 2015–2016 | 40 | 5 | 11 ^c | 3 |
| Cohort 5 | 2016–2017 | 26 | 4 | 3 | 7 |

Note. ASD = Achievement School District; iZones = Innovation Zones.

^aTwo schools were placed into the Nashville iZone in 2012–2013 even though they were not on the 2012 priority list. These two schools are on the 2014 priority list. ^bThe number of priority schools increases in 2014–2015, because Tennessee released a new list of priority schools in 2014. ^cTwo schools in the fourth iZone cohort were closed in the following year. 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 or 2014 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.

schools remain part of their LEA and are placed into an intradistrict network with other low-performing schools in the district. Local iZones are supported by full-time district employees and led by a director who reports to the superintendent.⁴ Like the ASD, iZone schools must also replace the principal. Unlike the ASD, iZone schools are not required to replace teachers, though almost all iZone schools did replace at least 50% of teachers on joining an iZone. Then, after initial staff replacements, the iZones had broad autonomy to manage their daily operations. Since 2012, two additional districts have opened iZones: Hamilton County Schools and Knox County Schools. In this article, we pool all four iZones together, but our conclusions do not change when we examine only the Memphis iZone, where the vast majority of priority schools are located.⁵

During its first 3 years, the ASD was funded by part of Tennessee’s RttT grant and with philanthropic support. When RttT funds were depleted, TDOE continued funding the ASD from its budget. When iZones were created, districts received School Improvement Grant (SIG) funding from TDOE to support all iZone schools. When SIG funding ended, districts used a combination of philanthropic support, state funds, and portions of their own budget to continue funding current schools and any new schools joining their iZones. Notably, iZone schools used their funding for performance pay incentives to recruit and retain effective teachers. Both the ASD and iZones have continued receiving funding and support, differentiating this article from studies that examine turnaround effects after the active reforms have ended. Furthermore, ongoing funding as part of the TDOE and district budgets show that these models have moved toward building capacity and sustaining the reforms.

Some priority schools not placed into either the ASD or an iZone were closed, and our communications with district leaders indicate that priority schools were only closed if they had very low enrollment. Closed schools remain part of our sample in the years they are open.⁶ Besides the priority designation itself, priority schools not taken over by the ASD,

placed into an iZone, or closed did not receive any other funding or interventions from the state or their LEAs.

ASD schools were chosen from the priority list based on school feeder patterns and whether a CMO could be matched with the school (Tennessee Department of Education, 2012). The matching process allowed CMOs to present their proposed reform plan to a committee of ASD leaders and local stakeholders (i.e., parents and community leaders). Under its feeder pattern approach, the ASD began taking over elementary schools in its first cohort and would later target middle and high schools receiving students from these ASD elementary schools. The Memphis iZone similarly used feeder patterns to choose schools, whereas all priority schools not placed into the ASD or closed in Nashville, Hamilton County, and Knox County were assigned to their respective iZones. Although these practices loosely guided which schools would join the ASD and iZones, our communications with state and district leaders suggest that the process was not systematic and often rushed because of political pressure to quickly intervene in priority schools. These pressures suggest that schools were not systematically chosen for turnaround because of their performance levels or potential for future growth. To mitigate concerns that schools were systematically chosen, we examine baseline characteristics and test score trends of ASD and iZone schools relative to comparison schools and found them to be highly similar.⁷

After Tennessee released an initial list of 83 priority schools in 2012, both the ASD and iZones began operating schools chosen from this list.⁸ In Table 1, we show the number of schools in each ASD and iZone cohort. Across the 6 years of turnaround examined in this article (2012–2013 through 2017–2018), five cohorts joined the ASD or an iZone, with no additional schools joining either model in 2017–2018. Beginning in 2012–2013, the first cohort includes six ASD, 13 iZone, and one school that was closed, leaving 65 priority schools that received no interventions. The second cohort was chosen from these remaining 65

priority schools to begin reforms in 2013–2014. The second cohort includes six ASD, 11 iZone, and three schools that were closed, leaving 45 priority schools that had not received any reforms by 2013–2014. New schools continued joining the ASD and iZones in Cohorts 3, 4, and 5, resulting in a total 26 ASD, 42 iZone, and 26 priority schools that have not received any turnaround interventions by 2017–2018. We show below that student characteristics are very similar in each of the five ASD and iZone cohorts, suggesting that different results across cohorts are not driven by changes in the types of schools being taken over.

Additional priority schools were identified in 2014 when Tennessee released a second priority list with 33 schools that were not on the 2012 list.⁹ Our comparison sample includes all nonturnaround schools on either the 2012 or 2014 priority list, but results are similar when we use only schools on one of the two lists (Appendix Table A3). Priority school designation can itself be considered an intervention because being labeled a low-performing school could motivate schools to demonstrate improvement, but all schools in our sample are priority schools. Thus, our analysis compares priority schools receiving one of two turnaround interventions with priority schools receiving no interventions. Finally, note that during this time period, no schools exited either the ASD or iZones.

Review of the Whole-School Reform Literature

Whole-school reforms have received substantial policy interest and investment, especially through federal initiatives including Project Follow Through (Egbert, 1981), New American Schools (Berends et al., 2002), Comprehensive School Reform (Aladjem et al., 2010), School Improvement Grants (Dragoset et al., 2017), Race to the Top (Heissel & Ladd, 2018; Henry et al., 2015), and NCLB (No Child Left Behind) waivers under the Obama administration (Bonilla & Dee, 2017; Dee & Dizon-Ross, 2017; Dougherty & Weiner, 2019; Hemelt & Jacob, 2017). Though specific interventions varied, these reform initiatives share the idea that incremental interventions are insufficient and that low-performing schools do not have individual capacity to improve. Under ESSA, the approach to whole-school reform is transitioning to flexible, state-formulated plans, and this flexibility suggests that states will need guidance on how to shape their reform plans in light of prior whole-school reform initiatives. However, evaluations accompanying these reform initiatives report mixed results (e.g., Dragoset et al., 2017), so to better understand the literature, we turn to the underlying theory of action for turnaround.

Broadly called “turnaround,” the most recent national efforts to improve low-performing schools have been primarily shaped by four models prescribed under the federal SIG program: transformation, turnaround, restart, and closure (Perlman & Redding, 2011). Local iZones in Tennessee

implemented reforms aligned with the transformation model: requiring schools to replace principals but not teachers. The ASD implemented the restart model by closing and reopening schools under new management. These reform initiatives share a theory of action with three stages: (1) disrupt barriers to improvement (Herman et al., 2008); (2) build an improvement infrastructure (Peurach & Neumerski, 2015); and (3) sustain the momentum (Copland, 2003; Kutash et al., 2010). This three-stage theory of action suggests that schools may need time to reestablish stable operations after initial, disruptive reforms and raises questions about sustainability.

Most existing research examines the first 3 years of turnaround (i.e., the disruption stage). These studies describe interventions that focus on building human capacity by recruiting effective teachers and leaders (Carlson & Lavertu, 2018; Kho et al., 2018), and some evaluations find larger positive results in schools that successfully recruit new teachers and principals relative to schools that do not bring in new staff (Anrig, 2015; Carlson & Lavertu, 2018; Dee, 2012; Strunk et al., 2016). These lessons are valuable, but the disruption stage is a tenuous time to evaluate reforms, because interventions can create rapid changes that may detract from potential positive effects.

In the second, infrastructure-building stage, researchers assert that successful reforms should focus on retaining effective staff because turnover could force principals to hire less-experienced or less-effective teachers (Calkins et al., 2007). Similarly, in-service development initiatives can be thwarted by turnover because teachers and leaders take newly gained skills with them when they leave (Henry & Harbatkin, 2018). Besides staff turnover, researchers also emphasize student attendance and mobility, because improving school operations means little when students are not in school (Dougherty & Weiner, 2019). These existing studies suggest that student attendance, student mobility, and staff retention help illuminate the reform process, but there is insufficient research documenting how ongoing reforms affect these student and staff characteristics.

Even fewer studies examine turnaround in the third stage, as reformers work to sustain the model. One model that has persisted is the portfolio district in New Orleans. Research in New Orleans suggests the model faced new challenges over time (Harris & Larsen, 2016, 2019; McEachin et al., 2016; Ruble, 2015), such as a shrinking pool of effective teachers and leaders available to transfer into turnaround schools. In New Orleans, efforts to sustain the model included recruiting charter operators and partnering with alternative certification programs like Teach for America and TeachNOLA (Harris & Larsen, 2016, 2019). However, even in New Orleans, there is little evidence on whether schools continue to successfully recruit and retain effective staff as the model expands over time. We help fill these gaps in the literature by examining the ASD and iZones as they mature into the capacity-building and sustaining stages.

Method

Data

Data are provided by TDOE and managed by the Tennessee Education Research Alliance. The administrative data sets capture characteristics of students, teachers, and principals for all Tennessee public schools. The student data sets contain a rich set of characteristics including gender, race, eligibility for free or reduced-price meals (FRPM), English language learner status (ELL), and special education status (SpED). Teacher and principal characteristics include gender, race, and degree attainment. Our analytic sample consists of students in tested subjects and grades, and we restrict the analysis to only priority schools. This leaves a sample of 315,663 student-year-subject observations with nonmissing covariates, with unique records for 63,419 students, 9,300 teachers, 311 principals, and 116 schools.

Measures

Our primary outcome of interest is student test scores. In Tennessee, students take end-of-grade exams in Grades 3 to 8 and end-of-course exams in high school subjects. Because of testing complications that occurred during the transition to new assessments in 2015–2016, EOG scores in Grades 3 to 8 were not reported for any Tennessee school; therefore, our analysis does not include any scores from 2015 to 2016.¹⁰ To have a common metric across years, we standardize the scale scores by subject, grade, year, and for the EOCs, semester.

Based on our literature review, we also examine effects on student mobility and attendance, along with turnover, effectiveness, and experience of teachers and leaders. By linking students and educators with schools in each year, we create student mobility and staff turnover measures by generating indicators that equal one when the student or educator is observed in a school that is different from her school in the prior year. These new-to-school indicators are beginning-of-year turnover measures. We test whether our results are robust when using an alternative, end-of-year turnover indicator for whether the student or educator is no longer in her current school in the following year. Consistent with existing literature, we reach similar conclusions when using this alternative mobility measure (Ronfeldt et al., 2013).

For students, attendance is the proportion of school days attended, and the mobility indicator captures whether the student is new to her school due to a nonstructural move.¹¹ For teachers, we examine three outcomes: turnover, years of experience, and prior year value-added scores for incoming teachers. In Tennessee, teachers in tested grades and subjects receive a value-added score on the Tennessee value-added assessment system (TVAAS) ranging from one to five. We use the TVAAS score to measure teacher effectiveness in the year before they enter an ASD or iZone school.

For principals, we use parallel measures that include principal turnover, observation scores, and years of principal experience. The principal observation scores range from one to five and are annual supervisor ratings given to school leaders using a rubric based on the Tennessee Instructional Leadership Standards or TILS (Grissom et al., 2018). Scores on the TILS rubric are given by the district superintendent or her designee (usually the principal's supervisor).¹² Grissom et al. (2018) find that principals' observation scores are internally consistent, stable over time, and predictive of other performance measures, such as student achievement growth and teachers' ratings of their school leadership.

Analytic Strategy

We use DID models that examine student outcomes (test scores, mobility, and attendance) and teacher and principal outcomes (turnover, effectiveness, and experience). We examine each outcome, y , for individual i in grade g , school s , and year t :

$$y_{igst} = \beta_0 + \beta_1 \text{After}_{gst} + \beta_2 (\text{ASD}_s * \text{After}_{gst}) + \beta_3 (\text{iZone}_s * \text{After}_{gst}) + \mathbf{X}'_{igst} \boldsymbol{\delta} + \mathbf{D}'_{st} \boldsymbol{\pi} + d_s + \varepsilon_{igst} \quad (1)$$

Equation 1 models y as a function of *After*, an indicator for years after turnaround begins, and interactions between *After* and indicators for schools that are ever *ASD* or *iZone*.¹³ The model also includes \mathbf{X}'_{igst} , a vector of individual control variables. For student-level outcomes, these individual controls include indicators for gender, race, FRPM, ELL, and SpED.¹⁴ For teacher and principal outcomes, individual control variables include indicators for gender, race, and graduate degree attainment.¹⁵ We include these individual control variables to mitigate any concerns that our results are driven by changes in student, teacher, or principal characteristics, but our results are not sensitive to their inclusion (Appendix Table A4).

Equation 1 also includes a school fixed effect (d_s) and time-varying school compositional characteristics (\mathbf{D}'_{st}). The school fixed effect controls for unobserved, time-invariant school characteristics that are associated with both turnaround and each outcome. Also, we include \mathbf{D}'_{st} to control for compositional characteristics that change over time, including student level covariates aggregated to the school level: gender, race, FRPM, ELL, and SpED.¹⁶ We follow standard practice in existing literature by controlling for these characteristics to ensure that our results are not driven by changes in the students enrolled in either ASD or iZone schools. Changes in student composition during the intervention period could mask or induce spurious effects unless controlled in the models. However, to rule out endogenous changes in composition, we provide estimates without

including these covariates in Appendix Table A4. All standard errors are clustered at the school-level to account for nonindependence between students within the same school.

When modeling test scores, we make several adjustments to Equation 1. First, we not only estimate Equation 1 for each subject (reading, math, and science) separately but also use a pooled sample of all subjects where we add subject indicators to control for any systematic variation between the three subjects. Second, we add a control for students' prior-year test score. We control for the prior-year achievement lag as our preferred model to facilitate comparisons with existing literature which primarily uses the prior-year lag. However, researchers have questioned the prior-year lag in this type of longitudinal model, arguing it may underestimate treatment effects because it adjusts for a score that may already be affected by turnaround (Sun et al., 2017). Therefore, we test two alternatives. First, we replace the prior-year lag with the average for each school-by-grade combination in all years prior to turnaround (i.e., a school-by-grade baseline lag). This alternative lag accounts for preintervention differences at the school-by-grade level, without controlling for scores that have been influenced by turnaround. We also replace the prior year test score with the average for each student in all years prior to attending a turnaround school (i.e., a student baseline lag). Models that use either the school-by-grade or student baseline lag lead to similar conclusions as our preferred prior-year lag (Appendix Table A5). Finally, for test scores outcomes, we add a grade fixed effect to ensure that results are not driven by unobserved differences across grades but excluding the grade fixed effect yields similar results (Appendix Table A4).

The coefficients of interest in Equation 1 are β_2 and β_3 for the ASD and iZones, respectively. The coefficients are interpreted as the pre–post turnaround difference for ASD or iZone schools minus the same pre–post difference for comparison schools. Equation 1 estimates the effects of ASD and iZone interventions under the identifying assumption that, conditional on covariates, comparison schools are a reasonable counterfactual for how ASD or iZone schools would have fared had they not begun turnaround. Below, we present validity and robustness checks to support this assumption.

To examine the ASD and iZone effects across time, we estimate event-study models that substitute the *After* variable with a set of six indicators for each year after schools begin turnaround. Interacting these six indicators with the ever-ASD and ever-iZone indicators allows us to estimate the cumulative effect of turnaround reforms in each year after schools join either model. Note that each cohort has a different number of years as part of a turnaround “treatment.” For example, the first cohort of ASD or iZone schools will have experienced six years of turnaround reforms between 2012–2013 and 2017–2018, and the second cohort

will have experienced 5 years of reforms between 2013–2014 and 2017–2018. For comparison schools, we center the *After* indicators on the year when they are designated as priority that is either 2012 or 2014. Despite missing test scores in 2015–2016, indicators for each postturnaround year allow us to estimate results for the fourth year of turnaround because Year 4 is 2016–2017 for cohort two and 2017–2018 for cohort three.

Finally, in a third specification, we separate the indicators for ever-ASD or ever-iZone into each of the five ASD and iZone cohorts. Interacting these cohort indicators with *After* yields effects for each cohort.¹⁷ When using the student mobility, teacher turnover, and principal turnover indicators as outcomes, our results are estimated using linear probability models. We present results from the linear probability models to ease interpretation, but estimates from logistic regression yield similar conclusions.

Below we also describe multiple robustness checks to rule out potential validity threats. For example, we test whether comparison schools are a valid counterfactual by examining whether test score trends in ASD, iZone, and comparison schools are parallel in the pret turnaround period. We also examine whether effects are driven by changes in student characteristics in these turnaround schools using a student fixed model to control for time-invariant student characteristics. Additionally, we test our models using an alternative comparison group and also examine whether our results are driven by mean reversion from potential dips in student performance prior to turnaround. Our checks provide evidence that our results are robust to alternative explanations.

Results

Descriptive Statistics

In Table 2, we present descriptive characteristics of students, teachers, and principals, showing that priority schools differ from nonpriority schools both before and after turnaround. Relative to nonpriority schools, priority schools tend to serve a larger population of Black students (87%–94%), FRPM eligible students (83%–92%), and students new to their school (28%–33%). Table 2 also shows that student characteristics are similar between ASD, iZone, and comparison schools before reforms began. When comparing baseline school characteristics before turnaround, we find only one significant difference in the proportion of minority students between iZone and comparison schools, but the magnitude of the difference (2 percentage points) is substantively small (Appendix Table A2). Moreover, turnaround reforms do not significantly change school compositional characteristics, suggesting that ASD or iZone effects are not driven by changes to the student population in these schools (Appendix Table A12).¹⁸

TABLE 2

Descriptive Characteristics of ASD Schools; iZone Schools; Non-ASD, Non-iZone Priority Schools; and All Other, Nonpriority Schools in Tennessee

| Characteristic | ASD | | iZone | | Non-ASD, non-iZone priority schools | | All other, nonpriority schools | |
|----------------------------------|--------|-------|--------|-------|-------------------------------------|-------|--------------------------------|-------|
| | Before | After | Before | After | Before | After | Before | After |
| Student characteristics | | | | | | | | |
| Female | .5 | .48 | .5 | .48 | .5 | .5 | .5 | .49 |
| Black | .92 | .93 | .87 | .87 | .94 | .89 | .2 | .21 |
| Hispanic | .057 | .049 | .066 | .082 | .048 | .054 | .06 | .085 |
| Asian | .0043 | .0035 | .0061 | .0052 | .003 | .0032 | .018 | .021 |
| White | .02 | .019 | .057 | .044 | .0078 | .012 | .72 | .68 |
| FRPM | .92 | .83 | .89 | .89 | .87 | .86 | .51 | .5 |
| ELL | .026 | .032 | .033 | .049 | .025 | .045 | .02 | .029 |
| SpED | .12 | .14 | .12 | .15 | .1 | .12 | .081 | .12 |
| Attendance rate | .94 | .92 | .93 | .92 | .94 | .93 | .95 | .95 |
| Proportion new to school | .33 | .31 | .28 | .29 | .31 | .31 | .15 | .15 |
| Teacher characteristics | | | | | | | | |
| Female | .79 | .82 | .77 | .78 | .76 | .77 | .79 | .78 |
| Black | .71 | .58 | .54 | .59 | .68 | .65 | .1 | .11 |
| White | .28 | .4 | .45 | .4 | .32 | .34 | .89 | .88 |
| Proportion new to school | .24 | .51 | .29 | .36 | .27 | .29 | .17 | .18 |
| TVAAS Score (1–5) | 2.5 | 2.7 | 2.5 | 3.1 | 2.6 | 2.8 | 3.3 | 3.2 |
| Years of experience | 12 | 5.2 | 11 | 8.7 | 11 | 10 | 13 | 12 |
| Master's degree or higher | .67 | .56 | .63 | .65 | .6 | .65 | .59 | .61 |
| Principal characteristics | | | | | | | | |
| Female | .65 | .8 | .56 | .6 | .64 | .56 | .55 | .57 |
| Black | .93 | .67 | .68 | .76 | .84 | .86 | .14 | .15 |
| White | .069 | .33 | .31 | .23 | .16 | .14 | .86 | .84 |
| Proportion new to school | .25 | .67 | .36 | .42 | .34 | .33 | .22 | .25 |
| Observation score (1–5) | 3.5 | 3.1 | 3.5 | 3.7 | 3.7 | 3.6 | 3.8 | 3.9 |
| Years of experience | 5.4 | 2.1 | 5 | 4.2 | 3.9 | 4.2 | 5.8 | 6.2 |
| Master's degree or higher | 1 | .76 | 1 | .99 | .92 | .93 | .98 | .98 |

Note. Attendance rate is the proportion of school days attended. The proportion of students new to a school do not include students who make structural moves. Principal years of experience only includes the years in which the individual was a principal, not all years of experience as an educator. The teacher TVAAS and principal observation scores range from 1 (*significantly below expectations*) to 5 (*significantly above expectations*). ASD = Achievement School District; iZone = Innovation Zone; FRPM = eligibility for free or reduced-priced meals; ELL = English language learner; SpED = special education status; TVAAS = Tennessee Value-Added Assessment System.

Table 2 also shows characteristics of teachers and principals in ASD and iZone schools. In ASD schools, both teacher and principal turnover increase after reforms begin. For example, the average proportion of new-to-school teachers rises from 24% to 51% after schools join the ASD. Part of this increased turnover is expected given staff replacement interventions in Year 1. However, increases in staff turnover are much less dramatic for iZone schools than ASD schools even though both models replaced teachers and principals in Year 1, suggesting that iZone schools were more successful at retaining teachers after the first year of intentional staff replacements.

There are also differences in the types of teachers and principals recruited to ASD and iZone schools. After schools

join the ASD, teachers are more likely to be White (increasing from 28% to 40%), have fewer years of experience (decreasing from 12 to 5.2 years), and are less likely to hold a graduate degree (decreasing from 67% to 56%). Average teacher experience also decreases after schools join an iZone (from 11 to 8.7 years). However, the decrease in teacher experience in iZone schools is smaller than in ASD schools, and average teacher TVAAS scores are noticeably higher after schools join an iZone, rising from an average of 2.5 to 3.1 out of 5. These trends in teacher experience and effectiveness are mirrored among principals. Additionally, we examine descriptive characteristics for each ASD and iZone cohort to better understand the types of schools chosen for either model over time. Appendix Table A6 shows that

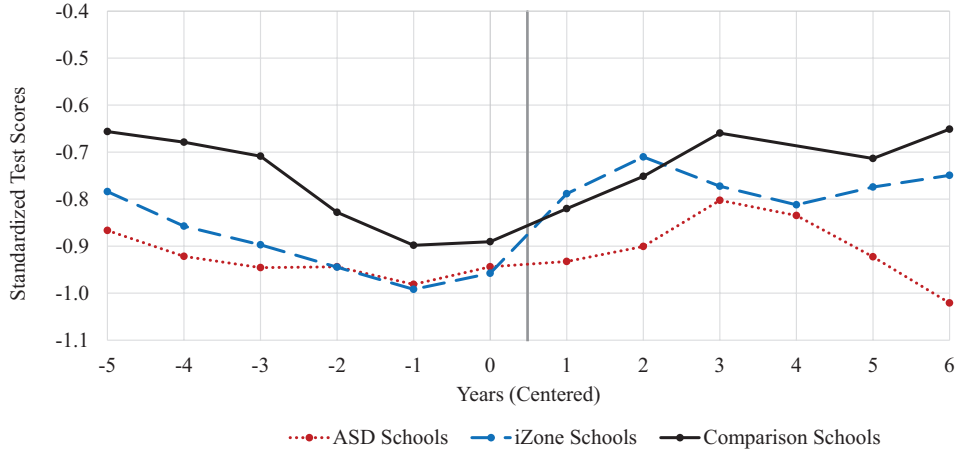


FIGURE 1. Average student test scores in each year before and after turnaround interventions began. Note. Years are centered such that zero is the year before school turnaround interventions began for ASD and iZone schools. For comparison schools, Year 0 is the year before these schools are designated as priority schools.

student characteristics are similar in each of the five ASD and iZone cohorts, suggesting that different outcomes across cohorts are not likely driven by changes in the types of schools chosen for turnaround.

Before turning to the DID results, Figure 1 shows descriptive trends in standardized test scores, averaged across all three subjects, for ASD, iZone, and comparison schools. The trends are shown across a year variable centered at zero on the baseline year before turnaround. For comparison schools, Year 0 is the year before their priority designation. The figure shows that, prior to turnaround, student achievement is slightly lower in ASD and iZone schools relative to comparison schools; however, the trends in student achievement are qualitatively similar with no sharp spikes in either ASD or iZone schools that differ from comparison schools. Moreover, no crossovers exist between turnaround and comparison schools prior to reforms. Student test scores in all schools exhibit some declines in the years before turnaround, but these downward trends are similar in ASD, iZone, and comparison schools. This figure provides descriptive evidence that student achievement trends in comparison schools provide a valid counterfactual for what would have occurred in turnaround schools had they not joined either the ASD or an iZone. To provide further evidence supporting the parallel trends assumption, we estimate a model that includes separate indicators for each year in our panel and interact them with the ASD and iZone indicators. This model examines student achievement in each pret turnaround year relative to the baseline year before turnaround (Appendix Table A7). Results show no significant difference in student achievement between either turnaround model and comparison schools in any pret turnaround year. Moreover, joint F tests of all pret turnaround years show no evidence of significant differences prior to turnaround.

Figure 1 also shows average standardized student achievement results after turnaround reforms are implemented, suggesting that test scores increase in all schools. Increases in all schools indicate that the priority designation alone may have spurred some improvements in these schools, because comparison schools received no other interventions. However, although all comparison schools show improvement, the increase in iZone schools surpasses comparison schools in the first 2 years. Then, the iZone gains diminish after Year 2, but they do not completely vanish, and average iZone scores remain higher than in the ASD. In contrast, test scores in ASD schools increase somewhat in the first 3 years, then decrease again, never catching up to either iZone or comparison schools and, by Year 6, average test scores in ASD schools are slightly lower than in any pret turnaround year.

DID Results

Turning to the DID results, Table 3 shows average ASD and iZone effects on standardized test scores averaged across all years and cohorts for all subjects and each subject (reading, math, and science). The table also shows separate models including either the prior-year lag or the school-by-grade baseline lag. Consistent with previous literature, we find that including the baseline lag yields somewhat larger effects than the prior-year lag, because turnaround reforms may have affected the prior-year lag. However, we prefer the prior-year lag, because students' most recent previous performance best accounts for unobserved factors that affect their current performance. Therefore, our preferred models should be interpreted as somewhat conservative. Hereafter, we focus on models using the prior-year lag; estimates using the baseline lag are available on request.

TABLE 3

ASD and iZone Effects on Standardized Test Scores Averaged Across All Years and Cohorts

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|-------------------------------|------------------|-------------------|-----------------|-------------------|------------------|-------------------|-----------------|------------------|
| | All Subjects | All Subjects | Reading | Reading | Math | Math | Science | Science |
| ASD | .0011 (.044) | .0022 (.031) | .03 (.047) | .0088 (.032) | -.015 (.054) | -.0095 (.044) | -.011 (.055) | .013 (.043) |
| iZone | .13** (.047) | .096** (.035) | .12** (.044) | .055 (.029) | .12* (.051) | .096* (.043) | .18** (.063) | .15** (.047) |
| School-by-grade baseline lag | .41*** (.037) | | .25*** (.05) | | .31*** (.056) | | .3*** (.063) | |
| Prior-year lag | | .62*** (.0043) | | .67*** (.0041) | | .57*** (.0082) | | .61*** (.005) |
| Student and school covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| School fixed effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Number of schools | 116 | 116 | 116 | 116 | 116 | 116 | 116 | 116 |
| Observations | 455,209 | 315,663 | 163,479 | 11,4299 | 153,253 | 105,023 | 138,477 | 96,341 |
| Adjusted R^2 | .110 | .416 | .127 | .478 | .094 | .369 | .108 | .403 |

Note. Standard errors in parentheses. *ASD* and *iZone* represent the interaction between indicators for schools that are ever-ASD or ever-iZone and an indicator for years after schools begin turnaround. Comparison group schools are all priority schools not receiving any turnaround interventions on either the 2012 or 2014 priority list. ASD = Achievement School District; iZone = Innovation Zone.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 3 shows that the average iZone effect is positive and statistically significant. Using a prior-year lag, the iZone effect is approximately 0.10 *SD*, which holds across subjects (though the coefficient is marginally significant for reading, $p = 0.06$). Similar to previous school evaluation studies, our estimates are somewhat larger in math (0.10 *SD*) than in reading (0.06 *SD*). The effect estimate is largest in science (0.15 *SD*). Since science test scores are not part of Tennessee's school accountability system, large gains in science suggest that positive iZone effects are not driven by a narrowed focus on subject-specific test preparation. For ASD schools, the coefficients are substantively small and statistically insignificant in all subjects.

To examine effects over time, Table 4 shows estimates from our event-study model. In ASD schools, we find no significant effects in any year or subject. In iZone schools, we find positive and significant effects concentrated in the first 2 years. For example, the cumulative effect of iZone reforms after 2 years is 0.15 *SD*. Positive effects of iZone reforms in the first 2 years are similar across reading, math, and science, except the estimate for reading in Year 2 is marginally significant ($p = 0.06$). However, the positive iZone effects diminish and become statistically insignificant in Years 3 and 4, suggesting that early gains were not sustained in these years. Then, for Years 5 and 6, the estimates for iZone schools increase but are not statistically significant, except the 6th-year estimate in reading (0.13 *SD*). The pattern across multiple years suggest that moderate to large positive effects of iZone interventions diminish after the first 2 years, but the coefficients in Years 5 and 6

are suggestive of potential positive effects, which may be insignificant due to low statistical power. There is less power to detect significant effects in Years 5 and 6, because only the first two iZone cohorts have received 5 years of reforms, and only the first cohort has undergone 6 years of reforms.

Table 5 examines separate effects for each cohort. For ASD schools, again we find no evidence of positive effects and even observe a 0.08 *SD* decrease in Cohort 2, driven mostly by math scores. For iZones, the overall positive effect is driven by the first two cohorts. The second and third cohorts produce mostly null effects, except a positive effect of 0.13 *SD* in science in Cohort 3. Finally, averaging all subjects, we find noticeable negative effects in the fifth iZone cohort of about 0.19 *SD*. These negative results are driven by scores in reading and math. Below, we discuss potential reasons for negative effects in the fifth iZone cohort.

Adding to our examination of test scores, Table 6 examines additional outcomes for students, teachers, and principals. We find that ASD and iZone reforms do not have a significant effect on student attendance or mobility, but teachers have a 23% increase in the probability of being new to the school after schools join the ASD. Part of the increased teacher turnover is explained by intentional staff replacements in Year 1, but below, we show that teacher turnover in the ASD continued to be high after Year 1. The coefficient for teacher turnover in iZone schools is not statistically significant. When we examine the effectiveness of incoming teachers, the coefficient is not significant for either model.

TABLE 4
ASD and iZones Effects on Test Scores by Turnaround Year

| | (1) | (2) | (3) | (4) |
|-------------------------------|-------------------|-------------------|-------------------|-------------------|
| | All Subjects | Reading | Math | Science |
| ASD | | | | |
| ASD Year 1 | -.037 (.064) | .017 (.046) | -.1 (.11) | -.033 (.097) |
| ASD Year 2 | .023 (.056) | .015 (.046) | .074 (.069) | -.026 (.081) |
| ASD Year 3 | .017 (.051) | .02 (.037) | -.015 (.078) | .064 (.073) |
| ASD Year 4 | -.056 (.055) | -.044 (.058) | -.07 (.073) | -.095 (.096) |
| ASD Year 5 | -.04 (.092) | -.038 (.054) | -.14 (.1) | -.0022 (.22) |
| ASD Year 6 | -.084 (.1) | -.056 (.082) | -.14 (.09) | -.097 (.17) |
| iZone | | | | |
| iZone Year 1 | .11* (.042) | .068* (.033) | .13* (.051) | .15* (.064) |
| iZone Year 2 | .15** (.044) | .071 (.038) | .21*** (.056) | .18** (.06) |
| iZone Year 3 | .027 (.045) | .013 (.043) | -.035 (.061) | .12 (.062) |
| iZone Year 4 | -.059 (.1) | -.087 (.094) | -.15 (.15) | .092 (.095) |
| iZone Year 5 | .11 (.07) | .1 (.054) | .051 (.082) | .097 (.11) |
| iZone Year 6 | .14 (.07) | .13** (.044) | .16 (.085) | .11 (.1) |
| Prior year lag | .62*** (.0043) | .67*** (.0042) | .57*** (.0082) | .61*** (.0049) |
| Student and school covariates | Yes | Yes | Yes | Yes |
| School fixed effect | Yes | Yes | Yes | Yes |
| Number of schools | 116 | 116 | 116 | 116 |
| Observations | 315,663 | 114,299 | 105,023 | 96,341 |
| Adjusted R^2 | .417 | .479 | .373 | .404 |

Note. Standard errors in parentheses. Comparison group schools are all priority schools not receiving any turnaround interventions on either the 2012 or 2014 priority list. ASD = Achievement School District; iZone = Innovation Zone.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Moreover, ASD teachers have fewer years of experience (4.9 fewer years than comparison school teachers) compared with iZone teachers (1.5 fewer years than comparison school teachers).

ASD principals have a 47% greater probability of being new to their school compared with comparison school principals. Similar to teachers, part of the elevated principal turnover in the ASD occurs because of mandated replacement policies in Year 1, but principal turnover remains high after the first year (Appendix Table A8). Average principal observation scores do not change

significantly in either ASD or iZone schools, but average principal experience decreases in ASD schools (about 3.4 years).

To conserve space, effect estimates for these student, teacher, and principal outcomes separated by year are in Appendix Table A8, but we note that ASD schools experience elevated teacher and principal turnover across the 6 years, whereas teacher turnover increases only in Year 1 for the iZones. To provide more context for ASD and iZone effects across cohorts, Table 7 shows all outcomes separated by cohort. Again, there is little evidence that either ASD or

TABLE 5
ASD and iZone Effects on Test Scores by Cohort

| | (1) | (2) | (3) | (4) |
|-------------------------------|-------------------|-------------------|-------------------|------------------|
| | All Subjects | Reading | Math | Science |
| ASD | | | | |
| ASD Cohort 1 | .043 (.052) | .008 (.052) | .023 (.063) | .092 (.068) |
| ASD Cohort 2 | -.082* (.041) | -.03 (.041) | -.16** (.053) | -.07 (.063) |
| ASD Cohort 3 | .0058 (.046) | .044 (.052) | .058 (.072) | -.075 (.088) |
| ASD Cohort 4 | .055 (.045) | .026 (.031) | .11 (.085) | .028 (.049) |
| ASD Cohort 5 | -.04 (.065) | -.06 (.033) | -.075 (.1) | .036 (.074) |
| iZone | | | | |
| iZone Cohort 1 | .16*** (.046) | .11** (.04) | .19*** (.051) | .19** (.064) |
| iZone Cohort 2 | .15** (.055) | .13** (.04) | .16** (.055) | .18* (.079) |
| iZone Cohort 3 | -.0021 (.033) | .0041 (.025) | -.079 (.081) | .13* (.058) |
| iZone Cohort 4 | -.048 (.056) | -.086 (.053) | -.029 (.08) | -.034 (.081) |
| iZone Cohort 5 | -.19** (.065) | -.26*** (.03) | -.29* (.13) | .075 (.12) |
| Prior-year lag | .62*** (.0043) | .67*** (.0041) | .57*** (.0082) | .61*** (.005) |
| Student and school covariates | Yes | Yes | Yes | Yes |
| School fixed effect | Yes | Yes | Yes | Yes |
| Number of schools | 116 | 116 | 116 | 116 |
| Observations | 315,663 | 114,299 | 105,023 | 96,341 |
| Adjusted R^2 | .417 | .479 | .371 | .404 |

Note. Standard errors in parentheses. Comparison group schools are all priority schools not receiving any turnaround interventions on either the 2012 or 2014 priority list. ASD = Achievement School District; iZone = Innovation Zone.
 * $p < .05$. ** $p < .01$. *** $p < .001$.

iZone schools have significant effects on student attendance or mobility, except students in the third ASD cohort have a lower probability of being new to their school, and the fifth ASD cohort has lower attendance rates.

Table 7 shows that the probability of a teacher being new to her school is increased in every ASD cohort (20%–38%), though the effect is marginally significant in cohort four. Likewise, the probability of a principal being new to her school is positive and significant in every ASD cohort (36%–70%). Moreover, average teacher and principal experience is lower than in comparison schools for every ASD cohort.

For iZones, the probability that a teacher is new to her school is also positive (8–15 percentage points) but smaller

in magnitude than in ASD schools. The effect of iZone reforms on teacher turnover is highest in Cohort 5 (15%). Besides teacher turnover, the fifth iZone cohort differs in other ways from the first four cohorts. Only in the fifth cohort do we observe a negative effect on prior-year TVAAS scores among incoming teachers, suggesting that the fifth iZone cohort is recruiting less-effective teachers than comparison schools. Moreover, the probability that a principal is new to her school is only significant in Cohort 5, and only in the fifth cohort do we find a negative effect on principal experience in iZone schools. Together, these results show that the fifth iZone cohort experienced higher staff turnover with incoming teachers who are less effective and principals who are less experienced.

TABLE 6

ASD and iZone Effects on Test Scores and Alternative Outcomes Across All Years and Cohorts

| | Student level | | | Teacher level | | | Principal level | | |
|----------------------------------|------------------|-------------------|-----------------------------|-----------------------------|--|-----------------------------|-------------------------------|-----------------------------------|-------------------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| | Test scores | Attendance rate | Student new to school (LPM) | Teacher new to school (LPM) | Prior year TVAAS for incoming teachers | Teacher years of experience | Principal new to school (LPM) | Principal observation score (1–5) | Principal years of experience |
| ASD | .0022 (.031) | -.011 (.0057) | -.045 (.027) | .23*** (.036) | -.2 (.29) | -4.9*** (.67) | .47*** (.092) | .011 (.11) | -3.4*** (.65) |
| iZone | .096** (.035) | -.00098 (.006) | -.011 (.019) | .055 (.029) | .049 (.21) | -1.5** (.48) | .11 (.097) | .1 (.11) | -1.1 (.69) |
| Individual and school covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| School fixed effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 315,663 | 272,112 | 272,202 | 26,573 | 1,259 | 26,281 | 753 | 667 | 753 |
| Adjusted R^2 | .416 | .013 | .014 | .018 | .011 | .073 | .056 | .034 | .100 |

Note. Standard errors in parentheses. Comparison group schools are all priority schools not receiving any turnaround interventions on either the 2012 or 2014 priority list. LPM denotes linear probability models for a dichotomous outcome. Student new to school does not include structural moves where students transfer to a new school level after completing the final grade offered at her previous school. The TVAAS index is a standardized, continuous value-added measure of effectiveness calculated only for test subject teachers. Principal observation scores range from 1 (*significantly below expectations*) to 5 (*significantly above expectations*). ASD = Achievement School District; iZone = Innovation Zones; TVAAS = Tennessee Value-Added Assessment System. * $p < .05$. ** $p < .01$. *** $p < .001$.

Alternative Specifications and Robustness Checks

To test the robustness of our DID estimates, we first use an alternative specification where we model the preturnaround trend using a comparative interrupted time series (CITS) model that includes a linear time trend interacted with ASD and iZone indicators (Appendix Table A9). The CITS results yield conclusions that are similar to our DID estimates. For example, the CITS model finds that, relative to comparison schools, the postturnaround deviation from the baseline trend for iZone schools is 0.11 *SD*. We do not use the CITS models as our main specification, because it estimates deviations from the baseline (pretturnaround) trend, and the baseline trend may not be an accurate representation of how these schools are performing after 6 years of reform when student achievement trends in these schools may have changed dramatically.

Next, we examine potential validity threats. First, we test whether results are driven by a dip in performance among turnaround schools in the year immediately prior to reforms. This pretturnaround dip may inflate our results because of mean reversion (Ashenfelter, 1978) and could occur if schools are discouraged by impending reforms in the following year. We examine this possibility using a placebo test that shifts the treatment year to the year prior to reforms and find that the estimates are not significant, providing evidence that no dip in performance exists in the baseline year (Appendix Table A10).

Then, we test whether our results are driven by changes in the student population in ASD or iZone schools by using both a student fixed effect model and testing whether the turnaround reforms significantly affected school compositional characteristics: enrollment and whether students are female, minority race, FRPM-eligible, ELL, and SpED. Effect estimates from the student fixed effects model produce similar conclusions (Appendix Table A11), and we find no evidence of changes in student characteristics after schools join either the ASD or an iZone (Appendix Table A12). Next, we test whether our results are robust to our choice of comparison group by using nonpriority schools that perform among the lowest 6% to 10% of schools in Tennessee. Our conclusions do not change when using this alternative comparison group (Appendix Table A13). Additionally, we restrict our comparison group to only schools designated as priority in 2012, instead of using both the 2012 and 2014 priority lists. Again, our conclusions do not change (Appendix Table A3). We also test whether our results are driven by students differentially leaving the test taking pool, by examining ASD and iZone effects on an indicator for missing test scores. We find that neither turnaround model significantly affected test score missingness, providing confidence that our results are not biased by differential selection of students into the sample (Appendix Table A14).

Finally, student movement between schools may bias our results. The potential bias may be concerning if

TABLE 7
ASD and iZone Effects on Test Scores and Alternative Outcomes By Cohort

| | Student Level | | | Teacher Level | | | Principal Level | | |
|----------------------------------|------------------|-------------------|-----------------------------|-----------------------------|--|-----------------------------|-------------------------------|-----------------------------------|-------------------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| | Test scores | Attendance rate | Student new to school (LPM) | Teacher new to school (LPM) | Prior year TVAAS for incoming teachers | Teacher years of experience | Principal new to school (LPM) | Principal observation score (1–5) | Principal years of experience |
| ASD | | | | | | | | | |
| ASD Cohort 1 | .043 (.052) | -.014 (.0079) | -.093 (.066) | .22*** (.033) | -.13 (.38) | -5.9*** (.46) | .7*** (.11) | -.23 (.16) | -2.5*** (.67) |
| ASD Cohort 2 | -.082* (.041) | -.0031 (.0062) | -.052 (.04) | .23*** (.056) | -.59 (.5) | -7.2*** (.72) | .36* (.16) | -.0032 (.16) | -4** (1.4) |
| ASD Cohort 3 | .0058 (.046) | -.014 (.0087) | -.065** (.02) | .2*** (.035) | .027 (.22) | -5.2*** (.5) | .38* (.16) | .19 (.2) | -2* (.97) |
| ASD Cohort 4 | .055 (.045) | -.005 (.0099) | .015 (.022) | .16 (.089) | -.62* (.31) | -4.1** (1.3) | .45*** (.088) | .25 (.24) | -4.6*** (1) |
| ASD Cohort 5 | -.04 (.065) | -.027* (.012) | .031 (.032) | .38*** (.043) | -.61* (.29) | -2.3* (1.1) | .46** (.15) | -.2 (.22) | -4.1*** (.94) |
| iZone | | | | | | | | | |
| iZone Cohort 1 | .16*** (.046) | -.0058 (.0062) | .00093 (.031) | .032 (.048) | .27 (.29) | -1.6 (.87) | -.048 (.14) | .45* (.18) | -1.5 (1.1) |
| iZone Cohort 2 | .15** (.055) | .015 (.0096) | -.02 (.027) | .079* (.036) | .073 (.22) | -2*** (.56) | .046 (.13) | -.0036 (.13) | 1 (.99) |
| iZone Cohort 3 | -.0021 (.033) | -.02 (.012) | -.024 (.025) | .093* (.04) | -.35 (.29) | -1.9*** (.52) | .52 (.35) | .059 (.25) | -2.6 (1.7) |
| iZone Cohort 4 | -.048 (.056) | -.0048 (.0051) | -.02 (.022) | .0096 (.06) | .25 (.25) | -.78 (.8) | .16 (.14) | .038 (.16) | -2.1 (1) |
| Zone Cohort 5 | -.19** (.065) | .014 (.012) | .03 (.024) | .15** (.045) | -.6** (.21) | -.49 (.84) | .55* (.21) | -.33 (.41) | -4.1* (2) |
| Individual and school covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| School fixed effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 315,663 | 272,112 | 272,202 | 26,573 | 1,259 | 26,281 | 753 | 667 | 753 |
| Adjusted R ² | .417 | .014 | .015 | .019 | .015 | .074 | .064 | .046 | .125 |

Note. Standard errors in parentheses. LPM denotes linear probability models for a dichotomous outcome. Student new to school does not include structural moves. The TVAAS index is a standardized, continuous value-added measure of effectiveness calculated only for test subject teachers. Principal observation scores range from 1 (*significantly below expectations*) to 5 (*significantly above expectations*). ASD = Achievement School District; iZone = Innovation Zone; TVAAS = Tennessee Value-Added Assessment System.

* $p < .05$. ** $p < .01$. *** $p < .001$.

schools retain higher-scoring students and push out lower-scoring students (Zimmer & Guarino, 2013). This issue is unlikely, because ASD and iZone schools were required to continue enrolling students from their neighborhood catchment areas. Nevertheless, we examine average student test scores among students who transfer into and out of schools in our sample (Appendix Table A15). We find no evidence of significant differences between incoming and outgoing students.¹⁹

Student movement could also bias results if they move between turnaround and nonturnaround schools. For example, if students move from a turnaround school to a comparison school, the turnaround effects may linger after they transfer. To examine this issue, we remove all observations where students move (1) from a turnaround to a comparison school, (2) from a comparison to a turnaround school, or (3) across the two turnaround models. Only about 4% of the students in our sample make these

transfers. These sample restrictions yield results that are similar to our main estimates (Appendix Table A16). Together, results from our checks support the robustness of our results.

Conclusions

This article examines the effects of Tennessee's ASD and iZones after 6 years, contributing to the sparse literature on turnaround reforms that have evolved beyond the initial disruption stage into the later stages of capacity-building and sustaining the momentum. We use newly updated data to expand on a prior evaluation by Zimmer et al. (2017), because the effects reported by Zimmer and colleagues may have not allowed enough time for either model to fully mature, nor did they allow sufficient time to examine the sustainability of iZone effects. We find that ASD interventions did not result in better or worse performance than comparison schools, whereas positive iZone effects are primarily observed in the first two cohorts.

Since ASD schools have not produced significant effects on student achievement in any subject, year, or cohort, we conclude that the ASD model is unlikely to be successful in the future. Researchers have documented structural problems with this original model including alienation of local communities around these schools (Glazer & Egan, 2018), and TDOE leaders have recently announced plans to release ASD schools back to their LEAs by 2022 (Bauman & Aldrich, 2020).

To help explain null ASD effects, we find high rates of teacher and principal turnover in ASD schools after Year 1 when staff replacement was no longer required. Organizational instability from high staff turnover may partly explain why ASD schools could not produce positive effects. This result aligns with prior literature which finds that high staff turnover is negatively associated with student achievement (Henry & Redding, 2020; Ronfeldt et al., 2013) and a prior study finding that staff turnover suppresses potential positive effects of school reform (Henry et al., 2020). Our results also suggest that the characteristics of teachers and principals recruited to ASD schools may help explain insignificant effects. Teachers recruited to ASD schools were less effective and less experienced than teachers in comparison schools. Moreover, ASD principals have about 3.4 fewer years of experience than principals in comparison schools. Moving forward under ESSA, these results highlight the importance of experienced and effective educators on the success of school reform.

The iZone schools may have more easily recruited effective educators partly because they offered pay bonuses. Previous studies have shown that pay incentives can have a positive effect on teacher recruitment (Springer et al., 2016), suggesting that pay incentives can be a useful tool for policymakers to consider in order to bring effective educators into turnaround schools. Moreover, ASD

schools may have had difficulty recruiting educators because the state-led model failed to gain trust from local communities. Previous work showed that the ASD faced community backlash when it began operating schools because critics saw the model as an attempt to remove schools from their local community (Glazer & Egan, 2018). Community mistrust may have dissuaded educators from coming to ASD schools, which suggests that policymakers thinking about state-led turnaround models must carefully attend to building local support.

For iZone schools, we find that positive effects are largely driven by the first two cohorts. Indeed, our effect estimates, averaging all cohorts together, are smaller than those reported by Zimmer et al. (2017). We estimate effect sizes (in *SD*) of 0.06 in reading, 0.10 in math, and 0.15 in science, whereas Zimmer et al. (2017) report effect sizes of 0.10 in reading, 0.20 in math, and 0.18 in science. The diminished effects in later cohorts suggest that the iZones may be having difficulty sustaining initial positive results.

To better understand effects in later iZone cohorts, we compare characteristics of students in each cohort and find them to be similar, suggesting that diminished effects in later cohorts are not driven by different types of schools joining the iZones. However, one potential reason for the diminished effects may be the inclusion of high schools. The iZones began taking over high schools starting with the third cohort, and the fifth iZone cohort consists entirely of high schools. This timing suggests that iZone interventions may be less effective in high schools. We test this potential explanation by interacting a high school indicator with the ASD and iZone indicators and find that this is indeed the case (Appendix Table A17). We urge future research to examine specific iZone interventions to better understand why they have not been as effective in high schools.

The negative effect of iZone schools in the fifth cohort is also noticeable, and our examination of teacher and principal outcomes suggest potential reasons to explain these negative effects. Relative to earlier cohorts, cohort five schools experienced the highest turnover rates, and educators in cohort five schools were less effective and less experienced prior to joining the iZone. These trends differ markedly from the first iZone cohort where teachers and principals were more effective than their counterparts in comparison schools. Comparing educator effectiveness and turnover across cohorts suggests that the local pool of effective educators may have diminished as more schools joined the iZone. This may have left too few effective educators available to staff schools in later cohorts, leading to less positive student achievement results.

Our findings have important implications for longer-term reform models under ESSA. Policymakers considering turnaround initiatives that rely on recruiting effective staff to low-performing schools would do well to focus on ways to continue expanding the local pool of effective teachers and leaders, perhaps by recruiting from broader geographic areas

or building more robust in-service professional development programs. Developing more effective teachers is especially important given previous research finding that recruiting effective teachers to iZone schools negatively affected the schools that lost these teachers (Kho et al., 2018). To sustain turnaround supports for low-performing schools without negatively affecting other schools, educational authorities should pay particular attention to developing teachers and principals in low-performing schools.

In both turnaround models, we find no notable effects on student mobility or attendance. Student mobility is a salient issue in both ASD and iZone schools, because the proportion of students making nonstructural moves into these schools (30%) is approximately double the rates in nonpriority schools (15%). Given research that finds that school mobility can negatively affect student achievement (Friedman-Krauss & Raver, 2015), reform initiatives like the ASD and iZones

should consider more actively addressing high-student mobility rates. Although average student attendance rates in ASD and iZone schools are only somewhat lower than in nonpriority schools (92% vs. 95%), low-student attendance rates can be a barrier to improvement if students are not in class to receive instruction. Future research should specifically examine how school reform models improve outcomes for chronically absent students, because they will likely need the most attendance support.

Overall, our results show a need for more evidence to elucidate school reforms in the later stages of capacity-building and sustaining the momentum. Even models that produce positive gains in early years may face different obstacles that require an evolving approach to supporting low-performing schools over time. We urge policymakers and educational authorities to carefully consider these lessons from Tennessee as they implement new reform models under ESSA.

Appendix

TABLE A1

Results After Restricting to Only Priority Schools in Memphis

| | (1) | (2) | (3) |
|-------------------------------|-----------------|------------------|-------------------|
| | Reading | Math | Science |
| ASD | .030 (.028) | .009 (.042) | .014 (.042) |
| iZone | .087* (.034) | .140** (.047) | .229*** (.044) |
| Student and school covariates | Yes | Yes | Yes |
| School fixed effect | Yes | Yes | Yes |
| Number of schools | 86 | 86 | 86 |
| Observations | 1,68,907 | 1,51,245 | 1,41,303 |
| Adjusted R^2 | .433 | .370 | .358 |

Note. Standard errors in parentheses. ASD = state-led Achievement School District; iZone = district-led local Innovation Zones.

* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE A2

Comparisons of the Two Treatment Groups (ASD and iZone) and the Comparison Group (All Other Priority Schools)

| | ASD | | iZone | |
|--------------------------------|----------------------------|--------------|----------------------------|--------------|
| | Non-ASD non-iZone priority | Ever treated | Non-ASD non-iZone priority | Ever treated |
| Student test scores | | | | |
| Average reading score | -0.943 | -1 | -0.943 | -0.988 |
| Average math score | -0.944 | -0.92 | -0.944 | -0.938 |
| Average science score | -1.1 | -1.15 | -1.1 | -1.14 |
| Student characteristics | | | | |
| Proportion female | .488 | .493 | .488 | .486 |
| Proportion minority | .988 | .985 | .988 | .967* |
| Proportion FRPM | .881 | .886 | .881 | .921 |
| Proportion SpED | .15 | .153 | .15 | .172 |
| Proportion ELL | .0261 | .0166 | .0261 | .0208 |

Note. ASD = Achievement School District; iZone = Innovation Zone; FRPM = eligibility for free or reduced priced meals; ELL = English language learner; SpED = special education status.

* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE A3
Results Using Only Schools on the 2012 Priority List

| | (1) | (2) | (3) |
|-------------------------------|-----------------|------------------|------------------|
| | Reading | Math | Science |
| ASD | .031 (.029) | .004 (.042) | -.020 (.039) |
| iZone | .067* (.029) | .118** (.042) | .125** (.042) |
| Student and school covariates | Yes | Yes | Yes |
| School fixed effect | Yes | Yes | Yes |
| Number of schools | 83 | 83 | 83 |
| Observations | 166,410 | 152,425 | 143,929 |
| Adjusted R^2 | .439 | .381 | .366 |

Note. Standard errors in parentheses. ASD = Achievement School District; iZone = Innovation Zone.
* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE A4
ASD and iZones Effects on All Outcomes: No Covariates

| | Student Level | | | Teacher Level | | | Principal Level | | | |
|----------------------------------|-----------------|-----------------------------------|-------------------|-----------------------------|-----------------------------|--|-----------------------------|-------------------------------|-----------------------------------|-------------------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| | Test scores | Test scores—no grade fixed effect | Attendance rate | Student new to school (LPM) | Teacher new to school (LPM) | Prior year TVAAS index for incoming teachers | Teacher years of experience | Principal new to school (LPM) | Principal observation score (1–5) | Principal years of experience |
| ASD | -.021 (.03) | -.02 (.03) | -.008 (.0058) | -.049 (.029) | .26*** (.038) | -.21 (.27) | -6*** (.75) | .46*** (.083) | .012 (.11) | -3.5*** (.64) |
| iZone | .079* (.036) | .076* (.037) | -.00042 (.006) | -.015 (.02) | .048 (.029) | .02 (.21) | -1.6** (.53) | .071 (.093) | .12 (.11) | -.97 (.7) |
| Individual and school covariates | No | No | No | No | No | No | No | No | No | No |
| Grade fixed effect | Yes | No | Yes | Yes | No | No | No | No | No | No |
| School fixed effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 315,886 | 316,064 | 272,571 | 272,661 | 27,268 | 1,274 | 26,966 | 779 | 693 | 779 |
| Adjusted R^2 | .405 | .403 | .004 | 0.000 | .013 | .006 | .021 | .054 | .001 | .093 |

Note. Comparison group schools are all priority schools not receiving any turnaround interventions on either the 2012 or 2014 priority list. LPM denotes linear probability models for a dichotomous outcome. ASD = Achievement School District; iZone = Innovation Zone; TVAAS = Tennessee Value-Added Assessment System.
* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE A5
ASD and iZones Effects on Test Scores Averaged Across All Years and Cohorts Using Various Lags

| | (1) | (2) | (3) | (4) |
|-------|----------------|------------------------------|----------------------|----------------|
| | No lag | School-by-grade baseline lag | Student baseline lag | Prior year lag |
| ASD | .01 (.04) | .00 (.04) | .03 (.04) | .00 (.03) |
| iZone | .14** (.05) | .13** (.05) | .14** (.04) | .10** (.03) |

(continued)

TABLE A5 (CONTINUED)

| | (1) | (2) | (3) | (4) |
|-------------------------------|---------|------------------------------|----------------------|-----------------|
| | No lag | School-by-grade baseline lag | Student baseline lag | Prior year lag |
| School-by-grade baseline lag | | .41*** (.04) | | |
| Student level baseline lag | | | .83*** (.01) | |
| Prior year lag | | | | .62*** (.00) |
| Student and school covariates | Yes | Yes | Yes | Yes |
| School fixed effect | Yes | Yes | Yes | Yes |
| Number of schools | 116 | 116 | 116 | 116 |
| Observations | 467,407 | 455,209 | 329,249 | 315,663 |
| Adjusted R ² | .107 | .110 | .527 | .416 |

Note. Comparison group schools are all priority schools not receiving any turnaround interventions on either the 2012 or 2014 priority list. ASD = Achievement School District; iZone = Innovation Zone.

* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE A6

Descriptive Characteristics by Cohort in the Postturnaround Period

| | ASD | | | | | iZone | | | | |
|---------------------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| | Cohort 1 | Cohort 2 | Cohort 3 | Cohort 4 | Cohort 5 | Cohort 1 | Cohort 2 | Cohort 3 | Cohort 4 | Cohort 5 |
| Student characteristics | | | | | | | | | | |
| Female | .48 | .5 | .49 | .46 | .47 | .48 | .49 | .49 | .43 | .47 |
| Black | .94 | .97 | .96 | .85 | .93 | .99 | .81 | .98 | .98 | .97 |
| Hispanic | .026 | .015 | .03 | .09 | .06 | .0083 | .16 | .011 | .0042 | .03 |
| Asian | .0037 | 0 | .0034 | .011 | .00092 | .00012 | .0086 | .00034 | 0 | .0005 |
| White | .024 | .01 | .0078 | .052 | .0069 | .0041 | .022 | .0038 | .013 | .0025 |
| FRPM | .9 | .9 | .79 | .73 | .7 | .93 | .89 | .92 | .92 | .88 |
| ELL | .021 | .01 | .021 | .01 | .045 | .0035 | .099 | .0033 | .0042 | .014 |
| SpED | .15 | .13 | .14 | .12 | .15 | .14 | .12 | .19 | .1 | .18 |
| Attendance rate | .93 | .92 | .91 | .94 | .91 | .94 | .94 | .89 | .93 | .91 |
| Proportion new to school | .3 | .34 | .29 | .31 | .38 | .36 | .32 | .33 | .38 | .23 |
| Teacher characteristics | | | | | | | | | | |
| Female | .82 | .84 | .75 | .85 | .8 | .82 | .8 | .65 | .82 | .62 |
| Black | .61 | .58 | .52 | .51 | .72 | .8 | .73 | .76 | .73 | .81 |
| White | .38 | .37 | .46 | .47 | .27 | .19 | .25 | .22 | .27 | .17 |
| Proportion New to school | .53 | .46 | .45 | .52 | .64 | .36 | .35 | .4 | .43 | .36 |
| TVAAS score (1–5) | 3.2 | 2.7 | 2.4 | 2.3 | 2.1 | 3.4 | 3.5 | 2.8 | 2.8 | 3 |
| Years of experience | 5.1 | 3.7 | 5.6 | 6 | 6 | 9.9 | 9.1 | 9.1 | 8.7 | 11 |
| Master's degree or higher | .55 | .6 | .49 | .63 | .51 | .7 | .68 | .61 | .5 | .62 |
| Principal characteristics | | | | | | | | | | |
| Female | .88 | .88 | .43 | .91 | .88 | .67 | .52 | .43 | 1 | .5 |
| Black | .64 | .79 | .57 | .55 | .75 | .95 | .97 | .88 | 1 | 1 |
| White | .36 | .21 | .43 | .45 | .25 | .05 | .034 | .13 | 0 | 0 |
| Proportion new to school | .79 | .56 | .67 | .73 | .63 | .3 | .47 | .81 | .33 | .33 |
| Years of experience | 1.9 | 1.8 | 2.9 | 2.4 | 1.4 | 4.4 | 5.5 | 3.4 | 11 | 4.8 |
| Master's degree or higher | .83 | .8 | .47 | .91 | .75 | 1 | 1 | 1 | 1 | 1 |

Note. Attendance rate is the proportion of school days attended. The proportion of students new to a school do not include students who make structural moves. Principal years of experience only includes the years in which the individual was a principal not all years of experience as an educator. The teacher TVAAS and principal observation scores range from 1 (*significantly below expectations*) to 5 (*significantly above expectations*). ASD = Achievement School District; iZone = Innovation Zone; FRPM = eligibility for free or reduced priced meals; ELL = English language learner; SpED = special education status; TVAAS = Tennessee Value-Added Assessment System.

TABLE A7

Student Achievement Effects in Each Pre- and Postturnaround Year Relative to the Baseline Year Before Turnaround (Year 0)

| | (1) | (2) | (3) |
|-------------------------------|------------------|-------------------|-------------------|
| | Reading | Math | Science |
| ASD Year -5 | -.032 (.030) | -.049 (.056) | -.025 (.041) |
| ASD Year -4 | -.026 (.035) | -.040 (.041) | -.044 (.051) |
| ASD Year -3 | -.045 (.026) | -.035 (.036) | -.016 (.050) |
| ASD Year -2 | .004 (.035) | -.048 (.036) | .031 (.042) |
| ASD Year -1 | .045 (.030) | .012 (.050) | .048 (.056) |
| ASD Year 1 | -.000 (.046) | -.051 (.110) | -.035 (.113) |
| ASD Year 2 | .036 (.039) | .081 (.063) | .016 (.068) |
| ASD Year 3 | .092** (.033) | .059 (.078) | .126 (.076) |
| ASD Year 4 | .039 (.051) | -.052 (.074) | -.007 (.093) |
| ASD Year 5 | -.021 (.036) | -.099 (0.076) | .033 (.196) |
| ASD Year 6 | -.019 (.081) | -.174* (.086) | -.087 (.131) |
| iZone Year -5 | .049 (.031) | .099 (.053) | .090 (.052) |
| iZone Year -4 | .019 (.025) | .011 (.045) | .025 (.030) |
| iZone Year -3 | .006 (.024) | .027 (.040) | .015 (.027) |
| iZone Year -2 | -.014 (.020) | -.040 (.031) | -.016 (.028) |
| iZone Year -1 | .001 (.023) | .058 (.039) | .002 (.027) |
| iZone Year 1 | .025 (.033) | .183*** (.053) | .141* (.059) |
| iZone Year 2 | .083* (.035) | .171** (.056) | .214*** (.057) |
| iZone Year 3 | .064 (.038) | .093 (.057) | .139** (.053) |
| iZone Year 4 | -.046 (.082) | -.066 (.121) | .232* (.096) |
| iZone Year 5 | .103* (.041) | .068 (.057) | .112** (.041) |
| iZone Year 6 | .135** (.042) | .086 (.079) | .063 (.100) |
| Student and school covariates | Yes | Yes | Yes |
| School fixed effect | Yes | Yes | Yes |
| Number of schools | 116 | 116 | 116 |
| Observations | 219,618 | 199,161 | 187,318 |
| Adjusted R^2 | .442 | .383 | .370 |

Note. Standard errors in parentheses. Year 0 is the reference year. ASD = Achievement School District; iZone = Innovation Zone.

* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE A8

ASD and iZone Effects on Test Scores and Alternative Outcomes by Year

| | Student level | | | Teacher level | | | Principal level | | |
|----------------------------------|-------------------|---------------------|-----------------------------|-----------------------------|---------------------------------------|-----------------------------|-------------------------------|-----------------------------------|-------------------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| | Test Scores | Attendance rate | Student new to school (LPM) | Teacher new to school (LPM) | TVAAS index for all incoming teachers | Teacher years of experience | Principal new to school (LPM) | Principal observation score (1–5) | Principal years of experience |
| ASD | | | | | | | | | |
| ASD Year 1 | −0.037 (0.064) | −0.018* (0.0074) | 0.0051 (0.02) | 0.48*** (0.074) | −0.084 (0.34) | −4.3*** (0.72) | 0.73*** (0.12) | −0.058 (0.19) | −3.1** (1) |
| ASD Year 2 | 0.023 (0.056) | −0.0048 (0.0078) | −0.026 (0.028) | 0.099* (0.046) | 0.24 (0.26) | −5.7*** (0.76) | 0.53*** (0.13) | 0.4* (0.17) | −4.5*** (0.79) |
| ASD Year 3 | 0.017 (0.051) | −0.0036 (0.0074) | −0.063* (0.026) | 0.089 (0.057) | −0.47 (0.41) | −5.6*** (0.59) | 0.08 (0.17) | 0.1 (0.15) | −3.6*** (1) |
| ASD Year 4 | −0.056 (0.055) | −0.022* (0.009) | −0.09** (0.032) | 0.14* (0.057) | −0.84* (0.41) | −5.4*** (0.67) | 0.099 (0.19) | 0.037 (0.17) | −1.8 (1.5) |
| ASD Year 5 | −0.04 (0.092) | −0.018 (0.0099) | −0.13** (0.039) | .073 (.068) | −0.87* (0.41) | −4.5*** (0.67) | 0.98*** (0.16) | 0.13 (0.13) | −2.9* (1.4) |
| ASD Year 6 | −0.084 (0.1) | −0.01 (0.0077) | −0.11* (0.042) | 0.046 (0.083) | −1.1* (0.48) | −4*** (1.1) | 0.45* (0.23) | 0.2 (0.17) | −2.4 (1.3) |
| iZone | | | | | | | | | |
| iZone Year 1 | 0.11* (0.042) | −0.003 (0.0066) | −0.027 (0.02) | 0.23*** (0.044) | 0.38 (0.31) | −1.8*** (0.38) | 0.22 (0.14) | −0.089 (0.15) | −1.2 (0.73) |
| iZone Year 2 | 0.15** (0.044) | 0.0013 (0.0066) | −0.00066 (0.021) | 0.014 (0.036) | −0.089 (0.21) | −1.8*** (0.45) | 0.056 (0.14) | 0.46*** (0.12) | −0.56 (0.86) |
| iZone Year 3 | 0.027 (0.045) | −0.0035 (0.0078) | 0.016 (0.024) | −0.032 (0.039) | 0.2 (0.31) | −1.3* (0.52) | 0.051 (0.13) | 0.2 (0.16) | −1.4 (0.86) |
| iZone Year 4 | −0.059 (0.1) | 0.006 (0.0095) | −0.024 (0.025) | 0.015 (0.043) | −0.11 (0.25) | −1.8** (0.61) | 0.34* (0.15) | −0.23 (0.17) | −1.9 (1) |
| iZone Year 5 | 0.11 (0.07) | 0.002 (0.0076) | −0.062 (0.035) | −0.055 (0.04) | −.1 (0.31) | −1.5* (0.71) | −0.078 (0.16) | −0.23 (0.19) | −1 (1.1) |
| iZone Year 6 | 0.14* (0.067) | −0.012 (0.0083) | 0.0073 (0.033) | −0.077 (.053) | −0.51 (0.39) | −1.4 (0.86) | −0.23 (0.2) | .19 (0.2) | 0.16 (1.4) |
| Individual and school covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| School fixed effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 315,663 | 272,112 | 272,202 | 26,573 | 1,259 | 26,281 | 753 | 667 | 753 |
| Adjusted R ² | .417 | .015 | .017 | .036 | .024 | .074 | .072 | .117 | .096 |

Note. Standard errors in parentheses. LPM denotes linear probability models for a dichotomous outcome. Student new to school does not include structural moves. The TVAAS index is a standardized, continuous value-added measure of effectiveness calculated only for test subject teachers. Principal observation scores range from 1 (*significantly below expectations*) to 5 (*significantly above expectations*). ASD = Achievement School District; iZone = Innovation Zone; TVAAS = Tennessee Value-Added Assessment System.

* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE A9

Effects of ASD and iZone Interventions on Student Test Scores Using a Comparative Interrupted Time Series Model

| | (1) | (2) | (3) | (4) |
|--------------------------------|--------------------|-------------------|--------------------|-------------------|
| | All subjects | Reading | Math | Science |
| Ever ASD * After treatment | -.048 (.045) | -.028 (.040) | -.060 (.064) | -.056 (.068) |
| Ever iZone * After treatment | .107* (.046) | .044 (.040) | .128* (.052) | .171* (.069) |
| Linear time trend | -.020*** (.006) | -.014* (.006) | -.026*** (.007) | -.019* (.008) |
| Ever ASD * Linear time trend | .017* (.008) | .015 (.008) | .017 (.009) | .018 (.011) |
| Ever iZone * Linear time trend | -.000 (.007) | .004 (.007) | -.003 (.008) | -.005 (.010) |
| Prior year lag | .604*** (.004) | .633*** (.004) | .589*** (.007) | .577*** (.006) |
| Student and school covariates | Yes | Yes | Yes | Yes |
| School fixed effect | Yes | Yes | Yes | Yes |
| Number of schools | 116 | 116 | 116 | 116 |
| Observations | 606,097 | 219,618 | 199,161 | 187,318 |
| Adjusted R^2 | .401 | .442 | .383 | .370 |

Note. Standard errors in parentheses. The linear time trend is centered such that Year 0 is the year prior to schools entering either the ASD or an iZone. *ASD* and *iZone* represent the interaction between indicators for schools that are ever ASD or ever iZone and an indicator for years after schools begin turnaround. Comparison group schools are all priority schools not receiving any turnaround interventions on either the 2012 or 2014 priority list. ASD = Achievement School District; iZone = Innovation Zone.

* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE A10

Placebo Test for Prior Year Dip in Performance by Centering Treatment on the Year Before Turnaround

| | (1) | (2) | (3) |
|-------------------------------|-----------------|-----------------|-----------------|
| | Reading | Math | Science |
| ASD | -.012 (.026) | -.010 (.049) | -.020 (.058) |
| iZone | .003 (.032) | .100 (.051) | .061 (.045) |
| Student and school covariates | Yes | Yes | Yes |
| School fixed effect | Yes | Yes | Yes |
| Number of schools | 116 | 116 | 116 |
| Observations | 114,299 | 102,674 | 94,446 |
| Adjusted R^2 | .477 | .407 | .450 |

Note. Standard errors in parentheses. ASD = Achievement School District; iZone = Innovation Zone.

* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE A11
Student Fixed Effect Model

| | (1) | (2) | (3) |
|-------------------------------|-------------------|-------------------|-------------------|
| | Reading | Math | Science |
| ASD | .034 (.018) | .021 (.022) | .017 (.022) |
| iZone | .096*** (.013) | .189*** (.015) | .124*** (.016) |
| Student and school covariates | Yes | Yes | Yes |
| Student fixed effect | Yes | Yes | Yes |
| Number of schools | 116 | 116 | 116 |
| Observations | 1,14,299 | 1,05,023 | 96,341 |
| Adjusted R Squared | 0.045 | 0.062 | .081 |

Note. Standard errors in parentheses. ASD = Achievement School District; iZone = Innovation Zone.
* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE A12
Total Enrollment and Student Characteristics as Outcomes

| | (1) | (2) | (3) | (4) | (5) | (6) |
|-------------------------------|---------------------|-------------------|-------------------|-------------------|------------------|------------------|
| | Total enrollment | Female | Minority | FRPM | ELL | SpED |
| ASD | -41.218 (33.383) | -0.020 (0.011) | -0.000 (0.002) | 0.005 (0.007) | 0.008 (0.005) | 0.002 (0.013) |
| iZone | 11.899 (23.308) | -0.006 (0.009) | 0.002 (0.002) | -0.001 (0.005) | 0.003 (0.005) | 0.012 (0.009) |
| Student and school covariates | Yes | Yes | Yes | Yes | Yes | Yes |
| School fixed effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Number of schools | 116 | 116 | 116 | 116 | 116 | 116 |
| Observations | 91,647 | 91,647 | 91,659 | 91,647 | 91,647 | 172,683 |
| Adjusted R^2 | .076 | .045 | .015 | .130 | .269 | .025 |

Note. Standard errors in parentheses. ASD = Achievement School District; iZone = Innovation Zone; FRPM = eligibility for free reduced-price meals; ELL = English language learner; SpED = special education status.
* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE A13
Alternative Comparison Group of Lowest Performing 6% to 10% of Schools

| | (1) | (2) | (3) |
|-------------------------------|-----------------|-------------------|-------------------|
| | Reading | Math | Science |
| ASD | .016 (.027) | .033 (.036) | .019 (.031) |
| iZone | .055* (.025) | .129*** (.037) | .143*** (.037) |
| Student and school covariates | Yes | Yes | Yes |
| School fixed effect | Yes | Yes | Yes |
| Number of schools | 158 | 158 | 157 |
| Observations | 180,432 | 169,416 | 155,244 |
| Adjusted R^2 | .519 | .413 | .448 |

Note. Standard errors in parentheses. The average student achievement score is a calculated measure of average achievement scores from all students and all subjects in each school. ASD = Achievement School District; iZone = Innovation Zone.
* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE A14

ASD and iZones Effects on Test Score Missingness Across All Years and Cohorts

| | (1) | (2) | (3) | (4) |
|-------------------------------|--------------|--------------|--------------|--------------|
| | All Subjects | Reading | Math | Science |
| ASD | .02 (.01) | .00 (.00) | .02 (.02) | .04 (.03) |
| iZone | .01 (.01) | .00 (.00) | .00 (.00) | .04 (.02) |
| Student and school covariates | Yes | Yes | Yes | Yes |
| School fixed effect | Yes | Yes | Yes | Yes |
| Number of schools | 110 | 110 | 110 | 110 |
| Observations | 390,171 | 130,057 | 130,057 | 130,057 |
| Adjusted R^2 | .052 | .016 | .009 | .114 |

Note. Comparison group schools are all priority schools not receiving any turnaround interventions on either the 2012 or 2014 priority list. ASD = Achievement School District; iZone = Innovation Zone.

* $p < .05$ ** $p < .01$. *** $p < .001$.

TABLE A15

Standardized Test Scores for Transferring Students by Turnaround Year

| | ASD | | | iZone | | | Non-ASD, non-iZone Priority | | |
|---------|-------------|--------------|-------------------------|-------------|--------------|-------------------------|-----------------------------|--------------|-------------------------|
| | Transfer in | Transfer out | p value of difference | Transfer in | Transfer out | p value of difference | Transfer in | Transfer out | p value of difference |
| Year 1 | | | | | | | | | |
| Reading | -1.02 | -1.00 | .87 | -0.90 | -0.93 | .63 | -0.92 | -0.93 | .97 |
| Math | -1.02 | -0.87 | .07 | -0.69 | -0.70 | .86 | -0.84 | -0.78 | .22 |
| Science | -1.16 | -1.12 | .68 | -0.92 | -0.97 | .56 | -0.98 | -0.99 | .84 |
| Year 2 | | | | | | | | | |
| Reading | -0.94 | -1.01 | .38 | -0.80 | -0.82 | .76 | -0.85 | -0.87 | .66 |
| Math | -0.76 | -0.86 | .38 | -0.69 | -0.68 | .85 | -0.79 | -0.79 | .91 |
| Science | -1.01 | -1.12 | .41 | -0.78 | -0.81 | .75 | -0.90 | -0.91 | .81 |
| Year 3 | | | | | | | | | |
| Reading | -0.93 | -0.84 | .45 | -0.80 | -0.90 | .17 | -0.79 | -0.87 | .13 |
| Math | -0.87 | -0.78 | .31 | -0.72 | -0.82 | .25 | -0.71 | -0.82 | .07 |
| Science | -0.88 | -0.91 | .80 | -0.83 | -0.86 | .73 | -0.82 | -0.93 | .07 |
| Year 4 | | | | | | | | | |
| Reading | -0.77 | -0.74 | .86 | -0.88 | -0.93 | .67 | -0.85 | -0.86 | .90 |
| Math | -0.80 | -0.82 | .86 | -0.87 | -0.91 | .75 | -0.86 | -0.87 | .89 |
| Science | -0.86 | -0.93 | .65 | -0.88 | -0.83 | .68 | -0.88 | -0.83 | .59 |
| Year 5 | | | | | | | | | |
| Reading | -0.91 | -0.77 | .19 | -0.90 | -0.92 | .83 | -0.84 | -0.84 | .97 |
| Math | -0.91 | -0.83 | .50 | -0.85 | -0.84 | .90 | -0.81 | -0.80 | .82 |
| Science | -1.00 | -0.96 | .84 | -0.89 | -0.80 | .52 | -0.86 | -0.86 | .96 |

Note. Year 6 could not be calculated because transfer out cannot be observed for the final year. ASD = Achievement School District; iZone = Innovation Zone.

* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE A16

Remove Observations Where Student Move Between Turnaround and Comparison Schools

| | Dropped movers from turnaround to nonturnaround schools | | | Dropped any movers from or to turnaround schools or between turnaround schools | | |
|-------------------------------|--|-----------------|------------------|---|-----------------|------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | Reading | Math | Science | Reading | Math | Science |
| ASD | .010 (.032) | -.013 (.045) | .008 (.043) | .013 (.032) | -.010 (.045) | .007 (.043) |
| iZone | .054 (.029) | .092* (.042) | .146** (.047) | .051 (.029) | .093* (.042) | .144** (.047) |
| Student and school covariates | Yes | Yes | Yes | Yes | Yes | Yes |
| School fixed effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Number of schools | 116 | 116 | 116 | 116 | 116 | 116 |
| Observations | 111,836 | 102,711 | 94,629 | 110,188 | 101,131 | 93,294 |
| Adjusted R^2 | .480 | .370 | .403 | .481 | .370 | .403 |

Note. Standard errors in parentheses. ASD = Achievement School District; iZone = Innovation Zone.

* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE A17

Interactions With an Indicator for High Schools

| | (1) | (2) | (3) | (4) |
|-------------------------------|------------------|-------------------|-------------------|-----------------|
| | All subjects | Reading | Math | Science |
| ASD | .01 (.03) | .00 (.03) | -.02 (.05) | .03 (.04) |
| ASD * High school | -.01 (.03) | .02 (0.06) | .11* (0.04) | -.11 (.13) |
| iZone | .13*** (.04) | .09** (.03) | .14*** (.04) | .16** (.05) |
| iZone * High school | -.20*** (.04) | -.16*** (0.04) | -.26*** (0.07) | -.06 (.06) |
| Prior year lag | .62*** (.00) | .67*** (.00) | .57*** (.01) | .61*** (.00) |
| Student and school covariates | Yes | Yes | Yes | Yes |
| School fixed effect | Yes | Yes | Yes | Yes |
| Number of schools | 116 | 116 | 116 | 116 |
| Observations | 3,15,663 | 1,14,299 | 1,05,023 | 96,341 |
| Adjusted R^2 | .417 | .478 | .370 | .403 |

Note. Standard errors in parentheses. ASD = Achievement School District; iZone = Innovation Zone.

* $p < .05$. ** $p < .01$. *** $p < .001$.

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.

ORCID iD

Lam D. Pham  <https://orcid.org/0000-0001-8031-7777>

Notes

1. Comparing the longer term effects of school turnaround with earlier school improvement initiatives like comprehensive school improvement also allows us to better assess models that explicitly focus on rapid change (i.e., turnaround) relative to models that require longer term implementation (i.e., school improvement). Also, some studies of turnaround find larger effects in math than in English Language Arts (ELA) in the early years of reform (Schueler et al., 2017), suggesting that gains may take longer to develop in some subjects.

2. Two other school reform models that have lasted beyond the first 3 years include the reforms in New Orleans (Harris & Larsen, 2019) and in Newark (Chin et al., 2019).

3. We find no significant differences between CMO-run and direct-run ASD schools. Results separating direct-run and CMO-run ASD schools are available on request.

4. Examining longer term effects on both models will help shed light on central office leadership as part of sustainable reforms. For the ASD, removing schools from their LEA could mean that reforms are shielded from changing district mandates; however, iZones could also support sustainability by building LEA capacity while schools are undergoing reform. Moreover, keeping schools within their LEA could support sustainable reforms because schools will not have to transition between state-led and district-led governance. No schools in Tennessee's ASD have transitioned back to their LEAs, but comparing the two models over time helps shed light on sustainable reforms at the district level.

5. See Appendix Table A1.

6. Our conclusions do not change when we completely exclude any schools that will eventually close.

7. See Appendix Table A2 and Figure 1.

8. Of the 83 priority schools on 2012 list, 69 are located in Memphis, and 50 of the 83 schools the 2014 list are located in Memphis. Restricting our analysis to only Memphis-area schools do not change our substantive conclusions. See Appendix Table A1.

9. The total number of schools on the 2014 priority list was 85 but two were schools that split off from existing ASD schools. Our analysis considers these split-off schools as part of the original school. While both the ASD and iZones operated new-start schools that did not exist previously, they were not priority schools when they opened, were usually the result of CMOs opening new schools independent of ASD operations, and are not included in our analysis.

10. We test the sensitivity of our results to an alternative sample where we include the end-of-course scores in high school subjects from 2015 to 2016, resulting in very minor changes. Much of this is because both the ASD and iZones operated very few high schools in 2015–2016, so including these scores add little to our analysis.

11. Nonstructural moves do not include transfers where students move to a different school level after completing the final grade offered at their previous school (e.g., moving to a high school after completing the final grade in a middle school).

12. For a complete list of all domains and indicators on the TILS rubric, see Grissom et al. (2018).

13. The model does not include separate indicators for schools that are ever ASD or an indicator for schools that are ever iZone schools, because the school fixed effect is perfectly collinear with these indicators.

14. We control for student race because of the rich literature showing that racial stereotypes can influence the effect of school interventions on student achievement (Cohen et al., 2006). We control for gender because gender has been shown to be correlated with differential test scores (Legewie & DiPrete, 2012). Likewise, we control for FRPM, ELL, and SpED because existing literature finds that social economic status (Chen & Weikart, 2008), English proficiency (Fry, 2008), and eligibility for special education services (Hibel et al., 2010) all have an effect on both the likelihood

that students attend low-performing schools and directly on their achievement.

15. We control for teacher and principal individual characteristics, because they have all been shown to affect teachers' and principals' decisions to either leave the profession or transfer schools, their effectiveness, and the likelihood that they will choose to work in a low-performing school (Borman & Dowling, 2008; Grissom, 2011; Guarino et al., 2006; Ingersoll & Conner, 2009).

16. We also test models that include a number of different school covariates such as total enrollment, but none produced substantively differing results.

17. Separating the *After* variable into sets of year and cohort indicators are especially important alternative models in light of recent evidence from Goodman-Bacon (2018) who provides new evidence on DID models averaging across multiple treatment times. Since our overall DID estimates are a weighted average of all combinations of estimators comparing earlier and later treated schools, using earlier treated schools as a control after their treatment can introduce bias if treatment effects vary over time and changes in treatment effects over time are incorporated into the overall DID estimate. To examine this potential for bias, we follow Goodman-Bacon's suggestion to examine event-study models with indicators for different treatment years and cohorts. Our event-study model estimates support the conclusions from our overall DID estimates but provide more nuanced estimates for different treatment times.

18. We also test for statistically significant changes to student characteristics after turnaround, as described in the section on falsification and robustness checks.

19. Three cases show marginally significant differences: ASD math scores in Year 1 and comparison schools' math and science scores in Year 3, but even in these cases, the small substantive differences between test scores for incoming and outgoing students suggest that schools are not strategically altering the student population.

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, & U.S. Department of Education. <https://www2.ed.gov/rschstat/eval/other/dramatic-school-improvement/exploratory-study.pdf>
- Anrig, G. (2015). *Lessons from school improvement grants that worked*. Century Foundation. <http://www.tcf.org/blog/detail/lessons-from-school-improvement-grants-that-worked>
- Ashenfelter, O. (1978). Estimating the effect of training programs on earnings. *Review of Economics and Statistics*, 60(1), 47–57. <https://doi.org/10.2307/1924332>
- Bauman, C., & Aldrich, M. (2020, January 6). *All 30 schools in Tennessee's turnaround district would exit by 2022 in a massive restructuring proposal*. Chalkbeat. <https://www.chalkbeat.org/posts/tn/2020/01/06/all-30-schools-would-exit-tennessees-achievement-school-district-by-2022-under-proposed-reset/>
- Berends, M., Bodilly, S. J., & Kirby, S. N. (2002). *Facing the challenges of whole-school reform: New American Schools after a*

- decade. RAND Corporation. https://www.rand.org/pubs/monograph_reports/MR1498.html
- Bloom, H. S., Ham, S., Melton, L., & O'Brien, J. (2001). *Evaluating the accelerated schools approach: A look at early implementation and impacts on student achievement in eight elementary schools*. Manpower Demonstration Research Corporation. <https://www.mdrc.org/publication/evaluating-accelerated-schools-approach>
- Bonilla, S., & Dee, T. (2017). *The effects of school reform under NCLB waivers: Evidence from focus schools in Kentucky* (NBER Working Paper No. 23462). National Bureau of Economic Research. <https://www.nber.org/papers/w23462.pdf>
- Borman, G. D., & Dowling, N. M. (2008). Teacher attrition and retention: A meta-analytic and narrative review of the research. *Review of Educational Research*, 78(3), 367–409. <https://doi.org/10.3102/0034654308321455>
- Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research*, 73(2), 125–230. <https://doi.org/10.3102/00346543073002125>
- Calkins, A., Guenther, W., Belfiore, G., & Lash, D. (2007). *The turnaround challenge: Why America's best opportunity to dramatically improve student achievement lies in our worst-performing schools*. Mass Insight Education & Research Institute.
- 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>
- Center for Research on Education Outcomes. (2019). *City Studies Project*. <https://cityschools.stanford.edu/>
- Chen, G., & Weikart, L. A. (2008). Student background, school climate, school disorder, and student achievement: An empirical study of New York City's middle schools. *Journal of School Violence*, 7(4), 3–20. <https://doi.org/10.1080/15388220801973813>
- Chin, M., Kane, T. J., Kozakowski, W., Schueler, B. E., & Staiger, D. O. (2019). School district reform in Newark: Within- and between-school changes in achievement growth. *ILR Review*, 72(2), 323–354. <https://doi.org/10.1177/0019793918774318>
- Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, 313(5791), 1307–1310. <https://doi.org/10.1126/science.1128317>
- Copland, M. A. (2003). Leadership of inquiry: Building and sustaining capacity for school improvement. *Educational Evaluation and Policy Analysis*, 25(4), 375–395. <https://doi.org/10.3102/01623737025004375>
- Dee, T. (2012). *School turnarounds: Evidence from the 2009 stimulus* (NBER Working Paper No. 17990). National Bureau of Economic Research. <https://doi.org/10.3386/w17990>
- Dee, T., & Dizon-Ross, E. (2017). *School performance, accountability and waiver reforms: Evidence from Louisiana* (NBER Working Paper No. 23463). National Bureau of Economic Research. <https://doi.org/10.3386/w23463>
- Dougherty, S. M., & Weiner, J. M. (2019). The Rhode to turnaround: The impact of waivers to No Child Left Behind on school performance. *Educational Policy*, 33(4), 555–586. <https://doi.org/10.1177/0895904817719520>
- Dragoset, L., Thomas, J., Herrmann, M., Deke, J., James-Burdumy, S., Graczewski, C., Boyle, A., Upton, R., Tanenbaum, C., & Giffin, J. (2017). *School improvement grants: Implementation and effectiveness (NCEE 2017-4013)*. National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, & U.S. Department of Education. <https://ies.ed.gov/ncee/pubs/20174013/pdf/20174013.pdf>
- Egbert, R. L. (1981). *Some thoughts about follow through thirteen years later* (ED244733). ERIC. <https://files.eric.ed.gov/fulltext/ED244733.pdf>
- Friedman-Krauss, A. H., & Raver, C. C. (2015). Does school mobility place elementary school children at risk for lower math achievement? The mediating role of cognitive dysregulation. *Developmental Psychology*, 51(12), 1725–1739. <https://doi.org/10.1037/a0039795>
- Fry, R. (2008). *The role of schools in the English language learner achievement gap*. Pew Research Center. <https://www.pewresearch.org/hispanic/2008/06/26/the-role-of-schools-in-the-english-language-learner-achievement-gap/>
- 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>
- Goodman-Bacon, A. (2018). *Difference-in-Differences with variation in treatment timing* (NBER Working Paper No. 25018). National Bureau of Economic Research. <https://doi.org/10.3386/w25018>
- Grissom, J. A. (2011). Can good principals keep teachers in disadvantaged schools? Linking principal effectiveness to teacher satisfaction and turnover in hard-to-staff environments. *Teachers College Record*, 113(11), 2552–2585.
- Grissom, J. A., Blissett, R. S., & Mitani, H. (2018). Evaluating school principals: Supervisor ratings of principal practice and principal job performance. *Educational Evaluation and Policy Analysis*, 40(3), 446–472. <https://doi.org/10.3102/0162373718783883>
- 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. <https://doi.org/10.3102/0162373709333886>
- 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. <https://doi.org/10.3102/00346543076002173>
- 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. <https://educationresearchalliancenola.org/files/publications/The-Effects-of-the-New-Orleans-Post-Katrina-School-Reforms-on-Student-Academic-Outcomes.pdf>
- Harris, D. N., & Larsen, M. F. (2019). *The identification of schooling preferences: Methods and evidence from post-Katrina New Orleans* [Technical report]. Education Research Alliance for New Orleans. <https://educationresearchalliancenola.org/files/publications/Harris-Larsen-How-Parents-Choose-2019-07-05.pdf>

- Heissel, J. A., & Ladd, H. F. (2018). School turnaround in North Carolina: A regression discontinuity analysis. *Economics of Education Review*, 62(February), 302–320. <https://doi.org/10.1016/j.econedurev.2017.08.001>
- Hemelt, S. W., & Jacob, B. (2017). *Differentiated accountability and education production: Evidence from NCLB waivers* (NBER Working Paper No. 23461). National Bureau of Economic Research. <https://doi.org/10.3386/w23461>
- Henry, G. T., Guthrie, J. E., & Townsend, L. W. (2015). *Outcomes and impacts of North Carolina's initiative to turn around the lowest-achieving schools*. Consortium for Educational Research and Evaluation–North Carolina. <https://cerenc.org/wp-content/uploads/2015/09/0-FINAL-Final-DST-Report-9-3-15.pdf>
- Henry, G. T., & Harbatkin, E. (2018, November 8–10). *The next generation of school reform: Improving the lowest performing schools without disrupting the status quo* [Paper presentation]. Annual Conference of the Association for Public Policy Analysis and Management, Washington, DC, United States.
- Henry, G. T., Pham, L., Kho, A., & Zimmer, R. (2020). Peeking into the black box of school turnaround: A formal test of mediators and suppressors. *Education Evaluation and Policy Analysis*. Advance online publication. <https://doi.org/10.3102/0162373720908600>
- Henry, G. T., & Redding, C. (2020). The consequences of leaving school early: The effects of within-year and end-of-year teacher turnover. *Education Finance and Policy*, 15(2), 332–356. https://doi.org/10.1162/edfp_a_00274
- Herman, R., Dawson, P., Dee, T., Greene, J., Maynard, R., Redding, S., & Darwin, M. (2008). *Turning around chronically low-performing schools. IES Practice Guide*. National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, & U.S. Department of Education. https://ies.ed.gov/ncee/wwc/Docs/PracticeGuide/Turnaround_pg_04181.pdf
- Hibel, J., Farkas, G., & Morgan, P. L. (2010). Who is placed into special education? *Sociology of Education*, 83(4), 312–332. <https://doi.org/10.1177/0038040710383518>
- Ingersoll, R. M., & Conner, R. (2009, April 13–17). *What the national data tell us about minority and Black teacher turnover* [Paper presentation]. Annual Meeting of the American Educational Research Association, San Diego, CA, United States.
- Kho, A., Henry, G. T., Zimmer, R., & Pham, L. (2018). *How has iZone teacher recruitment affected the performance of other schools?* Tennessee Education Research Alliance. https://peabody.vanderbilt.edu/TERA/files/iZone_Teacher_Recruitment_Effects.pdf
- Kutash, J., Nico, E., Gorin, E., Rahmatullah, S., & Tallant, K. (2010). *The school turnaround field guide*. FSG Social Impact Advisors. <https://www.wallacefoundation.org/knowledge-center/pages/turnaround-actors-school-turnaround-field-guide.aspx>
- Legewie, J., & DiPrete, T. A. (2012). School context and the gender gap in educational achievement. *American Sociological Review*, 77(3), 463–485. <https://doi.org/10.1177/0003122412440802>
- McEachin, A. J., Welsh, R. O., & Brewer, D. J. (2016). The variation in student achievement and behavior within a portfolio management model: Early results from New Orleans. *Educational Evaluation and Policy Analysis*, 38(4), 669–691. <https://doi.org/10.3102/0162373716659928>
- North Carolina Department of Public Instruction. (2018). *The Elementary and Secondary Education Act of 1965, as Amended by the Every Student Succeeds Act: North Carolina Consolidated State Plan*. <https://files.nc.gov/dpi/documents/succeeds/nc-essa-state-plan-final.pdf>
- Papay, J., & Hannon, M. (2018). *The effects of school turnaround strategies in Massachusetts* (EdWorking Paper No. 19-66). Brown University. <https://pdfs.semanticscholar.org/f2f7/b71ee-1768b593ec332d7f4be827d2f28bb64.pdf>
- Pearlman, C. L., & Redding, S. (Eds.). (2011). *Handbook on effective implementation of School Improvement Grants*. Academic Development Institute. http://www.centerii.org/handbook/Resources/Handbook_on_Effective_Implementation_of_School_Improvement_Grants.pdf
- Peurach, D. J., & Neumerski, C. M. (2015). Mixing metaphors: Building infrastructure for large scale school turnaround. *Journal of Educational Change*, 16(4), 379–420. <https://doi.org/10.1007/s10833-015-9259-z>
- Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal*, 50(1), 4–36. <https://doi.org/10.3102/0002831212463813>
- Ruble, W. (2015). *The effect of contracting out low performing schools on student performance* [Paper presentation]. Conference of the Association of Public Policy and Management in Miami, FL, United States.
- Schueler, B. E., Goodman, J. S., & Deming, D. J. (2017). Can states take over and turn around school districts? Evidence from Lawrence, Massachusetts. *Educational Evaluation and Policy Analysis*, 39(2), 311–332. <https://doi.org/10.3102/0162373716685824>
- Springer, M. G., Swain, W. A., & Rodriguez, L. A. (2016). Effective teacher retention bonuses: Evidence from Tennessee. *Educational Evaluation and Policy Analysis*, 38(2), 199–221. <https://doi.org/10.3102/0162373715609687>
- Strunk, K. O., Marsh, J. A., Hashim, A. K., & Bush-Mecenas, S. (2016). Innovation and a return to the status quo: A mixed-methods study of school reconstitution. *Educational Evaluation and Policy Analysis*, 38(3), 549–577. <https://doi.org/10.3102/0162373716642517>
- Stuit, D. A. (2010). *Are bad schools immortal? The scarcity of turnarounds and shutdowns in both charter and district sectors*. Thomas B. Fordham Institute. <https://fordhaminstitute.org/sites/default/files/publication/pdfs/fordhamimmortal10.pdf>
- 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>
- Tennessee Department of Education. (2012). *ESEA flexibility request*. <https://www2.ed.gov/policy/eseaflex/approved-requests/tn.pdf>
- Zimmer, R., & Guarino, C. (2013). Is there empirical evidence that charter schools “push out” low-performing students? *Educational Evaluation and Policy Analysis*, 35(4), 461–480. <https://doi.org/10.3102/0162373713498465>
- Zimmer, R., Henry, G. T., & Kho, A. (2017). The effects of school turnaround in Tennessee’s Achievement School District and Innovation Zones. *Educational Evaluation and Policy Analysis*, 39(4), 670–696. <https://doi.org/10.3102/0162373717705729>

Authors

LAM D. PHAM is a PhD candidate at Vanderbilt University's Peabody College in the Department of Leadership, Policy, and Organizations. His research interests include school improvement and reform, with a focus on personnel policies in low-performing schools.

GARY T. HENRY is dean of the University of Delaware's College of Education and Human Development and professor in the School of Education and the Joseph R. Biden, Jr. School of Public Policy & Administration. A renowned researcher in the field of education,

Henry specializes in education policy, educational evaluation, educator labor markets, and quantitative research methods.

ADAM KHO is an assistant professor of education in K–12 Policy and Leadership at the University of Southern California. His research interests include education policy and evaluation of programs serving traditionally disadvantaged students, with a focus on school improvement, school reform, and school turnaround.

RON ZIMMER is a professor and director of the Martin School of Public Policy and Administration at University of Kentucky. His research focuses on school choice and school finance.