



Do the Effects Persist? An Examination of Long-term Effects After Students Leave Turnaround Schools

Lam D. Pham

North Carolina State
University

Sean P. Corcoran

Vanderbilt University

Gary T. Henry

University of Delaware

Ron Zimmer

University of Kentucky

Whole-school reforms have received widespread attention, but a critical limitation of the current literature is the lack of evidence around whether these extensive and costly interventions improve students' long-term outcomes after they leave reform schools. Leveraging Tennessee's statewide turnaround reforms, we use difference-in-differences models to estimate the effect of attending a turnaround middle school on student outcomes in high school, including test scores, attendance, chronic absenteeism, disciplinary actions, drop out, and high school graduation. We find little evidence to support improved long-run student outcomes – mostly null effects that are nearly zero in magnitude. Our results contribute to a broad call for educational researchers to examine whether school reforms meaningfully affect student outcomes beyond short-term improvements in test scores.

VERSION: September 2024

Suggested citation: Pham, Lam D., Sean P. Corcoran, Gary T. Henry, and Ron Zimmer. (2024). Do the Effects Persist? An Examination of Long-term Effects After Students Leave Turnaround Schools. (EdWorkingPaper: 24-1041). Retrieved from Annenberg Institute at Brown University: <https://doi.org/10.26300/3q14-3z85>

Do the Effects Persist?

An Examination of Long-term Effects After Students Leave Turnaround Schools

Lam D. Pham
North Carolina State University

Sean P. Corcoran
Vanderbilt University

Gary T. Henry
University of Delaware

Ron Zimmer
University of Kentucky

Abstract

Whole-school reforms have received widespread attention, but a critical limitation of the current literature is the lack of evidence around whether these extensive and costly interventions improve students' long-term outcomes after they leave reform schools. Leveraging Tennessee's statewide turnaround reforms, we use difference-in-differences models to estimate the effect of attending a turnaround middle school on student outcomes in high school, including test scores, attendance, chronic absenteeism, disciplinary actions, drop out, and high school graduation. We find little evidence to support improved long-run student outcomes – mostly null effects that are nearly zero in magnitude. Our results contribute to a broad call for educational researchers to examine whether school reforms meaningfully affect student outcomes beyond short-term improvements in test scores.

Keywords: school turnaround, school reform, long-term effects, difference-in-differences, school improvement

Acknowledgements: 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.

Notwithstanding any Tennessee Department of Education (TDOE) data or involvement in the creation of this research product, the TDOE does not guarantee the accuracy of this work or endorse the findings. Any errors are the sole responsibility of the author(s).

Introduction

Policy makers have made substantial investments in building organizational capacity to support improvements in the public sector. In education, one salient example involves the Obama administration's efforts to support chronically low-performing schools by encouraging states and districts to adopt federally approved school turnaround policies (Duncan, 2009). These reform efforts have received billions of dollars in resources through Race to the Top (RttT) and School Improvement Grants or SIGs (Dragoset et al., 2019). More recently, under No Child Left Behind (NCLB) waivers and the current Every Student Succeeds Act (ESSA), the federal government has committed to continuing the long-standing national interest in reforms aimed at persistently low-performing schools. Supporting chronically low-performing schools is particularly important now because these schools serve disproportionately high shares of low-income and racially minoritized students, and they experienced the largest decreases in student achievement as a result of the COVID-19 pandemic (Dorn et al., 2020; Kogan & Lavertu, 2021; Kuhfeld et al., 2020). The long history of investment, current requirements under ESSA, and historically high-need exacerbated by COVID-19 suggest that school reforms will continue to prominently factor into national efforts to support equitable educational opportunities.

As part of these federal initiatives, states and districts across the country have now implemented whole-school reform policies for over a decade. While there is a long and growing list of studies that have examined the short-term effects of these reforms, few studies have examined long-term effects (for reviews, see Redding & Nguyen, 2020; Schueler et al., 2021). Among the smaller set of studies that have examined long-term effects, most of the interest has been in following new cohorts of schools placed into an existing reform model (Pham et al., 2020) or effects in reform schools after interventions have ended (Sun et al., 2021). Another

Long Term Outcomes of Turnaround

broadly important, but understudied, aspect of long-term effects is the impact on student outcomes after they leave the reform school. Our study sheds light on this much less understood aspect of long-term effects. To our knowledge, the only studies to examine this type of long-term effect focuses on the Recovery School District (RSD) reforms that essentially converted all schools in New Orleans into charter schools following Hurricane Katrina (Glenn & Harris, 2020; Harris & Larsen, 2023). While the RSD studies document positive long-term effects, the RSD is designed as a wholesale market-based reform, so these results may not readily apply to models that rely less on school choice. Additionally, given current ambiguities in the extant literature, we develop and contribute an overarching conceptual framework for delineating different types of long-term effects, with concrete nomenclature that will sharpen future academic and policy discourse on this topic.

Examining long-term effects on students is broadly consequential in educational research because interventions producing only short-term gains may not merit investment if they do not meaningfully improve students' educational trajectory and life opportunities. Additionally, effects across a large range of interventions, including many outside of school reform, may only be realized in the long-run. For instance, in the charter school literature, studies have suggested that charter schools may not have positive effects on test scores in the short run but can have long-term effects on educational attainment and labor market outcomes (Booker et al., 2011; Sass et al., 2016). However, the disruptive nature of school reform could also have adverse effects in the long run if students are receiving lower quality instruction while the school implements major changes. Examining this issue has implications for a wide array of educational interventions because reforms focusing solely on short-term outcomes, such as end-of-year test scores, could ultimately be detrimental to students' longer-term educational opportunities.

Long Term Outcomes of Turnaround

Finally, much of the existing literature focuses on student achievement with few studies using outcomes beyond test scores.

To help address these gaps in knowledge, we study Tennessee's two statewide reform models – the state-led Achievement School District (ASD) and district-led Innovation Zones (iZones). Tennessee is an highly informative and generalizable case because the state's reform models overlap substantially with current ESSA-aligned reform models used across the country in states such as Michigan, North Carolina, Florida, Illinois, Louisiana, and Massachusetts (Dragoset et al., 2019; Jochim, 2016; U.S. Department of Education, 2012). Initiated in 2012-13, Tennessee's reforms began when school turnaround, as defined by the SIG program, was the dominant approach to whole-school reform. Although states' current reform plans under ESSA are no longer dictated by the SIG-prescribed turnaround models, many states, including Tennessee, continue to use reform plans developed for turnaround (Rentner et al., 2017; TDOE, 2018). Substantial overlap between Tennessee's turnaround models and other reform efforts nationwide suggest that our results can inform similar initiatives in other states.

We describe both of Tennessee's turnaround models in detail below but note that the ASD relies on dramatic changes in school governance where chronically low-performing schools are removed from local districts, placed under state governance, and restarted, mainly by charter management organizations (CMOs). In contrast, iZone schools remain under the governance of their local district but are placed into an intra-district network that is supported by a team of full-time district staff. Although the ASD and iZones differ in their governance structures, both models replace teachers and principals when schools first begin implementing interventions, and the long-term effects of both can be estimated because they have been in place for over a decade. Previous research has examined both short-term impacts and effects over time as these reforms

Long Term Outcomes of Turnaround

matured (Pham et al., 2020; Zimmer et al., 2017). Specifically, previous research found that iZone reforms had positive effects on test scores, while ASD reforms produced null results (Pham et al., 2020). However, these prior studies have not examined whether effects either persist or materialize after students exit turnaround schools.

Our specific objective is to examine student outcomes in high school *after* they exit turnaround middle schools. We use a difference-in-differences (DiD) model to obtain intent to treat (ITT) effect estimates for students assigned to turnaround middle schools on their outcomes in high school. We examine a wide range of student outcomes when they reach high school, including attendance rates, chronic absenteeism, test scores, disciplinary outcomes, drop out, and graduation. We focus on middle schools, rather than elementary schools, because middle school is most proximate to meaningful long-term student outcomes in high school (e.g., graduation). Moreover, fewer studies have examined effects from turnaround interventions that occurred during middle school (Redding & Nguyen, 2020). To address concerns with estimating dynamic treatment effects (Callaway & Sant’Anna, 2021; Goodman-Bacon, 2021; Roth et al., 2023), we estimate DiD models using both a canonical approach and a “heterogeneity-robust” estimator recently developed by Callaway and Sant’Anna (2021). Results across both approaches lead to the same conclusions.

Overall, we find no evidence to support improved long-term student outcomes after they leave a turnaround middle school, with mostly null effects in high school for students assigned to either ASD or iZone middle schools. If anything, we find suggestive evidence that assignment to an ASD middle school produced negative effects on high school test scores, though these estimates are not consistently significant across all models. Likewise, assignment to iZone middle schools produced almost entirely null results except for a negative effect on math scores

Long Term Outcomes of Turnaround

that is marginally significant at the ten percent level. In the next section, we describe the turnaround context in Tennessee and review relevant literature. Then, we describe our data and methods. The final sections present results and conclusions.

School Turnaround in Tennessee

Tennessee's current turnaround models began under the state's 2010 First to the Top (FttT) legislation. FttT requires the Tennessee Department of Education (TDOE) to intervene in the state's priority schools, which are the lowest-performing five percent of schools. Beginning in 2012-13, Tennessee's priority schools were placed into the ASD, integrated into a district iZone, closed, or received no interventions. The ASD is Tennessee's boldest turnaround model – a statewide school district that removes priority schools from their local district to be governed directly by TDOE. If chosen for the ASD, priority schools cannot opt out, are required to replace the principal and at least 50% of teachers, and are restarted under new management. Most ASD schools are restarted under a CMO, except five that are directly managed by TDOE (Zimmer et al., 2017). The ASD relies on CMOs because TDOE leaders wanted to develop a model that would remove bureaucratic oversight and give school leaders flexibility in adapting reforms to individual school needs.

As the ASD began operating schools in summer 2012, TDOE also approved the creation of local iZones in Shelby County Schools (Memphis) and Metro-Nashville Public Schools (TDOE, 2012). Local iZones differ from the ASD primarily in that iZone schools remain part of their local education agency (LEA) and are placed into a district-within-a-district with other priority schools. Like the ASD, iZone schools must also replace the principal. In contrast to the ASD, iZone schools are not required to replace teachers, but in practice, iZone schools did replace at least 50 percent of teachers in the first year of reform (Henry et al., 2014). After the

Long Term Outcomes of Turnaround

first year of principal and teacher replacements, iZone school leaders have broad autonomy to manage daily school operations with support and oversight from a dedicated iZone office within the local district. Since 2012, two additional districts have opened iZones: Hamilton County Schools (Chattanooga) and Knox County Schools (Knoxville).

Importantly, neither ASD nor iZone schools are schools of choice because they are required to continue enrolling students from their local catchment area. This distinction sets our study apart from previous work examining long-term student outcomes in the New Orleans RSD, where market-based competition is a key feature of the model (Harris & Larsen, 2016). Instead, Tennessee's models are rooted in a belief that a set of disconnected interventions are insufficient to produce swift and dramatic improvements in student performance, which aligns with federal school reforms models that emphasize coordinated, schoolwide changes in how schools are governed, managed, and operated (Herman et al., 2008). These bold changes include teacher and principal replacement, flexibility and support for school leaders, and in the case of the ASD, even removal of chronically low-performing schools from local district governance.

Table 1 shows the number of schools in each turnaround model by school level (elementary, middle, and high).¹ Tennessee identified an initial list of 83 priority schools in 2012 (TN Press Release Center, 2012). A second list of 85 priority schools was released in 2014, though only 33 of these 85 schools were not already on the 2012 list (Carroll, 2014). Only schools on these two priority lists were eligible to join the ASD or an iZone, so as additional cohorts of schools were added to the ASD and iZones in each year, the remaining priority schools not receiving interventions grew smaller over time. Our analysis only compares students who were assigned to ASD or iZone middle schools, based on the feeder pattern of their elementary school, with students who were assigned to a non-ASD, non-iZone priority middle school, which we call

Long Term Outcomes of Turnaround

comparison schools. Table 1 shows that five cohorts of schools joined the ASD or an iZone between 2012-13 and 2016-17. After 2016-17, both models continued to operate existing schools, but no new schools were added to either model during the time of our study. As of 2018-19, the final year of data used in our analysis, no school selected for the ASD or iZones had exited. Thus, the ASD and iZone treatments are an absorbing state, such that once a school is treated, it remains treated through the last year of the panel. Below, we describe how we address potentially heterogeneous treatment effects across staggered turnaround adoption in the five ASD/iZone cohorts. However, it is important to highlight that the turnaround interventions did not change across the five cohorts. All ASD schools were removed from their local district; all iZone schools were placed into a district-led network that received additional support from district staff; and both implemented principal and teacher replacement.

To better understand how priority schools are selected to join either the ASD or an iZone, we interviewed leaders of both reform models. Our interviews suggested that there were no systematic inclusion criteria used to select schools, with the exception of targeting existing feeder patterns. However, in practice, even this criterion does not appear to strongly drive ASD or iZone school selection. Only four of the 26 ASD schools receive students from other ASD schools as part of the same feeder pattern, and only 8 of the 42 iZone schools receive students from other iZone schools as part of the same feeder pattern. Specifically, only 3 ASD and 4 iZone middle schools receive students from ASD and iZone elementary schools, respectively. Similarly, only 1 ASD and 4 iZone high schools receive students from ASD or iZone middle schools. Overall, our interviews with administrators suggested that priority schools were not systematically selected for specific turnaround strategies, but we cannot rule out potential selection based on unobserved school characteristics related to long-term student outcomes. To

Long Term Outcomes of Turnaround

address these concerns, we show evidence below of baseline equivalence and parallel trends between ASD, iZone, and comparison schools.

Literature Review

Whole-school reforms under RttT and SIGs emphasized school turnaround as federally specified models for supporting chronically low-performing schools (Dragoset et al., 2019; US Department of Education, 2009). These turnaround models emphasized rapidly improving school performance and relied on bold interventions, e.g., replacing the principal and at least half of all teachers (Aladjem et al., 2010; Herman et al., 2008; Peurach & Neumerski, 2015).

Two recent meta-analyses of the school turnaround literature (Redding & Nguyen, 2020; Schueler et al., 2021) report mixed results and show that the literature almost exclusively examines short-term outcomes, especially test scores. Redding and Nguyen (2020) review 35 studies of turnaround and find that these reforms are associated with improved test scores, student attendance, and graduation rates, but the latter two outcomes have been examined by a much smaller set of studies. All studies in Redding's and Nguyen's (2020) meta-analysis examined outcomes while students were still in the turnaround school, not after they left the turnaround school. In a second meta-analysis, Schueler et al. (2021) report a more expansive analysis that includes evaluations of any whole-school reform implemented since the adoption of NCLB. They find a moderate positive effect on math but no effect on English Language Arts achievement on high-stakes exams, positive impacts on low-stakes exams in STEM and humanities subjects, and no effect among the small set of studies that have examined non-test score outcomes. Among the studies reviewed by Redding and Nguyen (2020) and Schueler et al. (2021) are two evaluations of Tennessee's ASD and iZone reform efforts (Pham et al., 2020; Zimmer et al., 2017). These two studies found no effect in either the first three or six years of

Long Term Outcomes of Turnaround

implementation for the ASD but modest positive effects both in the first three and six years of implementation for iZone schools – though the positive iZone effects are concentrated in the initial cohorts with later cohorts showing no positive effects.

Collectively, the research examining turnaround reforms have helped to illuminate their short-term effects. However, with the exception of research on the RSD in New Orleans (Glenn & Harris, 2020; Harris & Larsen, 2016, 2023), research has not examined the long-term impact of these reform policies for students after they exit a turnaround school. The ASD and iZone models in Tennessee share commonalities with multiple other reform models across the country. First, Tennessee identifies its lowest-performing schools and the public attention to these schools aligns with accountability systems in states such as North Carolina (Henry & Harbatkin, 2020), Florida (Chiang, 2009; Rouse et al., 2013), and New York (Rockoff & Turner, 2010). Second, both the ASD and iZone models emphasize efforts to staff low-performing schools with effective principals and teachers, which mirrors turnaround models across the country including initiatives in California (Dee, 2012; Strunk et al., 2016; Sun et al., 2021), Ohio (Carlson & Lavertu, 2018), and Washington D.C. (Dee & Wyckoff, 2015), among many others (Dragoset et al., 2019). Furthermore, the ASD model, where low-performing schools are removed from local district governance, has been a topic of wide-ranging interest, with parallel policies in place across 34 states (Jochim, 2016). Comparable examples include charter conversion models in Boston and Baton Rouge (Abdulkadiroğlu et al., 2016) and state takeover of entire school districts that have occurred in 28 states since the 1980s (Schueler & Bleiberg, 2021). As the district-level, the iZones' district-within-a-district approach has been growing in popularity in districts such as Chicago, Denver, and Indianapolis (Iyengar et al., 2017).

Long Term Outcomes of Turnaround

Furthermore, less attention has been given to the effects of turnaround reforms implemented in middle schools (Redding & Nguyen, 2020; Schueler et al., 2021). Across their comprehensive review of this literature, Redding and Nguyen (2020) found only 4 of 35 studies (11%) that specifically focused on turnaround reforms implemented in middle schools. Of those four studies, only one met Redding's and Nguyen's criteria for a high-quality study with a low risk of bias. Additionally, researchers have long documented heterogeneity in the effects of school interventions by grade level (Edgerton & Desimone, 2019; Firestone & Herriott, 1982), so increased attention to reforms in middle schools will advance our collective understanding of how effective interventions differ across grade levels.

Overall, we contribute to the extant literature by:

1. developing a novel and broadly applicable framework for differentiating the currently ill-defined types of long-term effects in educational research;
2. focusing on turnaround models in Tennessee that are more comparable to reforms in other states than the whole district, market-based RSD model;
3. estimating effects on student outcomes after they leave a turnaround school,
4. examining the effects of reforms implemented during middle schools on high school outcomes; and
5. examining student outcomes beyond test scores to include important outcomes such as high school graduation.

Conceptual Framework for Differentiating Long-Term Effects

In reviewing literature on long-term effects of educational interventions, we find that the current academic conversation lacks clarity because prior literature conflates four distinct types of long-term effects: program maturation effects, dosage effects, institutionalized effects, and

Long Term Outcomes of Turnaround

persistent/latent effects. To sharpen academic discourse in this area, we developed a conceptual framework for understanding and delineating these different types of long-term effects, as shown in Figure 1. Although we tailor our discussion to long-term effects of whole-school reforms, this conceptual framework can be adapted to a broad range of educational interventions. We begin by separating long-term effects along two dimensions: (1) a temporal dimension that considers whether the interventions are still active or whether they have ended and (2) a level dimension that considers whether the research study is interested in examining effects on the school as whole or effects on students.

While interventions are actively being implemented, two types of long-term effects of interest include school-level program maturation effects and student-level dosage effects. Program maturation effects primarily capture the experience of new schools (or new cohorts) that join an existing reform model. Although these schools may be implementing reforms that are new to them, the reform program itself is not new because it has been implemented in other schools. Therefore, program maturation effects capture the long-run effect of a reform model as it grows, matures, and scales-up in more schools over time. One study of program maturation effects is the evaluation of the Public School Choice Initiative (PSCI) in Los Angeles Unified School District (Strunk et al., 2016). Using data across three cohorts of PSCI schools, Strunk and colleagues (2016) reported variation across cohorts: null effects on student achievement in Cohort 1 schools that began implementing PSCI reforms in 2010-11, positive effects on Cohort 2 schools that began in 2011-12, and negative effects on Cohort 3 schools that began in 2012-13. Strunk and colleagues (2016) could not extend their evaluation beyond three cohorts because PSCI was terminated after 2012-13; however, their discussion of program maturation effects provided a rich understanding of how early implementation hurdles diminished potential effects

Long Term Outcomes of Turnaround

in Cohort 1, while abrupt changes in program structure likely led to negative effects in Cohort 3. A growing number of evaluations have reported maturation effects across cohorts (Berends et al., 2002; Burns et al., 2023; Dragoset et al., 2019; Papay et al., 2022; Pham et al., 2020, 2024; Zimmer et al., 2017).

In contrast, dosage effects capture the impact on students who have attended reform schools over multiple years. Students who attend reform schools over a longer time period will have received a higher “dose” of the interventions, and some researchers have pointed out that reforms have a better chance of success if students experience increased supports for longer (Peurach & Neumerski, 2015). Abdulkadiroğlu and colleagues (2016) provide an example of dosage effects in their study of reforms that closed low-performing traditional public schools and re-opened them as charter schools. Overall, they found positive effects on student achievement from the charter conversions. However, when they examined students exposed to one, two, three, and four years of reforms following the charter conversion, they found that effects were larger for students after 1 or 2 years of exposure to reforms, relative to students exposed to three or more years of reforms. Thus, their examination of dosage effects at the student-level highlighted diminishing returns to the charter conversion model. Within the school reform literature, dosage effects are relatively common, though the vast majority of studies only report dosage through 2 - 4 years (Dee & Dizon-Ross, 2019; Henry & Harbatkin, 2020; Papay & Hannon, 2018; Schueler et al., 2017; Sun et al., 2017). An important area of need is for future studies to estimate dosage effects over a longer time period.

After a turnaround program’s core interventions have been completed, institutionalized effects occur at the school level and refer to whether school operations have changed in lasting ways without ongoing support from external actors. Institutionalized effects are relevant because

Long Term Outcomes of Turnaround

an implicit goal of many reform models is to improve organizational characteristics (e.g., culture, climate, practices, shared expectations) in low-performing schools so these schools can continue effectively serving students after external interventions and resources are no longer available.

Sun and colleagues (2021) provide an illustrative example of institutionalized effects. Using data from Boston, San Francisco, North Carolina, and Washington, the authors examined schools that received SIG funding to implement reforms. They find that SIG funds improved student achievement during the three years when the SIG reforms were actively implemented. After the SIG reforms ended, Sun and colleagues (2021) found that positive student achievement effects in these schools diminished but did not completely fade away and continued to be positive four years after the reforms ended. The authors explain that schools implementing more dramatic reforms later demonstrated a greater ability to maintain positive effects after the SIG reforms ended. Estimates of institutionalized effects are much less common than either maturation or dosage effects, but some studies of earlier whole-school reforms models have documented diminished institutionalized effects after active supports ended (e.g., Bifulco et al., 2005).

In contrast to the three previous types of long-term effects, this paper focuses on the fourth category of long-term effect, which we call persistent/latent effects at the student-level. Persistent/latent effects focus on how students fare after they leave reform schools. When reforms produce positive short-term effects that students continue to experience after leaving the reform school, we consider this a persistent long-term effect. On the other hand, it is also possible that a reform model does not produce short-term effects, but the latent effect becomes apparent after students leave the reform school. For example, reforms that focus on building foundational academic skills may not produce immediate short-term effects but may improve academic gains later, after students have had time to master the skills. Although persistent/latent

Long Term Outcomes of Turnaround

effects are rarely studied, they are arguably more meaningful than short-term effects, because these persistent/latent effects would show that the reforms successfully changed students' long-run educational trajectories in consequential ways, even after students leave the reform school. There are very few examples of persistent/latent effects in the school reform literature. One example involves studies of the RSD in New Orleans. Examining post-secondary outcomes after students leave RSD schools, Harris & Larsen (2023) find that students are more likely to attend college, and Glenn & Harris (2020) found that students are more likely to attend higher-quality colleges. Outside of the school reform literature, studies of persistent/latent effects are more common in the pre-school fadeout or convergence literature, which examine students outcomes after they leave pre-school (e.g., Bassok et al., 2019; Lipsey et al., 2018). For example, Lipsey and colleagues (2018) find that the positive effects of pre-school on student achievement largely disappear by the end of kindergarten and are even reversed by 2nd and 3rd grade. These studies of the persistent/latent effects of pre-school have had a substantial impact on current academic and policy conversations around pre-school. We believe other areas of educational research would benefit from similar attention to persistent/latent effects, and by differentiating persistent/latent effects from other long-term effects, our conceptual framework is intended to help increase attention to persistent/latent effects both within and outside of the school reform literature.

We recognize that studies of long-term effects may comprise some mixture or overlap between these four types of effects; however, scholarship on long-term effects would benefit from clearer language to facilitate communication on what types of effects are being reported. Better nomenclature will also clarify future comparisons of long-term effects across different contexts and reform models. Overall, this conceptual framework allows us to better clarify differences between our work and other studies of long-term effects of school turnaround which

Long Term Outcomes of Turnaround

have examined program maturation effects (Strunk et al., 2016), dosage effects (Abdulkadiroğlu et al., 2020), and institutionalized effects (Sun et al., 2021). However, we reiterate that persistent/latent effects of whole-school interventions are not well explored and should become a more central focus across a wide range of educational evaluations.

Methods

Data. Data are provided by TDOE and managed by the Tennessee Education Research Alliance. The administrative datasets are available for students in all Tennessee public schools. These data contain rich student characteristics and outcomes such as test scores, disciplinary actions, whether students graduate, gender, race, eligibility for free or reduced-price meals (FRPM), multilingual learner status (ML) and special education eligibility (SPED). Data are available from 2006-07 through 2019-20, except for test scores which are only available through 2018-19. Even though non-test score outcomes are available in 2019-20, we primarily use only data through 2018-19 because the COVID-19 pandemic may have affected students' educational outcomes in spring 2020 in ways that are still not well understood. In an auxiliary analysis, we show that our results are robust when we add data on all available outcomes in 2019-20 (Appendix Table A9).

Sample. Our goal is to estimate the long-term effect of turnaround reforms implemented in middle school on student outcomes in high school. Results in high school are particularly valuable for this study because high school outcomes like graduation are consequential for students' college and career options. Thus, we focus on turnaround middle schools because they are most proximate to outcomes measured while attending high school, which makes for a more straightforward definition of the treatment. To construct our sample, we use school feeder patterns to identify the middle school that students are assigned to attend when they complete the

Long Term Outcomes of Turnaround

final grade offered in their elementary school. Using these middle school assignments, we restrict our analytic sample to include only students who are assigned to an ASD, iZone, or comparison middle school. This sample allows us to estimate ITT estimates based on students' assigned middle school. We use this ITT sample instead of a treatment-on-the-treated (TOT) sample of students who actually attend ASD or iZone middle schools to account for student-level selection. Student-level selection bias could occur if choosing to attend a turnaround school is endogenous and related to unobserved factors that also affect student outcomes. This possibility is mitigated to some degree because Tennessee's turnaround schools are residentially assigned, not schools of choice. Nevertheless, parents could make a residential move that would allow their children to attend a non-turnaround school to shield their children from reform-induced disruptions. If these parents also tend to invest more resources in education such that their children are likely to have better schooling outcomes, their choice to avoid turnaround schools could negatively bias our estimate of the long-term reform effects. Our ITT approach addresses this issue by retaining all students who are assigned to ASD or iZone middle schools, regardless of whether they ultimately attend their assigned school. Below, we show that residential assignment to an ASD, iZone, or comparison school strongly predicts attendance in the assigned school (Appendix Table A3).

Thus, our preferred sample compares only students assigned to ASD or iZone middle schools with students assigned to similarly low-performing priority middle schools that receive no turnaround interventions. In auxiliary analyses, we restrict this preferred sample to include only students who can be observed attending all grades offered at an ASD, iZone, or comparison middle school and instrument the indicator for attending a school with the indicator for whether students were assigned to the school. This sample of students who attend all grades at the same

Long Term Outcomes of Turnaround

middle school allows us to avoid issues from students who move and experience the effects of different middle schools, but it is also a much smaller group of students than our preferred ITT sample: about one-third of students who ever attend a priority middle school in Tennessee attend all grades at that middle school. Results from this more restricted sample of students who attend all grades at a turnaround middle school lead us to similar conclusions, though significance levels vary between the two samples (see Appendix B). Finally, given available data, we can only examine students who are observed in a Tennessee public high school. Across all years, about 11 percent of students who are assigned to a priority middle school are not observed in any Tennessee high school. To examine potential bias from these missing students, we estimate the relative likelihood of attending a public high school in Tennessee between students assigned to ASD, iZone, or comparison schools. This analysis finds no significant differences (Appendix Table A10).

Table 2 shows a visual representation of our sample including the number of years we can observe outcome measures in high school for each student cohort. The table shows the four cohorts of students in turnaround middle schools that are included in our analysis,² with the years in which they were assigned to an ASD or iZone middle school in gray and their high school years observed in our data in white.³ For example, the first cohort of students would have started attending middle school in 2012-13 (the first year of turnaround reforms in Tennessee), attended middle school through 2014-15, and entered high school in 2015-16. These students are included in our analysis only between 2015-16 and 2018-19 when their outcomes are available in high school. Note that graduation from high school is primarily available only for the first cohort because most students spend four years in high school before graduating.

Long Term Outcomes of Turnaround

Part of our primary identification strategy relies on the DiD assumption of parallel trends in outcomes, which we show evidence to support further below. Although this assumption does not require ASD, iZone, and comparison schools to have similar baseline characteristics (only parallel *trends*), evidence of comparable baseline characteristics further supports our use of non-ASD, non-iZone priority schools as a valid counterfactual. Appendix Table A1 shows descriptive means for student demographics and each outcome of interest,⁴ within ASD, iZone and comparison middle schools before any reforms began in 2012-13. The table also shows results from *t*-tests comparing ASD versus comparison schools and iZone versus comparison schools. These tests show similar mean characteristics between all three groups of priority schools. The only statistically significant difference suggests that a larger proportion of students in iZone middle schools were white (6 percent) relative to comparison schools (2 percent). This difference makes sense in the Tennessee context because most priority schools are located in Memphis where most students are not white, but iZones were formed in other districts across the state, and schools in these districts tend to serve more white students. Nevertheless, a difference of 4 percentage points is quite modest, and we control for student race in all models.

Measures. We examine four categories of student outcomes in high school: test scores, attendance, discipline, and drop out/graduation. For test scores, we include scores on both the ACT and on state-required end-of-course (EOC) exams. For the ACT, we use only composite scores for students from the first time they take the exam to avoid bias from differential retake. Since the ACT is a norm-referenced test, we keep these scores on their original scale ranging from 1 to 36. Tennessee's end-of-course exams are administered in seven subjects: English I, English II, English III, Algebra I, Algebra II, Biology, and Chemistry. Given these different

Long Term Outcomes of Turnaround

courses, we standardize the EOC scores statewide by course, year, and semester to have a mean of zero and unit variance.

Attendance outcomes include the student's annual attendance rate and an indicator for chronic absenteeism. The attendance rate is the proportion of enrolled instructional days in which the student attended school, and chronic absenteeism is an indicator that equals one if students miss more than 10 percent of enrolled instructional days – the definition used in Tennessee's accountability policies (TDOE, 2021). Disciplinary actions are measured with indicators for whether the student commits a zero-tolerance disciplinary action and whether the student is expelled. Tennessee's zero-tolerance policy for offenses include possessing and/or using drugs, possessing a firearm, staff battery, and bullying. We chose zero-tolerance actions and expulsion as disciplinary outcomes because they are major offenses that would likely be recorded regardless of individual school or district discipline policies. Finally, we use state graduation records to code an indicator for whether the student drops out of school in each year and an indicator for whether the student receives a high school diploma in any year. Besides these outcomes, we also code whether students are assigned to ASD, iZone, or comparison middle schools based on the feeder pattern of their elementary school. Finally, we include a set of demographic student and school characteristics as covariates, as described below.

Analytic Method. Our analytic approach is a DiD model that estimates the ITT effect of assignment to a turnaround middle school. The DiD model allows us to address school-level selection, which could occur if schools chosen for either the ASD or an iZone differ systematically from non-turnaround priority schools prior to implementing any reforms. For example, if schools chosen for turnaround reforms have comparably better systems to improve

Long Term Outcomes of Turnaround

long-term student outcomes (e.g., more effective school leaders), these pre-existing school-level differences could positively bias our estimate of the long-term reform effects.

As a starting point, we model outcomes, y , for student i in high school s and year t :

$$y_{ist} = \beta_0 + \beta_1 ASD_i + \beta_2 iZone_i + \beta_3 BeganAfter_i + \beta_4 ASD_i * BeganAfter_i + \beta_5 iZone_i * BeganAfter_i + \beta_6 X_{it} + \beta_7 S_{st} + \theta_s + \phi_t + \varepsilon_{ist} \quad (1)$$

Equation 1 includes indicators for students who were assigned to middle schools that ever implement turnaround (ASD_i or $iZone_i$) and an indicator for whether the student began middle school after reforms were put into place ($BeganAfter_i$). It is important to emphasize that our data structure is closer to repeated cross-sections than a panel because the students who attend turnaround middle school before and after reforms began are different individuals. Appendix Figure A2 illustrates our data structure. The figure shows that we examine students records only in the years when they are in high school (e.g., grades 9 through 12). For outcomes such as EOC test scores, we may observe a student across several years if they take different EOCs each year of high school, but for some outcomes such as ACT scores, we observe each student only once in the year when they take the ACT for the first time. This repeated cross-section structure means that our DiD model compares students who were assigned to turnaround middle schools before reforms began with *different* students who were assigned to these middle schools in the years after reforms began, all relative to students assigned to comparison middle schools. Thus, the ASD_i and $iZone_i$ indicators vary between students but not across time. Likewise, the $BeganAfter_i$ indicator is also time invariant because regardless of what year(s) we observe the student in high school, this indicator equals one as long as the student began middle school after turnaround reforms had been implemented.

Long Term Outcomes of Turnaround

Additionally, Equation 1 includes vectors of student (X_{it}) and high school (S_{st}) characteristics as covariates. X_{it} includes the student's gender, race, eligibility for FRPM, ML status, and SPED status. S_{st} includes high school characteristics: proportion non-white students, proportion FRPM, proportion ML, and proportion SPED. Equation 1 also includes high school (θ_s) and year (ϕ_t) fixed effects and a residual error term (ε_{ist}) that includes unobserved factors affecting student outcomes, including measurement error. The high school fixed effect allows us to estimate within-high school effects, comparing students in a high school who attended ASD or iZone turnaround middle schools with other students in the same high school who attended a comparison middle school.⁵ Therefore, any effects a high school has on these outcomes should be common across both treatment and comparison groups, allowing us to isolate a middle school (treatment) effect from the high school effect. Additionally, because a few of the high schools attended by the treated students in our study are themselves ASD or iZone schools (3 ASD and 8 iZone high schools), we also estimate effects on a sample that excludes these ASD and iZone high schools and find similar results.

The coefficients of interest in Equation 1 are interpreted as the post-reform minus pre-reform difference in outcomes for students assigned to ASD (β_4) or iZone (β_5) middle schools, relative to the same difference for students assigned to priority middle schools that never implemented turnaround reforms. For tractability, we estimate linear probability models using OLS for binary outcomes (i.e., chronic absenteeism, zero tolerance action, expulsion, dropout, and graduation), so these coefficients can be interpreted as differences in probabilities.⁶ Assuming that pre-reform trends in student outcomes are parallel between students assigned to turnaround schools versus students assigned to comparison schools—and that these trends would continue post-reform—the DiD model in Equation 1 addresses the concern that schools chosen

Long Term Outcomes of Turnaround

for turnaround are systematically different from comparison schools even in the absence of reforms.

Recent developments in the DiD literature have extended the canonical model shown in Equation 1 to staggered DiD setups that address bias from potentially heterogeneous treatment effects across cohorts receiving turnaround reforms at different points in time. Because five different cohorts of ASD and iZone schools began reforms in each year between 2012-13 and 2016-17, results from Equation 1 are potentially biased from this staggered reform adoption if the long-term effects of turnaround are heterogeneous across the different ASD/iZone cohorts. To address this issue, we follow methods proposed by Callaway & Sant’Anna (2021), hereafter the CS approach. The CS approach begins by estimating separate effects for each treatment cohort in each year compared only to never-treated (or not yet treated) students to avoid problematic comparisons with already-treated students (Goodman-Bacon, 2018). We use only never-treated students who attend comparison middle schools as the control group and rely on the doubly robust DiD estimator (i.e., regression adjustment and inverse probability weighting) from Sant’Anna and Zhao (2020) to obtain each of these cohort-year specific effects. For reporting, we aggregate these separate cohort-year estimates using a simple weighted average of each cohort relative to its frequency in the treated population. We also report event-study parameters that are the weighted average effects for students who were assigned to ASD/iZone middle schools in each year after the school began reforms. For full transparency, we report both results from a canonical DiD model and the CS approach, which we implement using the *csdid* package in Stata (Rios-Avila et al., 2022). All models cluster standard errors at the school level.

Some researchers have argued that controlling for pre-treatment, or lagged, outcomes can lead to comparison units that have uncharacteristically low outcomes in the pre-treatment period,

Long Term Outcomes of Turnaround

which leads to bias from mean reversion (Daw & Hatfield, 2018). Therefore, our primary results using both the canonical and CS estimators do not include a pre-treatment, or lagged, outcome. This approach simplifies our interpretation of the results as effects on the level of each outcome. However, researchers have also found evidence that controlling for lagged outcomes measured prior to treatment can help reduce bias (Wilkins, 2018). Therefore, Appendix Table A4 shows auxiliary analyses that use the CS estimator with controls for lagged outcome measures. To avoid endogeneity from a covariate that may be affected by turnaround interventions in middle school, results in Appendix Table A4 are from a model that includes pre-treatment lags measured in the year prior to students attending middle school.⁷ Results from models that control for pretreatment lagged outcome measures are interpreted as effects on gains in each outcome, rather than levels. Results from models that control for lagged outcome measures align with our main conclusions, though the coefficients that are somewhat larger in magnitude.

To show descriptive evidence supporting the parallel trends assumption, Figure 2 uses our full ITT sample to graph trends in average EOC test scores on the y-axis for students who started middle school in each year before and after the school began turnaround reforms. Panels A-C in Figure 2 graphs trends in reading, math, and science EOC scores, with years on the x-axis centered such that year 0 is the year just before schools begin turnaround and year one is the first year of reforms. Although scores fluctuate from year-to-year, the figure shows little evidence of diverging trends in the EOC scores between students assigned to ASD, iZone, and comparison middle schools in years prior to reforms. In the post-reform years, we again do not see divergent trends in EOC scores between students assigned to ASD and iZone middle schools, relative to students assigned to comparison middle schools. Also, the figures suggest that, prior to reforms, students assigned to ASD middle schools had average EOC scores that were slightly higher than

Long Term Outcomes of Turnaround

the scores among students assigned to iZone and comparison middle schools. However, after reforms began, EOC scores for students assigned to ASD middle schools are no longer distinguishable from EOC scores among students assigned to iZone and comparison middle schools. To conserve space, we provide plots for the other outcomes of interest in Appendix C, all of which support parallel trends between ASD, iZone, and comparison schools. Below, we discuss formal statistical tests for parallel trends (Figure 3).

Results

Descriptive Results

Table 3 shows average characteristics of students assigned to ASD, iZone, and comparison middle schools in our ITT sample. For ASD and iZone schools, we show student characteristics before and after the school began implementing reforms. For comparison schools that never implement any turnaround interventions, we compare means from before and after 2012-13, the year when Tennessee first began operating turnaround schools. Tennessee's priority middle schools primarily serve Black students, ranging from 82 to 93 percent, depending on the turnaround model and years. Most students are eligible for free-or-reduced price meals (73 to 84 percent), and a small proportion are multilingual learners (1 to 3 percent). These student characteristics reflect the fact that most of Tennessee's priority schools serve historically Black communities in Memphis. Moreover, the table shows that student characteristics of ASD, iZone, and comparison middle schools do not substantially change after reforms began.

Table 3 also shows descriptive averages of student outcomes in high school. The table shows that students assigned to ASD middle schools have higher average EOC scores than students assigned to comparison schools in the years before reforms began. However, in the years after reforms began, students assigned to ASD middle schools have average EOC scores

Long Term Outcomes of Turnaround

that are lower than those assigned to comparison middle schools. This suggests that cohorts assigned to ASD middle schools post-reforms performed worse on their high school EOCs than the cohorts prior to reforms. However, students assigned to iZone middle schools post reforms had EOC test scores that were either similar or slightly higher than the pre-reform cohorts. While it is worth noting the relative achievement levels, what is more important to the validity of our DiD model is the pretreatment trends, which are largely parallel (Figure 2 and Appendix C). Nevertheless, Table 3 suggests that EOC scores largely do not improve for students assigned to ASD and iZone middle schools after reforms began relative to the cohorts before reforms were put in place. Likewise, none of the other long-term outcomes of interest meaningfully improved when comparing the pre- and post-reform cohorts of students assigned to turnaround middle schools. In fact, we observe a decrease in ACT scores, an increase in the chronic absenteeism rate, and a decrease in the graduation rate when we compare pre- and post-reform cohorts assigned to ASD middle schools.

Difference-in-Differences Results

Before turning to the DiD results, Appendix Table A3 shows that assignment to an ASD or iZone middle school strongly predicts whether attend that middle school. The table estimates our full DiD model (Equation 1) on whether the student attends a school that is (1) ever an ASD school, (2) an ASD school in the years after reforms began, (3) ever an iZone school, and (4) an iZone school in the years after reforms began. The results show that assignment to an ever ASD middle school strongly and significantly predicts attendance in an ever ASD middle school, and assignment to an ASD middle school in the years after the school began reforms significantly predicts attendance in an ASD school, post-reforms. The same is true for assignment and attendance in iZone schools. The results suggest that assignment to an ASD or iZone middle

Long Term Outcomes of Turnaround

school increases the probability that students will attend these schools by about 30 percentage points. These results support using indicators for assignment to ASD or iZone middle schools as a reduced-form ITT estimate, which can be rescaled to calculate TOT effects.

Table 4 shows results from estimating our canonical DiD model. The coefficients of interest are the DiD effect from interacting an indicator for assignment to an ASD or iZone middle school with an indicator for beginning middle school after reforms are put into place. Overall, the table shows primarily null long-term effects in high school for students who attended either ASD or iZone middle schools, and most of the coefficients are small in magnitude and precisely estimated.⁸ Similar to the descriptive results, DiD estimates for students assigned to ASD schools suggest negative effects, but the coefficients are not statistically significant for test scores, disciplinary actions, or dropout/graduation. The only outcomes where we observe a significant result is a 4 percentage point increase in the probability of being chronically absent. We also observe a potential decrease in attendance rate, but the result is modest (1 percentage point) and only marginally significant. For the iZones, we find no significant result on any outcome of interest, and most coefficients are very nearly zero.

Table 5 shows DiD results based on the CS approach to account for staggered adoption of turnaround reforms across five cohorts of schools. In alignment with results from the canonical DiD specification, we find no statistically significant effect on any test score, attendance, discipline, or dropout/graduation outcomes. Although the results are not statistically significant, we find that the DiD estimate for students assigned to ASD middle schools is consistently negative across all EOC subjects. Similarly, for students assigned to iZone middle schools, the coefficient on EOC math scores is negative and marginally significant. Similar results between the traditional specification and the CS approach suggest that effects were not heterogeneous

Long Term Outcomes of Turnaround

across different cohorts, which aligns with our understanding that across all five ASD and iZone cohorts, students experienced the same set of interventions, respectively.

Next, we use the CS approach to capture effects for students who began middle school in each year before and after turnaround reforms began, akin to an event-history model. Figure 3 plots each of these coefficients for EOC reading, math, and science scores. Similar plots for all other outcomes are in Appendix D. These event-history estimates allow us to formally test the parallel trends assumption by examining coefficients for students assigned to ASD or iZone schools prior to reforms. Figure 3 shows that coefficients in each of the pre-reform years are near zero and nonsignificant. We also conduct a joint significance test for all coefficients in the pre-reform years and find no significant differences. Furthermore, coefficients are not significant in the years after reforms began, and we do not observe strong trends across the years, except for potential trends downward in science EOC scores for students assigned to ASD middle schools. Year-by-year coefficients for the other outcomes are also primarily not significant (Appendix D).

Finally, the null results may be somewhat surprising for students assigned to iZone middle schools because previous research finds significant effects of iZone interventions on student achievement, at least while they are in the iZone school (Pham et al., 2020; Zimmer et al., 2017). Because results from the prior papers were averaged across all school levels, it is possible that effects for students in iZone middle schools were always statistically insignificant, and the positive effects reported in prior studies were driven by iZone elementary and high schools. To test this theory and better compare results in this paper with ASD/iZone effects on students while they were still in the turnaround school, we use methods similar to Pham et al. (2020) to estimate effects on the specific subsample of only middle schools. Results from these subsamples (Appendix Table A5) align with the previously reported findings: null ASD effects

Long Term Outcomes of Turnaround

and positive iZone effects. These results suggest that iZone middle schools had some positive short-term effects on student achievement while they were in middle school, but the positive effects do not carry over after students leave middle school.

Robustness Checks

A threat to the validity of our results could occur if ASD and iZone schools actively recruited or pushed out students who are systematically different from the types of students who would have attended these schools in the absence of turnaround. While this practice is unlikely because schools in both models are required to continue serving students in their local enrollment zones (TDOE, 2012), we empirically examine this possibility using characteristics of all students who enter and leave these middle schools in the post-turnaround years. Appendix Table A7 shows that the average rate of student in-migration and out-migration are similar in ASD, iZone, and comparison schools between 2012-13 and 2018-19. Additionally, Appendix Table A8 shows that the observed characteristics of incoming and outgoing students are nearly the same in all three groups of schools. These results suggest that our findings are not driven by differential student mobility into and out of turnaround middle schools. Finally, we test the robustness of our results by estimating models that remove all covariates. Appendix Tables A6 shows that our results are robust to this alternative specification.

Conclusion

While empirical literature on school reform has been growing in recent years (Redding & Nguyen, 2020; Schueler et al., 2021), long-term effects of reform are critically understudied. Indeed, our review of the literature found conceptually distinct categories of long-term effects that were previously commingled, which has muddled current academic and policy discourse between studies focusing on different types of long-term effects. In this study, we add clarity to

Long Term Outcomes of Turnaround

the study of long-term effects by developing a broadly applicable conceptual framework for delineating four categories of long-term effects: program maturation, institutionalized, dosage, and persistent/latent effects. Guided by this framework, we identified growing attention in the school reform literature to dosage (Abdulkadiroğlu et al., 2020), program maturation (Strunk et al., 2016), and institutionalized effects (Sun et al., 2019). However, we could find no study examining persistent/latent effects outside of the very distinct RSD model (Glenn & Harris, 2020; Harris & Larsen, 2023). Thus, in this study, we estimate persistent/latent effects within the context of Tennessee's much more generalizable statewide school reforms.

Overall, we find very little evidence of positive persistent/latent effects after students leave an ASD or iZone middle school. Results on high school test scores in reading, math, and science are negative and inconsistently significant for students assigned to ASD middle schools, which begins to suggest negative effects but will require additional studies to sufficiently test. We also find no significant changes in ACT scores nor did students experience any effect on the probability of graduating from high school. Results using the canonical DiD model suggests that students assigned to ASD middle schools have a significantly higher probability of being chronically absent in high school, but this result is not robust to using the CS approach to address staggered treatment timing. Long-term effects from students assigned to iZone middle schools are no more encouraging. None of our results are significant at the five percent level, and most are extremely small in magnitude (i.e., precisely estimated zeroes). The only coefficient that is marginally significant at the 10 percent level suggests that students assigned to iZone middle schools posted lower math EOC scores than students assigned to comparison middle schools. These null results across multiple student outcomes suggest that the turnaround reforms

Long Term Outcomes of Turnaround

implemented in Tennessee, which are similar to interventions used in many other states, are showing few signs of improving student outcomes after they leave the turnaround school.

Though only suggestive, evidence of negative effects on EOC scores from students attending ASD middle schools deserves further attention. Previous research found that the ASD reforms produced null short-term (Zimmer et al., 2017) and maturation effects (Pham et al., 2020), and our findings point toward a potentially detrimental latent effect from these reforms. Since the reforms likely disrupted the schools' staffing, culture, and daily operations when the principal and most teachers were replaced, part of the null or even negative effects in high school may be explained by these instabilities in students' instructional experiences. Previous research has shown that whole-school reforms requiring teacher and principal replacements can lead to high levels of instability in the school's daily routines and procedures (Malen et al., 2002; Rice & Malen, 2010). The theory of action for bold turnaround reforms implies that these disruptions could be worth the effort if they can overcome pre-existing barriers to improvement and are replaced with effective interventions and supports to quickly improve the school (Herman et al., 2008; Kutash et al., 2010); however, for the ASD, these disruptions were not met with effective interventions to support the improvement process.

The iZone model did improve student achievement while students attended the turnaround school (Pham et al., 2020; Zimmer et al., 2017).⁹ Yet, the positive iZone effects on achievement do not persist when these students reach high school. Most estimates are nearly zero and statistically insignificant, except a marginally significant negative effect on math test scores. These results are important to consider as the iZone model grows in prevalence in other districts throughout the country (Iyengar et al., 2017) because they suggest that Tennessee's iZone reforms achieved short-term gains but the gains faded after students left the turnaround school. A

Long Term Outcomes of Turnaround

likely explanation is that students in the lowest-performing schools continue to need support, and attending a few years at an improved iZone middle school is insufficient to change their educational trajectories. Our findings suggest that reform policies may need to be designed in a way that is connected across school levels to support students throughout their K-12 educational experience. Importantly, our results suggest that future reform models should put more thought into interventions that support persistent improvements. The emphasis in iZone schools was likely directed toward instructional improvements, but more comprehensive supports at the system level, including supporting students' social-emotional learning (Carr, 2021), investing in wraparound services (Hill, 2020), and soliciting support from the local community around the school (Anderson-Butcher et al., 2010; Oakes et al., 2017) may be needed for improved outcomes to persist.

This paper contributes novel insights to the literature on school reform by highlighting persistent/latent effects on students' long-run outcomes as an important but understudied aspect of turnaround. We find scant evidence of long-term positive effects from either of Tennessee's two turnaround models, but a critical avenue for future research in this area is an in-depth explanation of why the iZone model showed positive short-term effects that did not persist. Our results also suggest that systemic reforms (beyond individual school buildings) may be needed for positive effects on students to persist. In general, we emphasize the need for future evaluations of school reform to regularly examine, and clearly differentiate, multiple types of long-term effects. Ongoing research in this area will provide important evidence to help policymakers and educators to fully understand returns to investments in school reform.

Endnotes

1. One ASD and one iZone school was classified as an serving an “other” grade level (i.e., K-8 schools), but these schools were subsequently closed and not included in our analysis.
2. In auxiliary analyses that include outcomes in 2019-20, we are able to add a fifth cohort of students who began middle school in 2016-17 (and who then have outcomes in high school beginning in 2019-20).
3. For presentational clarity, the table shows students who attend middle school for three years (e.g., 6th through 8th grade), but the logic is similar for students attending middle schools with a different grade configuration.
4. Table 2 does not include ACT scores and graduation outcomes because neither measure is available for students in middle school.
5. To provide evidence of within-school variation in the students coming from ASD, iZone, or comparison middle schools, we note that of the 785 high school-by-year records in our sample, 651 (83%) contain some combination of students from ASD, iZone, and comparison middle schools. That is, 17% of the high school-by-year records contain only students from ASD, iZone, or comparison middle schools. The within-high school average percent of students from ASD, iZone, and comparison middle schools are 24.7%, 50.8% and 24.5%, respectively.
6. We reach similar conclusions when using logistic regression for the binary outcomes.
7. Note that we can only add pre-treatment lags for outcomes that are measurable before high school (i.e., end-of-grade test scores in the same subject, attendance rate, chronic absenteeism, zero tolerance actions, and expulsion). We cannot add lags to models that examine effects on high school graduation or ACT scores.
8. Although standard errors appear to be somewhat larger for ACT scores than for other outcomes, note that ACT scores are on their original 1-36 point scale instead of standard deviation units.
9. Since these previous studies examined all iZone schools, pooling together all levels, it is possible that the overall results may have obscured less positive results in iZone middle schools. Therefore, we replicate methods from Pham and colleagues (2020) and estimate the effect of reforms on student test scores but restrict the sample to only middle school students. Consistent with the prior published effects on all iZone schools, the iZone effect in only middle schools is positive and significant ($\beta = 0.09 SD$).

References

- Abdulkadiroğlu, A., Angrist, J. D., Hull, P. D., & Pathak, P. A. (2016). Charters without Lotteries: Testing Takeovers in New Orleans and Boston. *American Economic Review*, *106*(7), 1878–1920.
- Abdulkadiroğlu, A., Pathak, P. A., Schellenberg, J., & Walters, C. R. (2020). Do parents value school effectiveness? *American Economic Review*, *110*(5), 1502–1539.
- Aladjem, D. K., Birman, B. F., Orland, M., Harr-Robins, J., Heredia, A., Parrish, T. B., & Ruffini, S. J. (2010). Achieving Dramatic School Improvement: An Exploratory Study. A Cross-Site Analysis from the Evaluation of Comprehensive School Reform Program Implementation and Outcomes Study. *Office of Planning, Evaluation and Policy Development, US Department of Education*.
- Anderson-Butcher, D., Lawson, H. A., Iachini, A., Flaspohler, P., Bean, J., & Wade-Mdivanian, R. (2010). Emergent evidence in support of a community collaboration model for school improvement. *Children & Schools*, *32*(3), 160–171.
- Bassok, D., Gibbs, C. R., & Latham, S. (2019). Preschool and Children’s Outcomes in Elementary School: Have Patterns Changed Nationwide Between 1998 and 2010? *Child Development*, *90*(6), 1875–1897.
- Baum, C. F., Schaffer, M. E., & Stillman, S. (2007). Enhanced routines for instrumental variables/generalized method of moments estimation and testing. *The Stata Journal*, *7*(4), 465–506.
- Berends, M., Bodilly, S. J., & Kirby, S. N. (2002). *Facing the challenges of whole-school reform: New American Schools after a decade*. Rand Corporation.
- Bifulco, R., Duncombe, W., & Yinger, J. (2005). Does whole-school reform boost student performance? The case of New York City. *Journal of Policy Analysis and Management*, *24*(1), 47–72.
- Booker, K., Sass, T. R., Gill, B., & Zimmer, R. (2011). The effects of charter high schools on educational attainment. *Journal of Labor Economics*, *29*(2), 377–415.
- Burns, J., Harbatkin, E., Strunk, K. O., Torres, C., Mcilwain, A., & Frost Waldron, S. (2023). The Efficacy and Implementation of Michigan’s Partnership Model of School and District Turnaround: Mixed-Methods Evidence From the First 2 Years of Reform Implementation. *Educational Evaluation and Policy Analysis*, *45*(4), 622–654.
- Callaway, B., & Sant’Anna, P. H. (2020). Difference-in-differences with multiple time periods. *Journal of Econometrics*.

Long Term Outcomes of Turnaround

- Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2), 200–230.
- 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.
- Carroll, D. (2014). *Tennessee announces “priority” and “focus” schools*. <https://www.wrcbtv.com/story/26321066/tennessee-announces-priority-and-focus-schools>
- Carr, R. (2021). *Accountability in the Classroom: Using Social-emotional Learning to Guide School Improvement*. Rowman & Littlefield Publishers.
- Chiang, H. (2009). How accountability pressure on failing schools affects student achievement. *Journal of Public Economics*, 93(9), 1045–1057.
- Daw, J. R., & Hatfield, L. A. (2018). Matching and regression to the mean in difference-in-differences analysis. *Health Services Research*, 53(6), 4138–4156.
- Dee, T. (2012). *School turnarounds: Evidence from the 2009 stimulus*. National Bureau of Economic Research.
- Dee, T., & Dizon-Ross, E. (2019). School Performance, Accountability, and Waiver Reforms: Evidence from Louisiana. *Education Evaluation and Policy Analysis*, 41(3), 316–349.
- Dee, T. S., & Wyckoff, J. (2015). Incentives, Selection, and Teacher Performance: Evidence from IMPACT. *Journal of Policy Analysis and Management*, 34(2), 267–297.
- Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2020). COVID-19 and learning loss—Disparities grow and students need help. *McKinsey & Company, December*, 8, 6–7.
- Dragoset, L., Thomas, J., Herrmann, M., Deke, J., James-Burdumy, S., & Lee Luca, D. (2019). The Impact of School Improvement Grants on Student Outcomes: Findings From a National Evaluation Using a Regression Discontinuity Design. *Journal of Research on Educational Effectiveness*, 12(2), 215–250.
- Duncan, A. (2009). *Turning around the bottom 5 percent*. U.S. Department of Education.
- Edgerton, A. K., & Desimone, L. M. (2019). Mind the Gaps: Differences in How Teachers, Principals, and Districts Experience College- and Career-Readiness Policies. *American Journal of Education*, 125(4), 593–619.
- Firestone, W. A., & Herriott, R. E. (1982). Prescriptions for effective elementary schools don't fit secondary schools. *Educational Leadership*, 40(3), 51–53.

Long Term Outcomes of Turnaround

- Glenn, B., & Harris, D. N. (2020). Climbing the College Ladder? The Effects of New Orleans School Reforms on College Outcomes and the Quality of Colleges That Students Attended. EdWorkingPaper No. 20-339. *Annenberg Institute for School Reform at Brown University*.
- Goodman-Bacon, A. (2018). *Difference-in-differences with variation in treatment timing*. National Bureau of Economic Research.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2), 254–277.
- 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.
- Harris, D. N., & Larsen, M. F. (2023). Taken by storm: The effects of Hurricane Katrina on medium-term student outcomes in New Orleans. *Journal of Human Resources*, 58(5), 1608–1643.
- Henry, G. T., & Harbatkin, E. (2020). The Next Generation of State Reforms to Improve their Lowest Performing Schools: An Evaluation of North Carolina’s School Transformation Intervention. *Journal of Research on Educational Effectiveness*, 13(4), 702–730.
- Henry, G. T., Zimmer, R., Attridge, J., Kho, A., & Viano, S. (2014). *Teacher and student migration in and out of Tennessee’s Achievement School District*. Tennessee Consortium on Research, Evaluation & Development.
https://peabody.vanderbilt.edu/TERA/files/ASD_Teacher_Student_Migration.pdf
- Herman, R., Dawson, P., Dee, T., Greene, J., Maynard, R., Redding, S., & Darwin, M. (2008). *Turning Around Chronically Low-Performing Schools. IES Practice Guide*. National Center for Education Evaluation and Regional Assistance.
- Hill, R. A. (2020). Wraparound: A Key Component of School-Wide Culture Competence to Support Academics and Socio-Emotional Well-Being. *Peabody Journal of Education*, 95(1), 66–72.
- Iyengar, N., Lewis-LaMonica, K., & Perigo, M. (2017). *School District Innovation Zones: A New Wave of District-led Efforts to Improve Economic Mobility* (p. 46). The Bridgespan Group.
- Jochim, A. (2016). Measures of Last Resort: Assessing Strategies for State-Initiated Turnarounds. Linking State and Local School Improvement. *Center on Reinventing Public Education*.
- Kogan, V., & Lavertu, S. (2021). The COVID-19 pandemic and student achievement on Ohio’s third-grade English language arts assessment. *Ohio State University*.

Long Term Outcomes of Turnaround

- Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, J. (2020). Projecting the Potential Impact of COVID-19 School Closures on Academic Achievement. *Educational Researcher*, 49(8), 549–565.
- Kutash, J., Nico, E., Gorin, E., Rahmatullah, S., & Tallant, K. (2010). The school turnaround field guide. *Wallace Foundation*.
- Lipsey, M. W., Farran, D. C., & Durkin, K. (2018). Effects of the Tennessee Prekindergarten Program on children’s achievement and behavior through third grade. *Early Childhood Research Quarterly*, 45, 155–176.
- Malen, B., Croninger, R., Muncey, D., & Redmond-Jones, D. (2002). Reconstituting schools: “Testing” the “theory of action.” *Educational Evaluation and Policy Analysis*, 24(2), 113–132.
- Oakes, J., Maier, A., & Daniel, J. (2017). Community Schools: An Evidence-Based Strategy for Equitable School Improvement. *National Education Policy Center*.
- Papay, J., & Hannon, M. (2018). *The Effects of School Turnaround Strategies in Massachusetts* [Working Paper].
- Papay, J. P., Kraft, M. A., & James, J. K. (2022). Operator versus Partner: A Case Study of Blueprint School Network’s Model for School Turnaround. *Journal of Research on Educational Effectiveness*, 15(2), 367–393.
- Peurach, D. J., & Neumerski, C. M. (2015). Mixing metaphors: Building infrastructure for large scale school turnaround. *Journal of Educational Change*, 16(4), 379–420.
- Pham, L. D., Henry, G. T., Kho, A., & Zimmer, R. (2020). Sustainability and maturation of school turnaround: A multi-year evaluation of Tennessee’s achievement school district and local innovation zones. *AERA Open*, 6(2).
- Pham, L. D., Matthews, G. F., & Drake, T. A. (2024). Beyond Prescriptive Reforms: An Examination of North Carolina’s Flexible School Restart Program. *Educational Evaluation and Policy Analysis*, 01623737231218735.
- Redding, C., & Nguyen, T. D. (2020). The relationship between school turnaround and student outcomes: A meta-analysis. *Educational Evaluation and Policy Analysis*, 42(4), 493–519.
- Rentner, D. S., Kober, N., & Frizzell, M. (2017). Planning for Progress: States Reflect on Year One Implementation of ESSA. *Center on Education Policy*.
- Rice, J. K., & Malen, B. (2010). *School Reconstitution as an Education Reform Strategy*. National Education Association.

Long Term Outcomes of Turnaround

- Rios-Avila, F., Callaway, B., & Sant'Anna, P. H. (2022). *csdid: Difference-in-differences with multiple time periods in stata*. Working paper, Boston University.
- Rockoff, J., & Turner, L. J. (2010). Short-run impacts of accountability on school quality. *American Economic Journal: Economic Policy*, 2(4), 119–147.
- Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *arXiv Preprint arXiv:2201.01194*.
- Rouse, C. E., Hannaway, J., Goldhaber, D., & Figlio, D. (2013). Feeling the Florida Heat How Low-Performing Schools Respond to Voucher and Accountability Pressure. *American Economic Journal: Economic Policy*, 5(2), 251–281.
- Sant'Anna, P. H., & Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, 219(1), 101–122.
- Sass, T. R., Zimmer, R. W., Gill, B. P., & Booker, T. K. (2016). Charter high schools' effects on long-term attainment and earnings. *Journal of Policy Analysis and Management*, 35(3), 683–706.
- Schueler, B. E., Asher, C. A., Larned, K. E., Mehrotra, S., & Pollard, C. (2021). Improving Low-Performing Schools: A Meta-Analysis of Impact Evaluation Studies. *American Educational Research Journal*, 00028312211060855.
- Schueler, B. E., & Bleiberg, J. F. (2021). Evaluating Education Governance: Does State Takeover of School Districts Affect Student Achievement? *Journal of Policy Analysis and Management*, pam.22338.
- 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.
- Strunk, K. O., Marsh, J. A., Hashim, A. K., Bush-Mecenas, S., & Weinstein, T. (2016). The impact of turnaround reform on student outcomes: Evidence and insights from the Los Angeles Unified School District. *Education Finance and Policy*, 11(3), 251–282.
- Sun, M., Kennedy, A. I., & Loeb, S. (2021). The Longitudinal Effects of School Improvement Grants. *Educational Evaluation and Policy Analysis*, 43(4), 647–667.
- Sun, M., Kennedy, A., & Loeb, S. (2019). The Longitudinal Effects of School Improvement Grants. *Paper Presented at the Annual Conference of the Association for Public Policy Analysis and Management*.

Long Term Outcomes of Turnaround

- 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.
- TDOE. (2018). *Every Student Succeeds Act: Building on Success in Tennessee*. https://www.tn.gov/content/dam/tn/education/documents/TN_ESSA_State_Plan_Approved.pdf
- Tennessee Department of Education. (2012). *ESEA Flexibility Request*. <https://www2.ed.gov/policy/eseaflex/approved-requests/tn.pdf>
- Tennessee Department of Education. (2021). *Chronically Out of School*. Tennessee Report Card. <https://reportcard.tnedu.gov/state/0/attendance>
- TN Press Release Center. (2012, August 14). *State Names 250 Schools to List of Priority, Focus Schools*. TNReport. <http://tnreport.com/2012/08/14/state-names-250-schools-to-list-of-priority-focus-schools/>
- US Department of Education. (2009). *The Race to the Top Begins: Remarks by Secretary Arne Duncan*.
- U.S. Department of Education. (2012). *Elementary and Secondary Education Act flexibility: Request*. <https://www2.ed.gov/policy/eseaflex/tn.pdf>
- Wilkins, A. S. (2018). To lag or not to lag?: Re-evaluating the use of lagged dependent variables in regression analysis. *Political Science Research and Methods*, 6(2), 393–411.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.
- 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.

Long Term Outcomes of Turnaround

Table 1. Number of Schools in the Sample by ASD, iZone, and Comparison Groups and by School Level

Cohort Number	First year of turnaround	Comparison				ASD				iZone			
		Elem.	Middle	High	All Levels	Elem.	Middle	High	All Levels	Elem.	Middle	High	All Levels
Cohort 1	2012-2013	28	19	15	65 ^a	3	3	0	6	7	6	0	13 ^b
Cohort 2	2013-2014	16	12	14	45	5	1	0	6	5	5	1	11
Cohort 3	2014-2015	23	15	12	59 ^c	3	0	2	5	1	0	3	4
Cohort 4	2015-2016	12	8	11	40	3	2	0	5	7	3	1	11 ^d
Cohort 5	2016-2017	11	6	4	26	1	2	1	4	0	0	3	3

Note. Elem stands for elementary school. The ASD and iZone rows are schools added from lists of priority schools released by TDOE in either 2012-2013 or 2014-2015. The comparison school rows represent the remaining priority schools that have not joined either the ASD or and iZone in each year. The number of comparison schools decreases each year as more schools join the ASD or an iZone, except in 2014-2015 when additional priority schools were added.

^a The number of comparison schools at all levels does not equal to the sum of elementary, middle, and high schools because there some comparison schools that do fit into an “other” level category (e.g., K-8 schools). The number of priority schools closed in each of the five cohorts are 1, 3, 8, 3, and 7 respectively.

^b Two 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.

^c The number of priority schools increases in 2014–2015, because Tennessee released a new list of priority schools in 2014.

^d Two 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 are not included in this study.

Long Term Outcomes of Turnaround

Table 2. Visual Representation of Each Cohort of Students in the Years they Attend Middle and High School including the Number of Years the Outcomes can be Measured

	2012-13	2013-14	2014-15	2015-16	2016-17	2017-18	2018-19
Cohort 1	Assigned to Priority Middle School	Assigned to Priority Middle School	Assigned to Priority Middle School	Test Scores Attendance Discipline Drop Out	Test Scores Attendance Discipline Drop Out	Test Scores Attendance Discipline Drop Out	Test Scores Attendance Discipline Drop Out Graduation
Cohort 2		Assigned to Priority Middle School	Assigned to Priority Middle School	Assigned to Priority Middle School	Test Scores Attendance Discipline Drop Out	Test Scores Attendance Discipline Drop Out	Test Scores Attendance Discipline Drop Out
Cohort 3			Assigned to Priority Middle School	Assigned to Priority Middle School	Assigned to Priority Middle School	Test Scores Attendance Discipline Drop Out	Test Scores Attendance Discipline Drop Out
Cohort 4				Assigned to Priority Middle School	Assigned to Priority Middle School	Assigned to Priority Middle School	Test Scores Attendance Discipline Drop Out

Note. For presentational clarity, the table shows students who attend middle school for three years (e.g., 6th through 8th grade), but the logic is similar for students attending middle schools with a different grade configuration.

Long Term Outcomes of Turnaround

Table 3. Descriptive Characteristics of Students who Attended ASD, iZone, or Comparison Middle Schools in the Years Before and After Turnaround Interventions Began

	ASD		iZone		Comparison	
	Before	After	Before	After	Before	After
Student Characteristics						
Female	0.50	0.50	0.49	0.49	0.49	0.49
FRPM	0.75	0.81	0.80	0.73	0.84	0.75
ML	0.02	0.02	0.01	0.02	0.01	0.03
SPED	0.15	0.18	0.17	0.17	0.16	0.16
Asian	0.01	0.01	0.01	0.01	0.00	0.00
Black	0.85	0.91	0.85	0.82	0.93	0.90
Latino/a/x	0.10	0.06	0.05	0.09	0.06	0.09
White	0.04	0.03	0.09	0.08	0.00	0.01
Other Race	0.001	0.001	0.001	0.001	0.001	0.001
Student Outcomes in High School						
Reading (Standardized)	-0.45 (0.99)	-0.60 (0.92)	-0.59 (1.13)	-0.52 (0.95)	-0.62 (1.07)	-0.52 (0.90)
Math (Standardized)	-0.51 (0.97)	-0.68 (0.94)	-0.60 (1.03)	-0.61 (0.97)	-0.63 (1.02)	-0.62 (0.94)
Science (Standardized)	-0.50 (1.10)	-0.58 (0.97)	-0.66 (1.20)	-0.54 (1.06)	-0.73 (1.20)	-0.54 (1.00)
ACT Score (1-36)	15.55 (3.47)	14.51 (2.67)	15.65 (3.82)	15.32 (3.64)	15.28 (3.02)	14.98 (3.05)
Attendance Rate	0.91 (0.11)	0.89 (0.12)	0.89 (0.13)	0.90 (0.13)	0.90 (0.12)	0.90 (0.12)
Chronic Absenteeism	0.29	0.36	0.35	0.34	0.32	0.32
Zero Tolerance Action	0.01	0.02	0.01	0.02	0.01	0.01
Expulsion	0.01	0.02	0.02	0.02	0.02	0.02
Dropout	0.02	0.02	0.03	0.01	0.03	0.01
High School Graduation	0.87	0.71	0.86	0.73	0.85	0.79
Student-Year Observations	21645	2466	29283	8506	11840	6855

Note. Standard deviations in parentheses. Students are only included in the sample if they either attended all grades at a priority (ASD, iZone, or comparison) middle school during the years before or after turnaround interventions were implemented.

Long Term Outcomes of Turnaround

Table 4. DiD Effect of Assignment to ASD or iZone Middle Schools on Outcomes in High School

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	EOC Reading	EOC Math	EOC Science	ACT Composite Score	Attendance Rate	Chronic Absenteeism (Binary)	Zero Tolerance (Binary)	Expulsion (Binary)	Dropout (Binary)	Graduate (Binary)
Treatment Indicators										
Assigned to ASD Middle School	0.071** (0.023)	0.059* (0.024)	0.114*** (0.024)	-0.076 (0.117)	0.008** (0.002)	-0.027** (0.010)	0.001 (0.002)	0.001 (0.002)	-0.007*** (0.002)	0.009 (0.009)
Assigned to iZone Middle School	-0.007 (0.025)	-0.008 (0.030)	0.013 (0.032)	-0.195 (0.128)	-0.000 (0.003)	0.002 (0.011)	0.001 (0.001)	0.002 (0.002)	-0.004+ (0.002)	0.006 (0.009)
Began Middle School After Turnaround Began	-0.000 (0.032)	0.032 (0.043)	0.111* (0.053)	-0.251+ (0.138)	0.015*** (0.003)	-0.065*** (0.011)	0.001 (0.002)	0.004* (0.002)	-0.014*** (0.003)	-0.140*** (0.017)
Assigned to ASD Middle School*Began After	-0.057 (0.045)	-0.075 (0.053)	-0.061 (0.079)	0.021 (0.179)	-0.009+ (0.005)	0.043* (0.021)	-0.004 (0.003)	-0.002 (0.003)	0.007 (0.004)	-0.031 (0.032)
Assigned to iZone Middle School*Began After	0.008 (0.034)	-0.008 (0.048)	0.027 (0.058)	0.086 (0.140)	0.001 (0.004)	-0.004 (0.015)	-0.001 (0.002)	-0.004 (0.003)	0.001 (0.003)	-0.047 (0.034)
Student Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R Squared	0.135	0.073	0.072	0.120	0.023	0.025	0.005	0.006	0.002	0.098
N	72472	57417	32409	10438	136959	137040	79220	79220	138609	24117

Note. + p<0.1, * p<0.05, ** p<0.01, *** p<0.001 All models include high school and year fixed effects. Standard errors in parentheses clustered at school level. All estimates are from an ITT sample of students assigned to ASD, iZone, or comparison middle schools based on their elementary school. Student characteristics included as controls include the student's gender, race, eligibility for FRPM, ML status, and SPED status. High school characteristics included as controls include proportion non-White students, proportion FRPM, proportion ML, and proportion SPED.

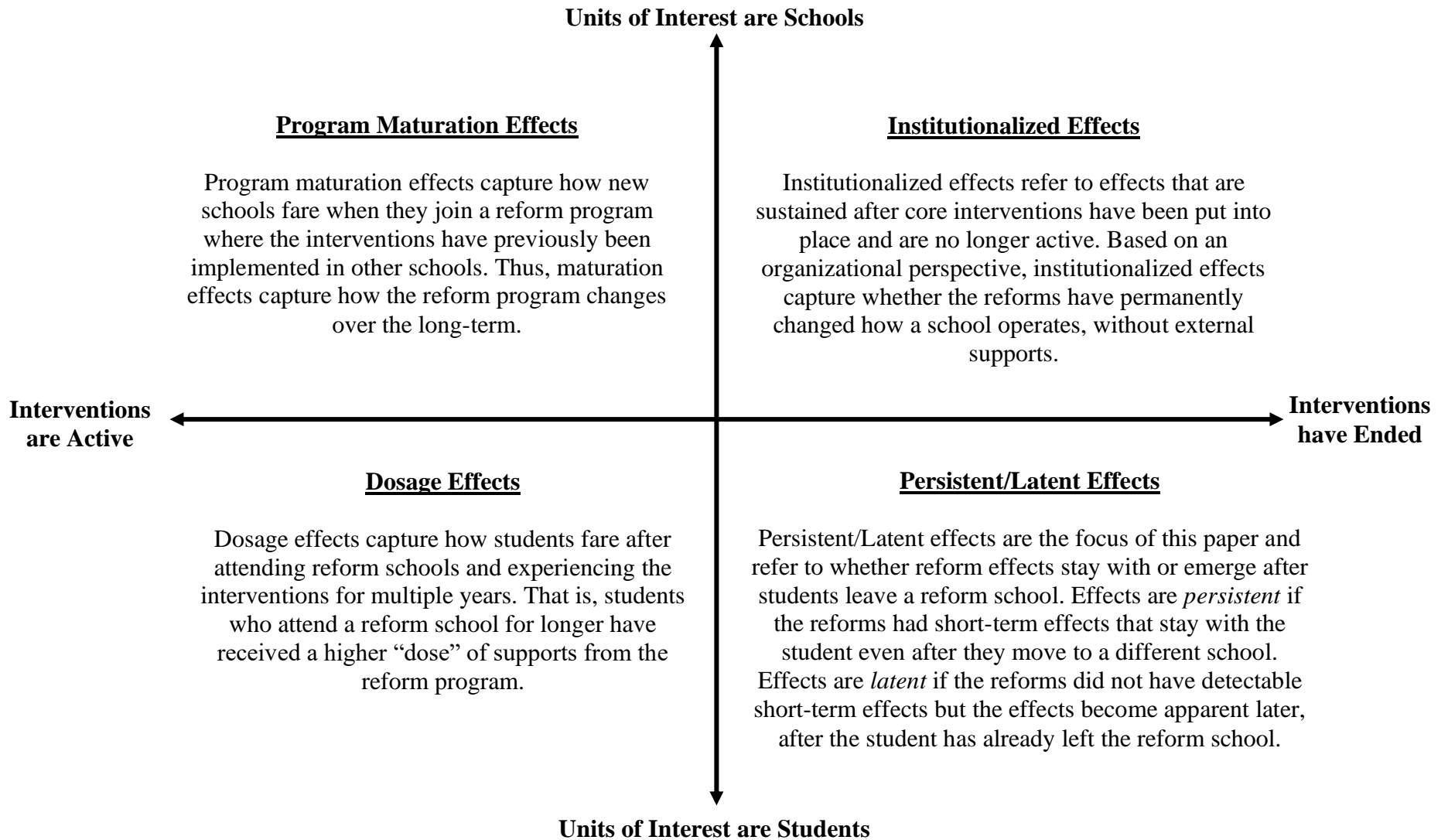
Long Term Outcomes of Turnaround

Table 5. DiD Effect of Assignment to ASD or iZone Middle Schools Using Callaway and Sant’Anna (2020) Estimator

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	EOC Reading	EOC Math	EOC Science	ACT Composite Score	Attendance Rate	Chronic Absenteeism (Binary)	Zero Tolerance (Binary)	Expulsion (Binary)	Dropout (Binary)	Graduate (Binary)
ASD Effect	-0.01 (0.09)	-0.07 (0.13)	-0.14 (0.31)	0.30 (0.32)	0.01 (0.06)	-0.07 (0.25)	-0.01 (0.01)	-0.01 (0.01)	0.02 (0.05)	-0.11 (0.08)
iZone Effect	0.02 (0.05)	-0.09+ (0.05)	-0.03 (0.11)	0.24 (0.27)	0.01 (0.01)	-0.02 (0.04)	0.01 (0.01)	0.00 (0.01)	-0.00 (0.01)	0.05 (0.03)

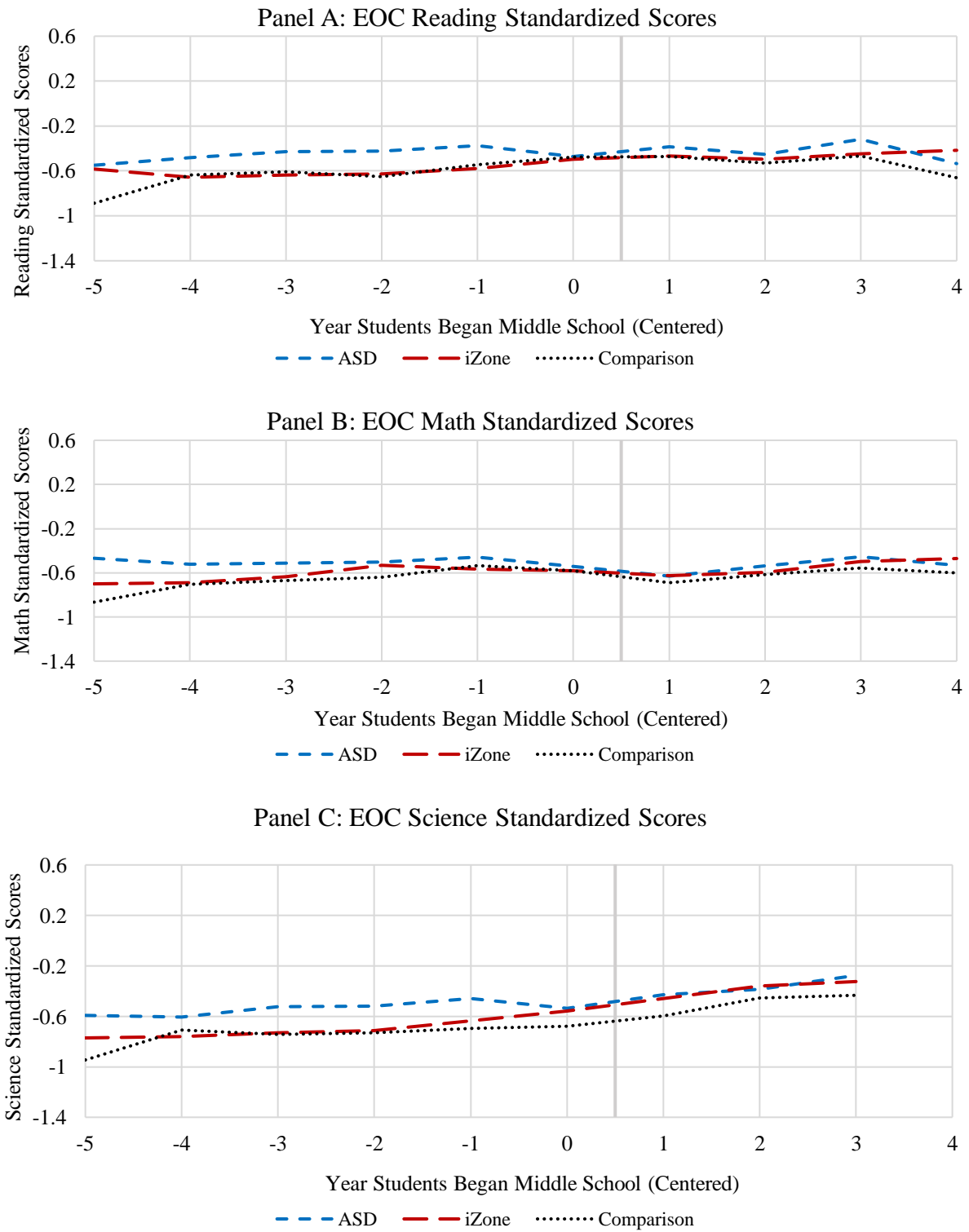
Note. + p<0.1, * p<0.05, ** p<0.01, *** p<0.001 Standard errors in parentheses clustered at school level. Results in this table use the DiD estimator proposed by Callaway and Sant’Anna (2020), which relies on comparisons with students who never attend an ASD or iZone school, and avoids potentially problematic comparisons with already treated students. All models control for the same student and school characteristics as in Equation 1. Student characteristics included as controls include the student’s gender, race, eligibility for FRPM, ML status, and SPED status. High school characteristics included as controls include proportion non-White students, proportion FRPM, proportion ML, and proportion SPED.

Figure 1. Conceptual Framework for Defining Long-Term Effects of Whole-School Reform



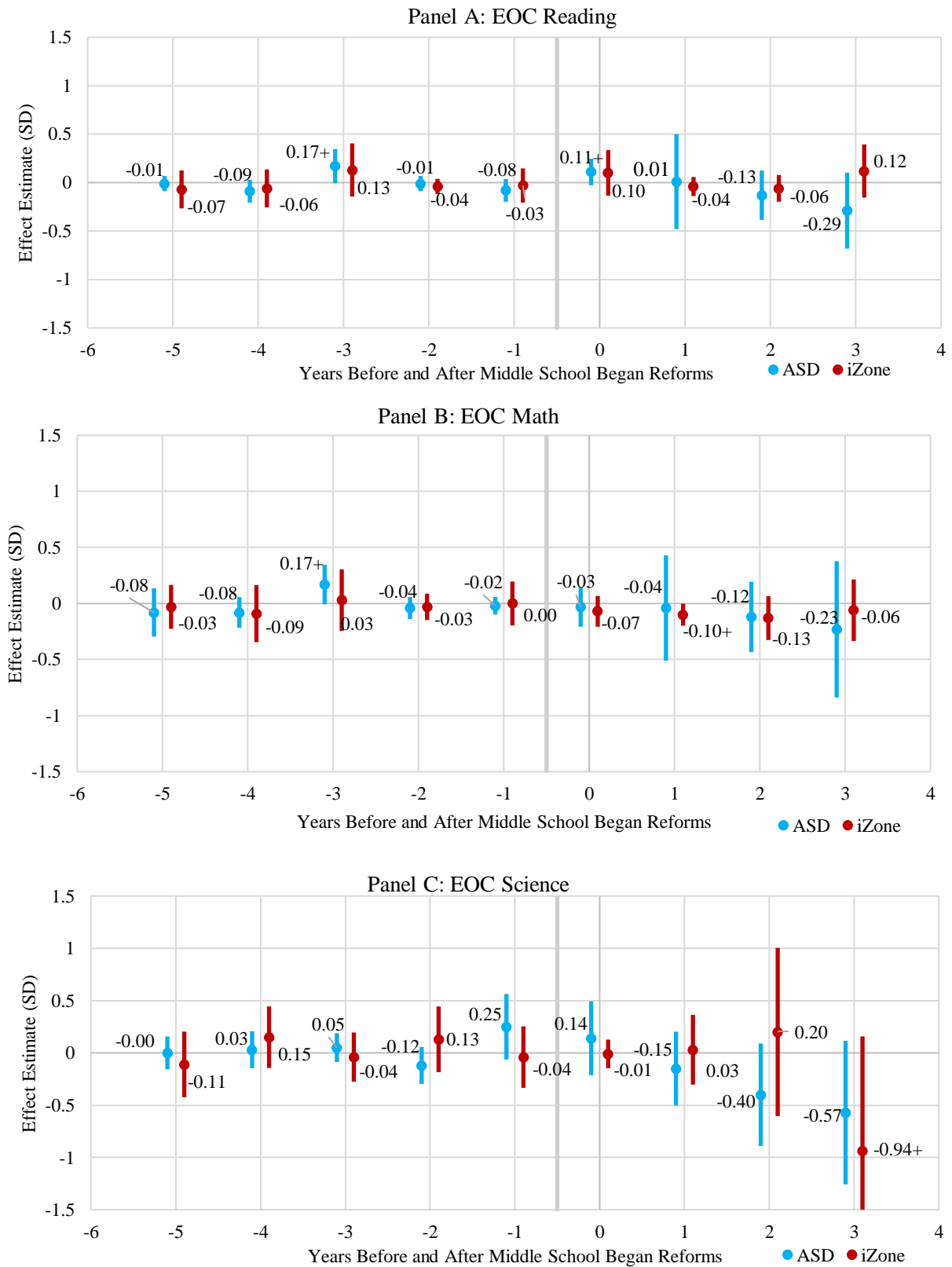
Long Term Outcomes of Turnaround

Figure 2. Trend in Average High School EOC Scores for Students who Began Middle School in Each Year Before and After Turnaround Reforms Began



Long Term Outcomes of Turnaround

Figure 3. DiD Estimate on Test Scores by Years before and After Reforms Began



Appendix A: Auxiliary Analyses and Checks

Appendix Table A1. Baseline Comparison of ASD, iZone, and Comparison Middle Schools in the Year Before Reforms Began

	ASD	iZone	Comparison
Outcomes of Interest			
Reading Score (Standardized)	-0.97 (0.20)	-1.04 (0.17)	-0.96 (0.17)
Math Score (Standardized)	-0.87 (0.20)	-0.93 (0.26)	-1.00 (0.23)
Science Score (Standardized)	-1.05 (0.16)	-1.11 (0.17)	-1.07 (0.22)
Attendance Rate	0.94 (0.01)	0.93 (0.03)	0.95 (0.02)
Chronic Absenteeism Rate	0.18 (0.04)	0.20 (0.11)	0.15 (0.08)
Proportion Zero Tolerance Action	.003 (0.00)	.002 (0.00)	.003 (0.00)
Proportion Expelled	0.05 (0.03)	0.04 (0.03)	0.05 (0.04)
Student Demographic Characteristics			
Proportion Female	0.48 (0.02)	0.47 (0.04)	0.47 (0.03)
Proportion FRPM	0.87 (0.09)	0.92 (0.04)	0.86 (0.13)
Proportion ELL	0.03 (0.02)	0.02 (0.03)	0.02 (0.03)
Proportion SPED	0.19 (0.03)	0.20 (0.02)	0.18 (0.04)
Proportion Asian	0.00 (0.01)	0.01 (0.01)	0.00 (0.01)
Proportion Black	0.92 (0.05)	0.89 (0.13)	0.95 (0.04)
Proportion Hispanic	0.05 (0.04)	0.04 (0.06)	0.03 (0.05)
Proportion White	0.02 (0.04)	0.06* (0.07)	0.01 (0.01)
Proportion Other Race	0.001 (0.00)	0.001 (0.00)	0.001 (0.00)

Note. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$ Standard deviations in parentheses. t -tests compare ASD and iZone schools with comparison schools. The baseline year is the year before schools begin turnaround reforms.

Long Term Outcomes of Turnaround

Appendix Figure A2. Illustration of the Data Structure

Student	Lagged Academic Year (2016 for 2015-16)	Grade	Standardized EOC Test Score	ACT Test Score	Assigned to ASD Middle School (ASD_i)	Assigned to an iZone Middle School ($iZone_i$)	Began Middle School After Reforms Began ($Began\ After_i$)
A	2016	9	-0.23	.	1	0	1
A	2017	10	0.34	.	1	0	1
A	2018	11	-0.12	.	1	0	1
A	2019	12	.	29	1	0	1
B	2014	9	0.17	.	1	0	0
B	2015	10	0.21	.	1	0	0
B	2016	11	0.28	.	1	0	0
B	2017	12	0.39	34	1	0	0
C	2016	9	-0.17	.	0	1	1
C	2017	10	.	.	0	1	1
C	2018	11	-0.21	.	0	1	1
C	2019	12	0.07	19	0	1	1
D	2014	9	0.51	.	0	1	0
D	2015	10	-0.23	.	0	1	0
D	2016	11	.	25	0	1	0
D	2017	12	-0.08	.	0	1	0
E	2014	9	0.93	.	0	0	0
E	2015	10	-0.74	.	0	0	0
E	2016	11	.	28	0	0	0
E	2017	12	0.11	.	0	0	0

Long Term Outcomes of Turnaround

Appendix Table A3. Effect of Assignment to a Turnaround Middle School on the Actual Attendance in a Turnaround Middle School

	(1)	(2)	(3)	(4)
Outcome:	Attend Ever ASD	Attended ASD Post Reforms	Attend Ever iZone	Attended iZone Post Reforms
Treatment Indicators				
Assigned to ASD Middle School	0.29*** (0.03)	0.02** (0.01)	-0.01 (0.01)	-0.01* (0.01)
Assigned to iZone Middle School	-0.07*** (0.02)	0.01** (0.00)	0.27*** (0.02)	0.04*** (0.01)
Began Middle School After Turnaround Began	-0.00 (0.01)	0.04** (0.01)	0.07*** (0.01)	0.02 (0.02)
Assigned to ASD Middle School*Began After	-0.04 (0.06)	0.37*** (0.04)	-0.02 (0.02)	-0.01 (0.02)
Assigned to iZone Middle School*Began After	0.02 (0.01)	-0.03* (0.01)	-0.02 (0.02)	0.27*** (0.02)
R Squared	0.11	0.16	0.10	0.19
N	131059	131059	131059	131059

Note. + p<0.1, * p<0.05, ** p<0.01, *** p<0.001 All models include high school and year fixed effects. Standard errors in parentheses clustered at school level. All estimates are from an ITT sample of students assigned to ASD, iZone, or comparison middle schools based on their elementary school. Student characteristics included as controls include the student's gender, race, eligibility for FRPM, ML status, and SPED status. High school characteristics included as controls include proportion non-White students, proportion FRPM, proportion ML, and proportion SPED.

Long Term Outcomes of Turnaround

Appendix Table A4. DiD Effect using Callaway and Sant'Anna (2020) Estimator with Lagged Outcomes Measured in the Year Prior to Middle School

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	EOC Reading	EOC Math	EOC Science	Attendance Rate	Chronic Absenteeism (Binary)	Zero Tolerance (Binary)	Expulsion (Binary)
ASD Effect	-0.10*	-0.14	-0.37***	0.01	-0.06	-0.02	-0.00
	(0.05)	(0.10)	(0.11)	(0.07)	(0.29)	(0.01)	(0.01)
iZone Effect	0.03	-0.18**	0.07	0.01	-0.03	0.01	0.00
	(0.06)	(0.07)	(0.28)	(0.01)	(0.04)	(0.01)	(0.01)

Note. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ Standard errors in parentheses clustered at school level. Results in this table use the DiD estimator proposed by Callaway and Sant'Anna (2020). All models control for the same student and school characteristics as in Equation 1 along with a lagged outcome control measured in the year before middle school. Note that lagged outcomes in elementary school are not available for ACT, dropout, or graduation outcomes. Student characteristics included as controls include the student's gender, race, eligibility for FRPM, ML status, and SPED status. High school characteristics included as controls include proportion non-White students, proportion FRPM, proportion ML, and proportion SPED.

Long Term Outcomes of Turnaround

Appendix Table A5. Effect of ASD and iZone Reforms on Test Scores while Students are Still in Middle School – Sample Restricted to only ASD, iZone, or Comparison Middle Schools

	(1) All Subjects	(2) Reading	(3) Math	(4) Science
ASD	0.02 (0.04)	-0.00 (0.04)	0.03 (0.06)	0.03 (0.05)
iZone	0.09* (0.04)	0.05 (0.04)	0.12* (0.05)	0.11 (0.06)
R-squared	0.46	0.54	0.40	0.45
Observations	159623	53360	53229	53034

Note. + $p < 0.1$, * $p < .05$. ** $p < .01$. *** $p < .001$. Standard errors in parentheses. Results presented in this table follow similar methods as Pham et al. (2020). *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.

Long Term Outcomes of Turnaround

Appendix Table A6. DiD Effect Using Callaway and Sant’Anna (2020) Estimator with No Covariates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	EOC Reading	EOC Math	EOC Science	ACT Composite Score	Attendance Rate	Chronic Absenteeism (Binary)	Zero Tolerance (Binary)	Expulsion (Binary)	Dropout (Binary)	Graduate (Binary)
ASD Effect	0.05 (0.06)	-0.03 (0.07)	-0.28* (0.13)	0.05 (0.29)	-0.01+ (0.01)	0.03 (0.02)	-0.01 (0.01)	0.00 (0.01)	0.00 (0.01)	-0.07 (0.06)
iZone Effect	0.03 (0.04)	-0.05 (0.05)	-0.08 (0.16)	0.48 (0.34)	0.01 (0.00)	-0.01 (0.01)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	0.03 (0.02)

Note. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ Standard errors in parentheses clustered at school level. Results in this table use the DiD estimator proposed by Callaway and Sant’Anna (2020), which relies on comparisons with students who never attend an ASD or iZone school, and avoids potentially problematic comparisons with already treated students. Models do not include any student or school covariates.

Long Term Outcomes of Turnaround

Appendix Table A7. Average Student Incoming and Exit Rates by ASD, iZone, and Comparison Middle Schools

	2013	2014	2015	2016	2017	2018	2019	2020
Incoming								
ASD	0.33	0.33	0.37	0.31	0.34	0.33	0.31	0.32
iZone	0.31	0.33	0.38	0.34	0.31	0.33	0.32	0.28
Comparison	0.33	0.33	0.32	0.3	0.31	0.28	0.29	0.27
Outgoing								
ASD	0.33	0.38	0.32	0.32	0.32	0.29	0.25	.
iZone	0.35	0.32	0.32	0.36	0.32	0.34	0.31	.
Comparison	0.37	0.39	0.38	0.34	0.31	0.28	0.3	.

Note. Column labels represent lagged academic years (e.g., 2013 for 2012-13). Exit rates are not available in 2020 because we need the 2021 school enrollment data to calculate exit rates after 2020.

Long Term Outcomes of Turnaround

Appendix Table A8. Characteristics of Incoming and Outgoing Students in ASD, iZone, and Comparison Schools

	<u>ASD</u>		<u>iZone</u>		<u>Comparison</u>	
	Incoming	Outgoing	Incoming	Outgoing	Incoming	Outgoing
Outcomes of Interest						
Reading Score (Standardized)	-0.92 (0.95)	-0.90 (0.96)	-0.90 (0.96)	-0.92 (0.96)	-0.91 (0.93)	-0.91 (0.94)
Math Score (Standardized)	-0.90 (0.94)	-0.93 (0.92)	-0.81 (0.95)	-0.87 (0.98)	-0.84 (0.96)	-0.89 (0.95)
Science Score (Standardized)	-0.89 (0.93)	-0.92 (0.95)	-0.86 (0.97)	-0.92 (1.00)	-0.90 (0.89)	-0.94 (0.90)
Attendance Rate	0.91 (0.11)	0.89 (0.13)	0.90 (0.12)	0.90 (0.12)	0.92 (0.11)	0.91 (0.11)
Chronic Absenteeism Rate	0.30 (0.46)	0.39 (0.49)	0.32 (0.47)	0.36 (0.48)	0.28 (0.45)	0.29 (0.45)
Proportion Zero Tolerance Action	0.01 (0.10)	0.01 (0.11)	0.01 (0.09)	0.01 (0.11)	0.01 (0.07)	0.01 (0.09)
Proportion Expelled	0.00 (0.04)	0.00 (0.05)	0.02 (0.14)	0.03 (0.16)	0.03 (0.18)	0.04 (0.19)
Student Demographic Characteristics						
Proportion Female	0.47 (0.50)	0.47 (0.50)	0.47 (0.50)	0.46 (0.50)	0.47 (0.50)	0.46 (0.50)
Proportion FRPL	0.79 (0.41)	0.84 (0.37)	0.87 (0.34)	0.89 (0.31)	0.85 (0.36)	0.87 (0.34)
Proportion ELL	0.06 (0.24)	0.04 (0.19)	0.07 (0.26)	0.04 (0.20)	0.04 (0.20)	0.03 (0.16)
Proportion SPED	0.17 (0.37)	0.18 (0.39)	0.19 (0.39)	0.20 (0.40)	0.18 (0.39)	0.19 (0.39)
Proportion Asian	0.00 (0.06)	0.01 (0.08)	0.00 (0.06)	0.01 (0.07)	0.00 (0.04)	0.00 (0.04)
Proportion Black	0.87 (0.34)	0.87 (0.33)	0.82 (0.39)	0.82 (0.38)	0.92 (0.26)	0.93 (0.25)
Proportion Hispanic	0.09 (0.28)	0.07 (0.26)	0.10 (0.30)	0.08 (0.27)	0.06 (0.25)	0.06 (0.23)
Proportion White	0.04 (0.19)	0.05 (0.21)	0.08 (0.27)	0.09 (0.29)	0.01 (0.09)	0.01 (0.10)
Proportion Other Race	0.00 (0.04)	0.00 (0.05)	0.00 (0.04)	0.00 (0.05)	0.00 (0.03)	0.00 (0.03)

Note. Standard deviations in parentheses.

Long Term Outcomes of Turnaround

Appendix Table A9. DiD Effect Using Callaway and Sant'Anna (2020) Estimator with Outcomes from the 2019-20 School Year

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	EOC Reading	EOC Math	EOC Science	ACT Composite Score	Attendance Rate	Chronic Absenteeism (Binary)	Zero Tolerance (Binary)	Expulsion (Binary)	Dropout (Binary)	Graduate (Binary)
ASD Effect	0.03 (0.08)	-0.03 (0.12)	-0.08 (0.31)	0.13 (0.34)	0.00 (0.02)	-0.02 (0.08)	-0.01 (0.01)	0.00 (0.01)	0.01 (0.02)	-0.03 (0.06)
iZone Effect	-0.02 (0.07)	-0.15+ (0.08)	-0.24* (0.11)	-0.01 (0.25)	0.01 (0.01)	-0.01 (0.03)	0.01+ (0.00)	0.00 (0.00)	-0.01 (0.01)	0.03 (0.03)

Note. + p<0.1, * p<0.05, ** p<0.01, *** p<0.001 Standard errors in parentheses clustered at school level. Results in this table use the DiD estimator proposed by Callaway and Sant'Anna (2020), which relies on comparisons with students who never attend an ASD or iZone school, and avoids potentially problematic comparisons with already treated students. Student characteristics included as controls include the student's gender, race, eligibility for FRPM, ML status, and SPED status. High school characteristics included as controls include proportion non-White students, proportion FRPM, proportion ML, and proportion SPED.

Long Term Outcomes of Turnaround

Appendix Table A10. DiD Effect Using Callaway and Sant’Anna (2020) Estimator – Outcome is a Dichotomous Indicator for whether the Student is Ever Observed in a Tennessee Public High School

	(1)
	Attended Public School in TN
ASD Effect	-0.06 (0.04)
iZone Effect	-0.07 (0.04)

Note. + p<0.1, * p<0.05, ** p<0.01, *** p<0.001 Standard errors in parentheses clustered at school level. Results in this table use the DiD estimator proposed by Callaway and Sant’Anna (2020), which relies on comparisons with students who never attend an ASD or iZone school, and avoids potentially problematic comparisons with already treated students. Student characteristics included as controls include the student’s gender, race, eligibility for FRPM, ML status, and SPED status. High school characteristics included as controls include proportion non-White students, proportion FRPM, proportion ML, and proportion SPED.

Appendix B: Instrumented Difference-in-Differences Approach

Although the DiD approach addresses school-level selection bias, estimating Equation 1 using OLS recovers an ITT effect. To recover a TOT effect that addresses selection bias at the student-level, we use a generalized method of moments approach (Baum et al., 2007) that instruments enrollment in all grades offered at an ASD or iZone middle school using the indicator for whether the student was assigned to an ASD or iZone middle school in their final year in elementary school, per the district's school feeder pattern. We call this approach an instrumented difference-in-differences (DDIV). This DDIV approach uses the difference-in-differences (DD) model to address school level selection (e.g., schools selected for turnaround based on unobserved characteristics that also affect student outcomes) and uses an instrumental variable approach to address student level selection (e.g., students systematically leaving turnaround schools for reasons related to their long-term outcomes in high school).

Specifically, we use school feeder patterns to construct indicators for whether students are assigned to an ASD or iZone middle school based on the elementary school they attended one year before middle school. We then predict whether students attended an ASD or iZone middle school using the indicators for whether students are *assigned* to an ASD or iZone middle school. Note that in this analysis, we use indicators for attendance in all grades at the turnaround middle school to avoid bias from students who move and experience the effects of a different middle school. Since we have two treatment indicators in Equation 1 of the main text (ASD_i or $iZone_i$) that are each interacted with an indicator for the years after reforms ($BeganAfter_i$), we include two instruments (indicators for assignment to ASD or iZone middle schools) and also interact each instrument with the *BeganAfter* indicator. Thus, our model is just identified with two endogenous predictors and two exogeneous instruments. We include both instruments, their

Long Term Outcomes of Turnaround

interactions with the *BeganAfter_i* indicator, and all covariates used as controls in our first stage equations (Wooldridge, 2010). The DDIV model includes all covariates in Equation 1.

This DDIV approach scales the ITT effect on student outcomes by how well assignment to an ASD or iZone middle school predicts actually attending an ASD or iZone middle school. Under the instrument relevance and exogeneity assumptions and the exclusion restriction, this approach produces an intent-to-treat estimate that addresses our primary student selection concerns. To support the relevance assumption, we show first stage F-statistics that indicate being assigned to attend an ASD or iZone middle school strongly predicts the probability that students are observed attending an ASD or iZone middle school. We argue that this instrument satisfies the exclusion restriction because assignment to a turnaround middle school is unlikely to affect student outcomes unless the student actually attends the assigned school. Also, the exogeneity restriction is likely satisfied because any omitted factors associated with being zoned to a turnaround middle school likely occurs because these students live in disadvantaged neighborhoods, but students zoned to comparison priority schools live in similar neighborhoods, often very close to each other. Finally, it is important to note that this approach identifies a local average treatment effect, and the causal estimate relates to compliers only; that is, students whose attendance at an ASD or iZone middle school was determined by residential assignment. All standard errors are clustered at the school level.

Appendix Table B1 below shows results from this DDIV approach. The table shows that our results are generally in the same direction as the results reported in the main text. However, the point coefficients are generally scaled up because assignment to a turnaround middle school does not perfectly predict actual attendance in these schools. The TOT effect is thus larger in

Long Term Outcomes of Turnaround

magnitude and statistically significant when we examine EOC test scores for students in ASD schools.

Long Term Outcomes of Turnaround

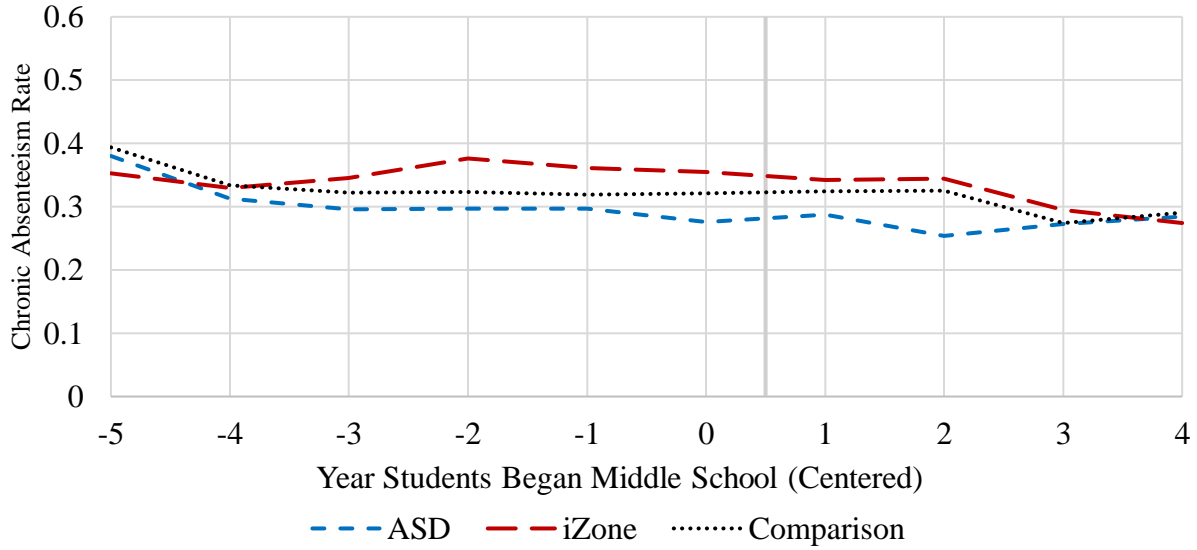
Appendix Table B1. Effect of Attending ASD and iZone Middle Schools on Outcomes in High School from DDIV Model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	EOC Reading	EOC Math	EOC Science	ACT Composite Score	Attendance Rate	Chronic Absenteeism (Binary)	Zero Tolerance (Binary)	Expulsion (Binary)	Dropout (Binary)	Graduate (Binary)
Treatment Indicators										
ASD Middle School	0.17*** (0.05)	0.27** (0.09)	0.23** (0.08)	-1.30+ (0.77)	0.01 (0.01)	-0.03 (0.03)	-0.01 (0.01)	0.01 (0.01)	-0.03*** (0.01)	-0.02 (0.03)
iZone Middle School	0.07 (0.07)	0.15 (0.11)	0.15 (0.13)	-2.03+ (1.08)	-0.01 (0.01)	-0.00 (0.04)	-0.01 (0.01)	0.00 (0.01)	-0.03* (0.01)	-0.04 (0.05)
After Turnaround	0.16** (0.06)	0.17* (0.08)	0.19* (0.08)	-0.84 (0.56)	0.03*** (0.01)	-0.13*** (0.03)	0.00 (0.01)	0.02+ (0.01)	-0.03*** (0.01)	-0.14* (0.06)
ASD Middle School*Began After Turnaround	-0.36* (0.16)	-0.48*** (0.14)	-0.33* (0.16)	1.16 (0.90)	0.00 (0.01)	0.05 (0.07)	-0.03+ (0.02)	-0.04* (0.02)	0.03 (0.02)	-0.00 (0.10)
iZone Middle School*Began After Turnaround	-0.14 (0.09)	-0.22+ (0.13)	0.07 (0.13)	1.25 (0.80)	-0.01 (0.01)	0.06 (0.06)	0.00 (0.01)	-0.02 (0.02)	0.01+ (0.01)	0.16 (0.14)
Student Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First Stage F	16.43	13.80	17.04	7.84	12.51	12.52	9.06	9.02	12.58	2.20
R Squared	0.31	0.17	0.17	0.10	0.08	0.07	0.01	0.01	0.01	0.10
N	26036	20361	11834	4103	53350	53376	26804	26804	62640	12338

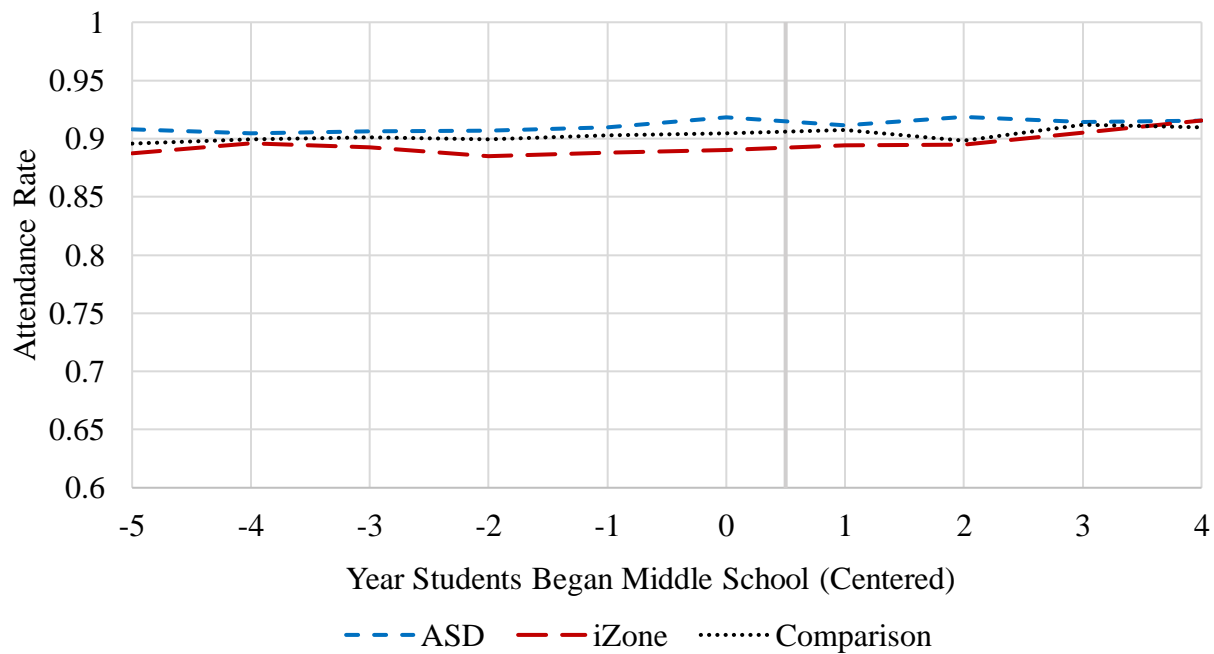
Note. + p<0.1, * p<0.05, ** p<0.01, *** p<0.001 All models include high school and year fixed effects. Standard errors in parentheses clustered at school level. Lagged outcome variables in the year prior to middle school are not available for ACT scores, drop out, or graduation. Student characteristics included as controls include the student's gender, race, eligibility for FRPM, ML status, and SPED status. High school characteristics included as controls include proportion non-White students, proportion FRPM, proportion ML, and proportion SPED.

Appendix C: Trends by Year

Appendix Figure C1. Trend in Average High School Chronic Absenteeism Rates for Students who Began Middle School in Each Year Before and After Turnaround Reforms Began

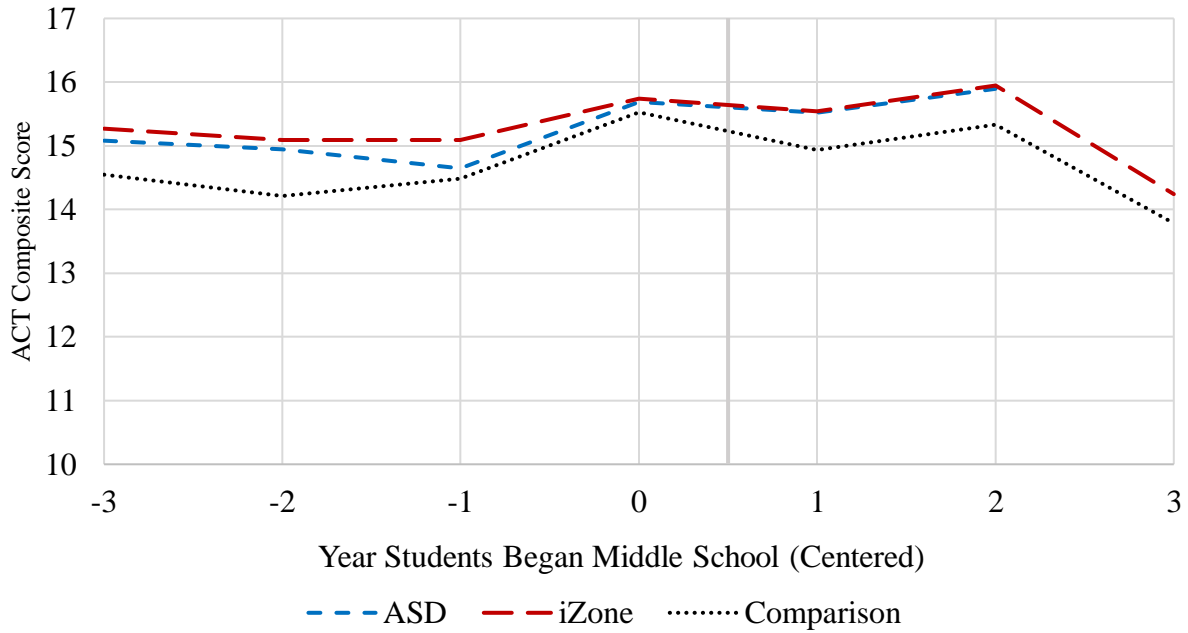


Appendix Figure C2. Trend in Average Attendance Rate for Students who Began Middle School in Each Year Before and After Turnaround Reforms Began

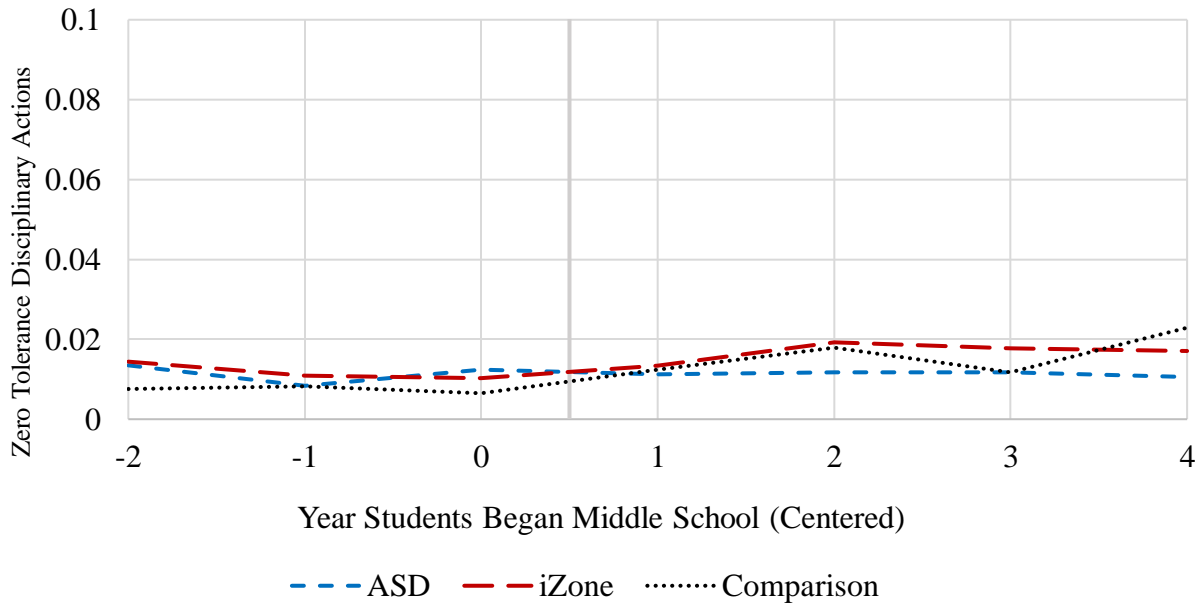


Long Term Outcomes of Turnaround

Appendix Figure C3. Trend in Average ACT Composite Score for Students who Began Middle School in Each Year Before and After Turnaround Reforms Began

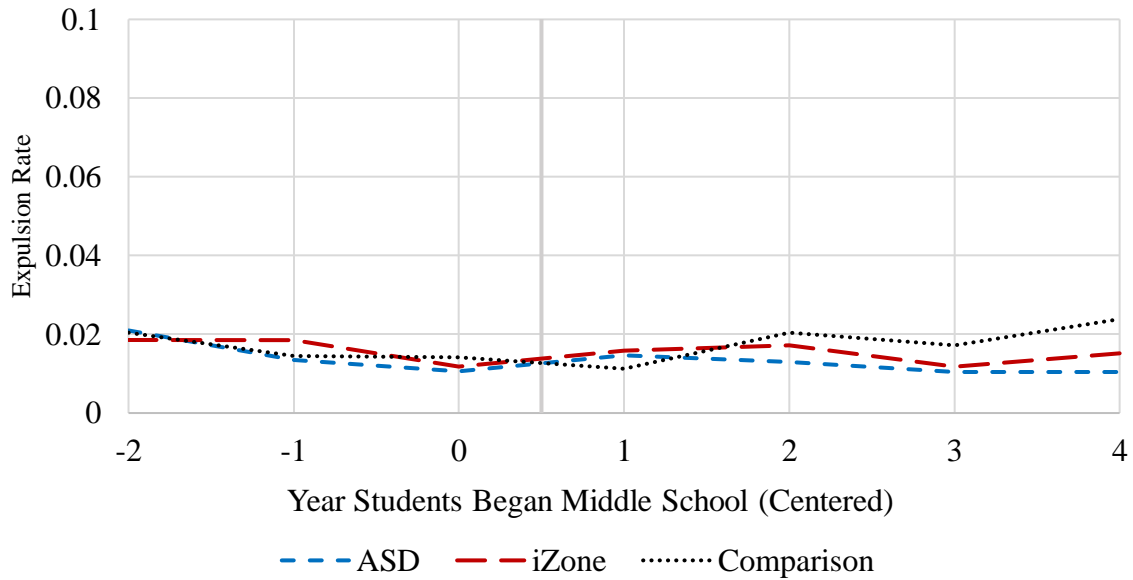


Appendix Figure C4. Trend in Average Zero Tolerance Disciplinary Actions for Students who Began Middle School in Each Year Before and After Turnaround Reforms Began

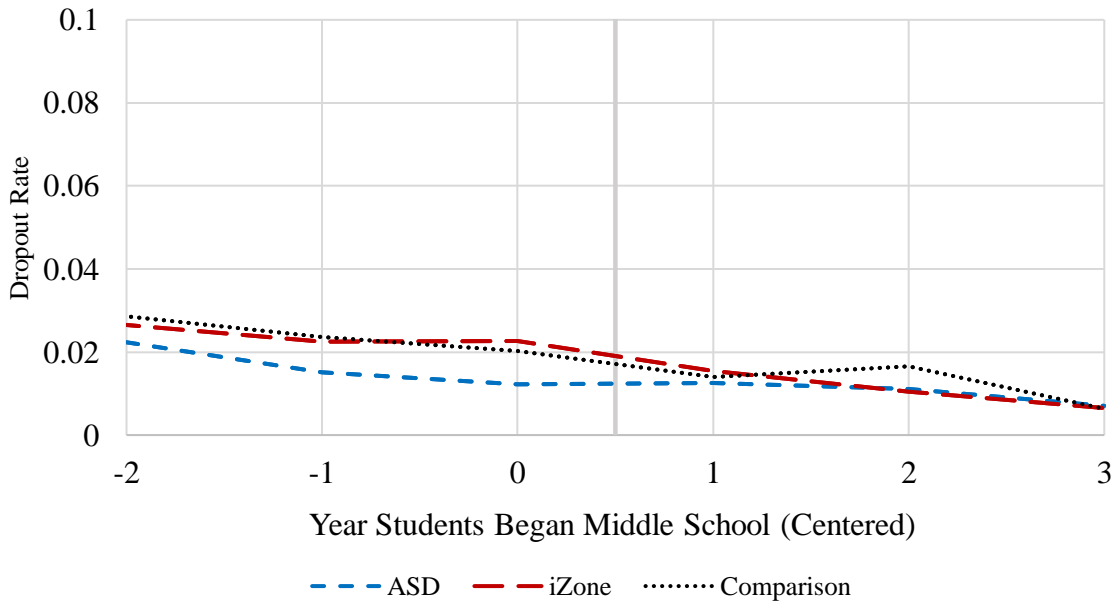


Long Term Outcomes of Turnaround

Appendix Figure C5. Trend in Average Expulsion Rate for Students who Began Middle School in Each Year Before and After Turnaround Reforms Began

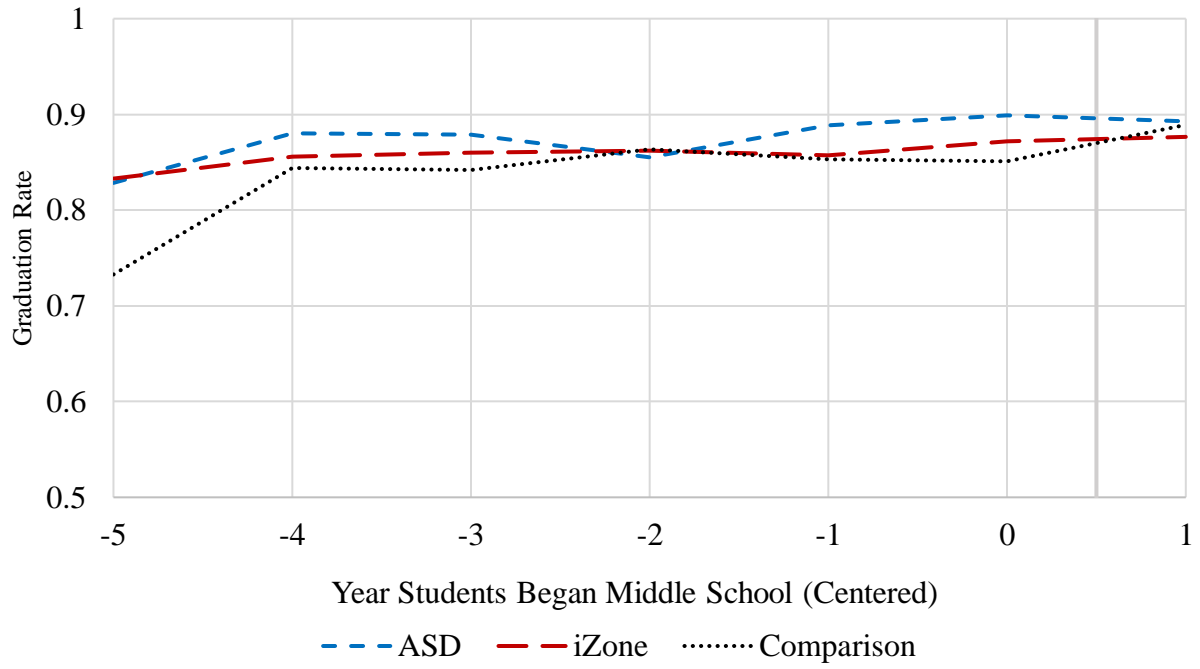


Appendix Figure C6. Trend in Average Dropout Rate for Students who Began Middle School in Each Year Before and After Turnaround Reforms Began



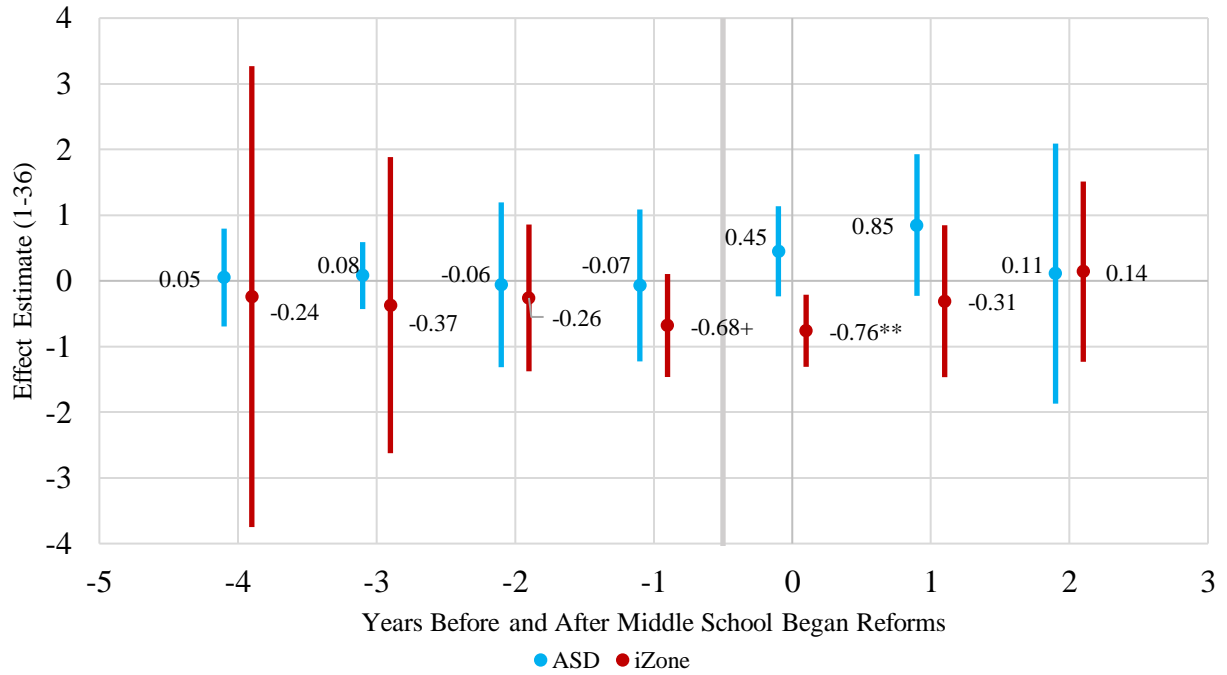
Long Term Outcomes of Turnaround

Appendix Figure C7. Trend in Average Graduation Rate for Students who Began Middle School in Each Year Before and After Turnaround Reforms Began



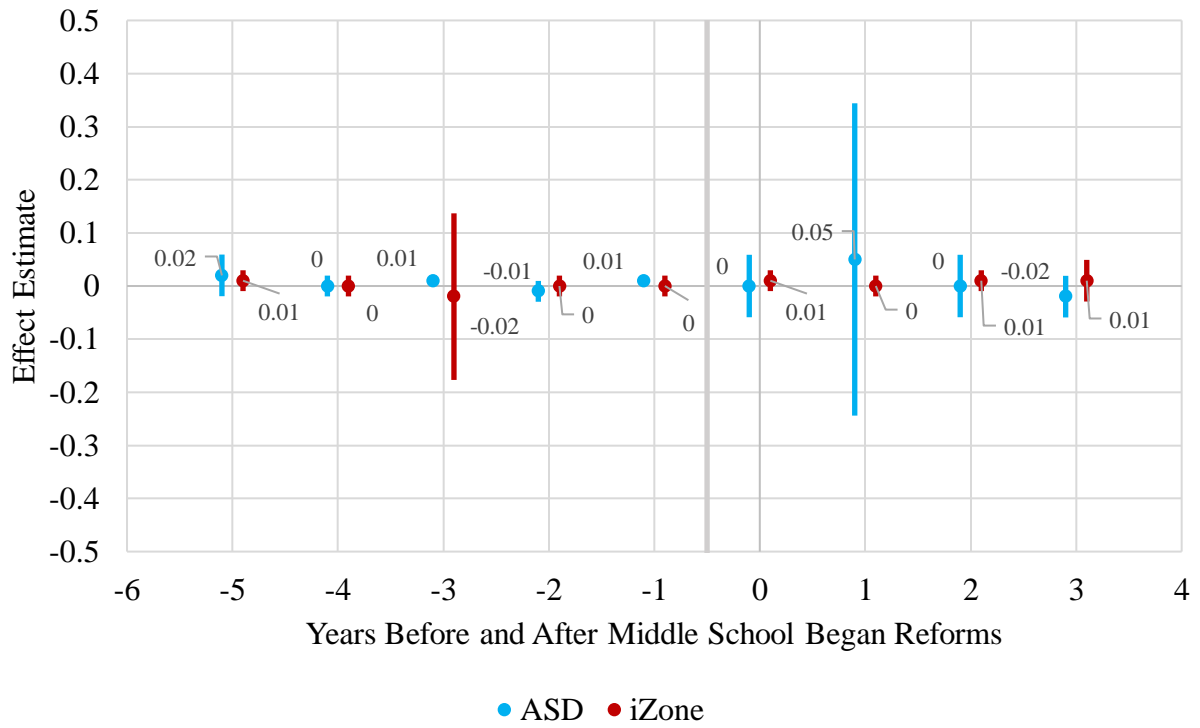
Appendix D

Appendix Figure D1. DiD Estimate by Years before and After Reforms Began – Outcome is ACT Composite Score



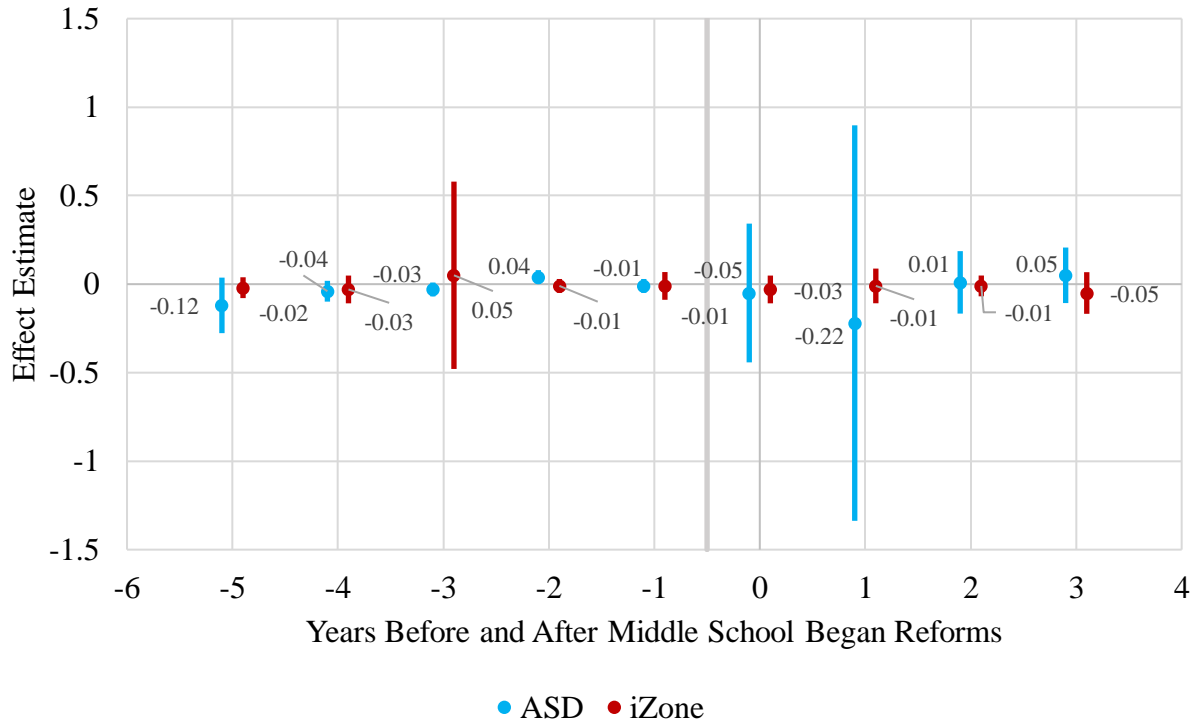
Long Term Outcomes of Turnaround

Appendix Figure D2. DiD Estimate by Years before and After Reforms Began – Outcome is Attendance Rate



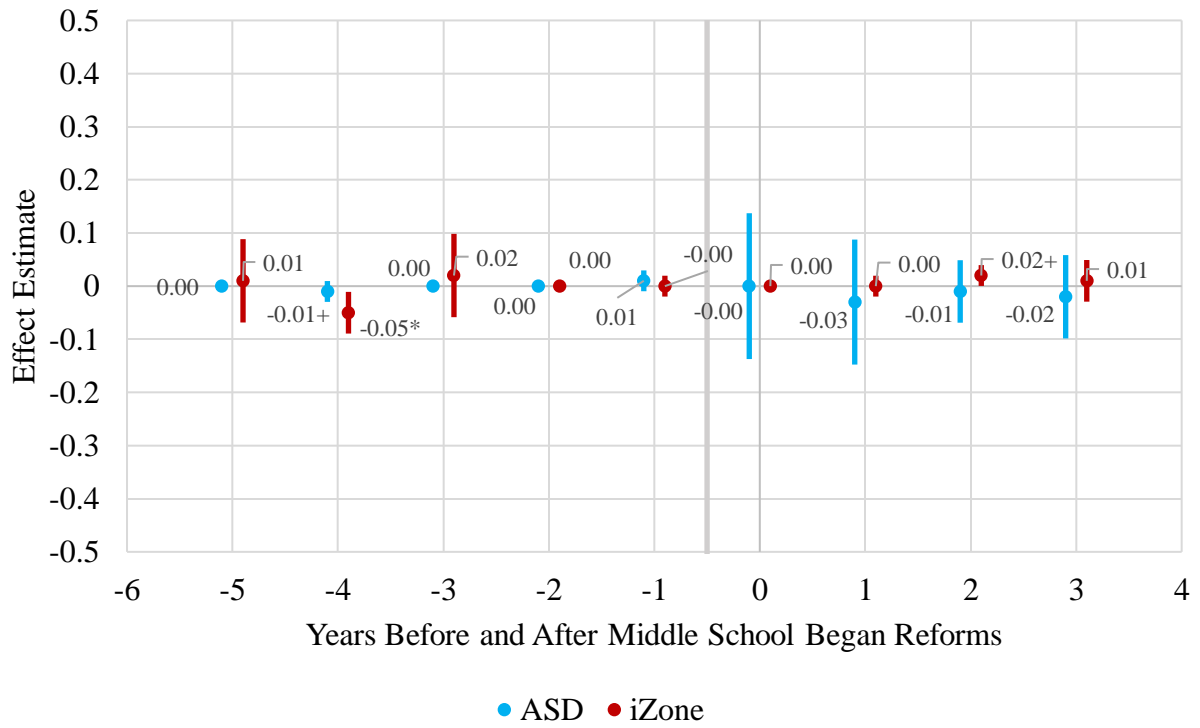
Long Term Outcomes of Turnaround

Appendix Figure D3. DiD Estimate by Years before and After Reforms Began – Outcome is a Dichotomous Indicator for Chronic Absenteeism



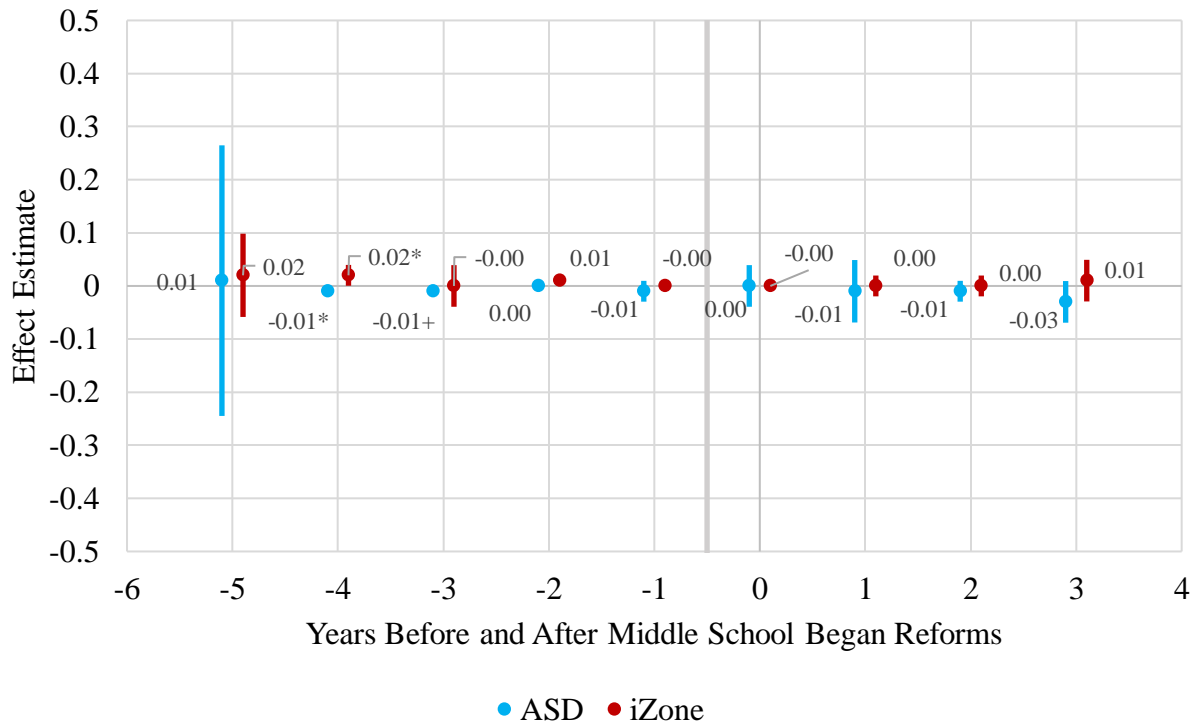
Long Term Outcomes of Turnaround

Appendix Figure D4. DiD Estimate by Years before and After Reforms Began – Outcome is a Dichotomous Indicator for Receiving a Zero Tolerance Disciplinary Action



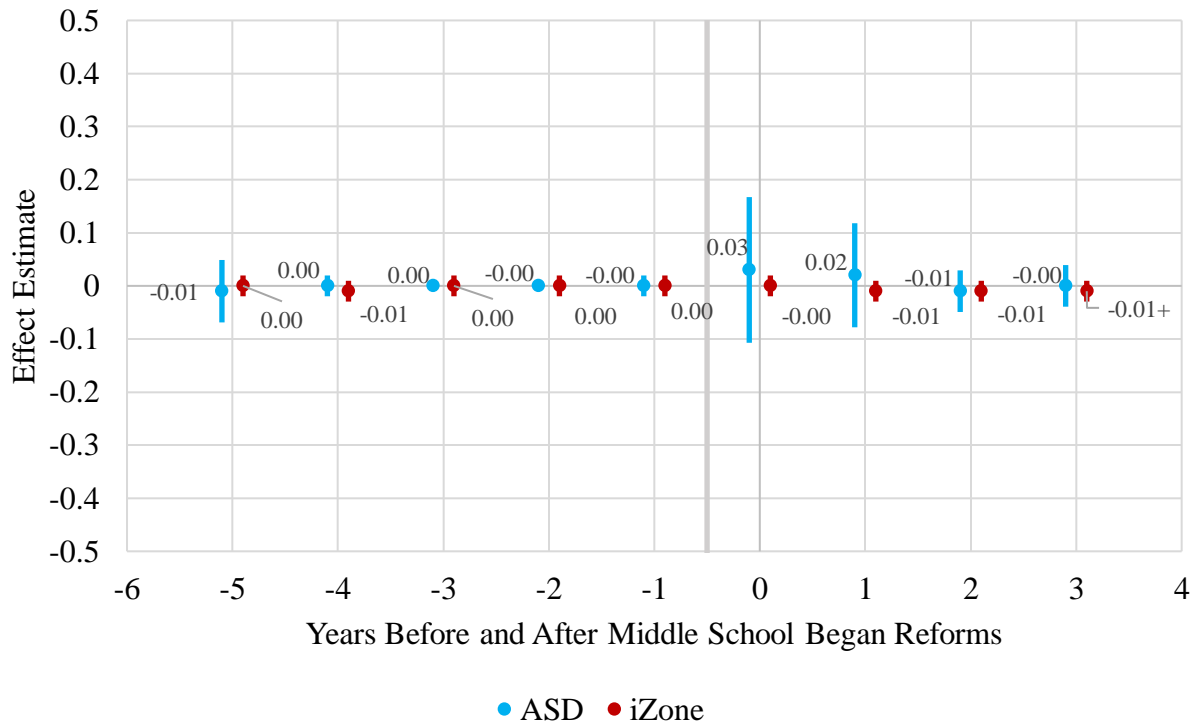
Long Term Outcomes of Turnaround

Appendix Figure D5. DiD Estimate by Years before and After Reforms Began – Outcome is a Dichotomous Indicator for Expulsion



Long Term Outcomes of Turnaround

Appendix Figure D6. DiD Estimate by Years before and After Reforms Began – Outcome is a Dichotomous Indicator for Dropping Out



Long Term Outcomes of Turnaround

Appendix Figure D7. DiD Estimate by Years before and After Reforms Began – Outcome is a Dichotomous Indicator for Graduating

