The Role of Governance and Management in School Turnaround Policies: The Case of Tennessee's Achievement School District and iZones

Ron Zimmer, Gary T. Henry & Adam Kho

Vanderbilt University

May 9, 2016

Abstract:

In recent years, the federal government has invested billions of dollars through Race to the Top and School Improvement Grants to address chronically low performing schools. These grants required prescriptive turnaround reforms including the option of a "restart" approach, which required a replacement of the management of schools, often with outside providers such as charter management organizations. In some cases, districts spearhead the reform, but in others, the state may intercede, taking over schools and changing governance from the local school district to the state. This latter restart approach assumes that districts do not have the capacity or the will or both to manage meaningful reforms. In this paper, we examine Tennessee's use of the Race to the Top grant to implement reform models that included both a change in governance—i.e., state takeover of schools with management of schools outside of the district—and maintaining governance and management of reforms within the district. Our study examines whether it is necessary to have the governance and management of schools outside of the district in order to have a meaningful change in performance. We find that schools managed by districts are more successful at improving chronically low performing schools than reforms requiring governance and management outside of district auspices.

ACKNOWLEDGMENTS: We wish to thank Achievement School District and Tennessee Department of Education for their feedback and assistance. We would also like to thank the Lam Pham and Samantha Viano for research assistance. Funding for this research was provided by the state of Tennessee's Race to the Top grant from the United States Department of Education and the Walton Family Foundation. Any opinions or errors are solely attributable to us, however.

I. Introduction

A steadfast but elusive goal of policymakers is to improve the performance of chronically low performing schools. In pursuit of this goal, from 1988 through 2005, the federal government distributed nearly $2 billion in grants for Comprehensive School Reform (CSR), which required schools to choose and adopt a whole school reform model that prescribed the teaching and learning process to be followed throughout school. While the rigorous evaluation of some of the demonstrations of whole school reform models have shown positive effects on student achievement, e.g. *Success For All*, the highest quality studies of the overall effects of implementing CSR at scale have found few positive and some negative effects (Borman, Slavin, Chueng, Chamberlain, Madden, & Chambers, 2007; Bifulco, Duncombe & Yinger, 2005; Gross, Booker, & Goldhaber, 2009).

While CSR programs proved inconsistent and mainly ineffective in addressing the challenges of chronically low performing schools, other initiatives such as state accountability policies as well as *No Child Left Behind* (2001) heightened the awareness of pervasiveness of low performing schools through greater testing and reporting of school results. In response, the federal government encouraged states and districts to adopt federally approved reforms for low performing schools that can be broadly characterized as "turnaround" policies by providing more than $7 billion dollars in resources through Race to the Top (RttT) and School Improvement Grants (SIGs). In addition, the federal government put into place other federal policies such as NCLB Waivers to increase state accountability pressures on persistently low performing schools (U.S. Department of Education, 2010).

In the last two decades, the federal initiatives transitioned from voluntary CSR policies to mandatory turnaround policies. This transition is a reflection of a belief that reforming the teaching

and learning processes, often referred to as the "technical core of schooling", would not be sufficient to produce swift and dramatic increases in student performance (Herman, Dawson, Dee, Greene, Maynard, Redding, 2008). The "theory of change" for federal school turnaround suggests more fundamental changes such as personnel replacement and removal of chronically low performing schools from the districts that seemed to lack the capacity and/or will to improve them—in other words, changes to governance, management, and personnel are necessary for meaningful improvement in these schools. The prescribed turnaround models include replacing principals and all or most of the teachers and fundamental changes to the political authority traditionally vested in local school districts, which may lead to state takeover or restarting the schools under the auspices of a charter management organization (CMO).

In this study, we are able to examine the effects of turnaround that changed the governance of schools from local to state and management of schools from public to private. In addition to these types of reform, we examine a group of schools that implemented turnaround models that brought them under the management of a special district—a district-within-a district—that replaced personnel, altered school operations and reformed teaching and learning practices. This current study focuses on turnaround policies in Tennessee funded by the $500 million Race to the Top (RttT) award from the federal government in 2010. With these resources, Tennessee pursued three distinct turnaround strategies, two of which involved a change in governance and management. To implement the first two strategies, Tennessee removed a group of low performing schools from their local districts and placed them in the state's Achievement School District (ASD). These schools were then either directly run by the ASD or matched with a CMO, which was granted autonomy for operating these schools. In the third turnaround approach, three districts established internal local Innovation Zones, labeled iZones.

In this paper we are able to directly examine the effects of all three approaches, which provides insights into the need to alter governance and management of chronically low performing schools to improve their performance. The broad scope of the analysis is in contrast to other contemporaneous studies, which generally focus on only one type of turnaround or changes only in personnel and school operations and structures. It also provides an opportunity to examine whether it is possible to turnaround schools under district management or whether it is necessary to have outside providers. Ultimately, the findings from this study provide salient information at a critical time. With the recent reauthorization of Elementary and Secondary Education Act, states now have more flexibility for how they improve chronically low performing schools. Therefore, it is critical for states to know whether more or less intrusive approaches that affect governance and management of these schools are required for effectively reforming them.

I. Background on Turnaround

The theory of change for federally subsidized turnaround policies, which included the possibility of changes in governance and management as well as personnel replacement, more rigorous teacher evaluation, school autonomy, and changes to school operations, was codified into four models of reform authorized for receipt of federal RttT and SIG funds: (1) "transformation", (2) "turnaround" (with a narrower definition than the term used in the broader context of reform), (3) "restart," and (4) "closure" (Perlman & Redding, eds., 2010). The most frequently implemented model is transformation, which mandates principal replacement, more rigorous teacher evaluation, and increased learning time (US Department of Education, 2011, 2014). The turnaround model goes further by requiring replacement of the principal and at least half of the teaching staff and greater autonomy for the new principal. Restart requires fundamental change to schools by

transferring the school to an independent entity such as a CMO. Restart may be done through the district or the state and the latter case adds a change in governance to this model by placing the school under the direct control of the state. The restart approach assumes that districts, because of entrenched bureaucracy, will not have the capacity or the will or both to turnaround low performing schools (Chubb and Moe, 1990), which raises questions of the role and importance of management and governance in reforming schools.

Often management and governance are used interchangeably in the context of reforms, but they can be thought of as distinct concepts. More specifically, governance can be thought of as the locus of authority for establishing public schools, holding them accountable, and making decisions about who manages them. The actual management of schools can be thought of as a separate concept which involves the responsibility for choosing school leaders and establishing the personnel and operating procedures, such as the length of the school year and school day. By examining the ASD in Tennessee, we can examine the effects of removing schools from district governance and restarting them under the management of the state ASD district or a CMO. In addition, we can examine the effects of maintaining governance of turnaround schools within the district and adopting the transformation model, which affects personnel, operations, and teaching and learning processes.

II. Literature Review

While there have been a number of published studies examining CSR models (Berends, Bodilly, and Kirby, 2002; Desimone 2002, Borman, et al., 2003; Bifulco, Duncombe, and Yinger 2005, Gross, Booker, and Goldhaber, 2009) and school closures (de la Torre, et al., 2012; Engberg et al., 2013, Brummet, 2014; Ruble 2015), research of the use of federal turnaround reforms

including those subsidized by RttT and SIGs are only recently emerging. The earliest work examined the use of state takeover as a means of improving chronically low performing schools—Pennsylvania's takeover of the Philadelphia school district resulting in the turnover of the management of 45 low performing schools to Education Management Organizations (EMO) in the early 2000s. Philadelphia's example could be best characterized as a restart approach focusing on governance as schools were turned over and managed by outside providers. Researchers found that these schools did not outperform the gains compared to other schools within Philadelphia (Gill et al., 2007; MacIver and MacIver, 2006) despite the additional resources these schools were provided.[1] It is worth noting certain restrictions were placed on schools on what reforms they could implement, including a restriction to maintain the schools as neighborhood schools.

More recently, with the incentives associated with RttT and SIGs as well as the NCLB waivers, the number of locations adopting state takeover and turnaround strategies has grown. As the number has grown, so too has the research (although much is unpublished at this point) with various research designs examining different approaches with mixed results (Dee, 2012; Strunk et al., in press; Ruble, 2015; Schuler et al., 2015; Papay, 2015; Doughtery and Weiner, 2015; Heissel and Ladd, 2016; Henry and Guthrie, 2016). A summary of these turnaround studies highlighting the changes in governance (state takeover), system management (local, state or CMO), site management and operations, and teaching and learning along with the effect estimates are presented in Table 1. Of the four studies in which state takeover was involved, two directly run by the state and one by CMOs found positive effects while the fourth, run by the state, did not. Of the three studies with reforms primarily replacing personnel and changing school operations, two

[1] Peterson and Chingos (2007) refined the analysis of these schools by comparing the performance of for-profit EMOs to non-profit EMOs and found that for-profit EMOs outperformed non-profit EMOs.

found positive effects and one found negative effects in reading, and one found positive effects in math while the other two found no effects in math.

Table 1. Summary of Turnaround Studies

Functional Requirements of Public Schools	**Philadelphia (Gill et al., 2007)**	**New Orleans (Ruble, 2015)**	**Los Angeles (Strunk et al., in press)**	**Massachusetts (Papay, 2015)**	**Lawrence, MA (Schueler et al., 2015)**	**Rhode Island (Dougherty & Weiner, 2015)**	**North Carolina (Henry & Guthrie, 2016)**
Governance Authority, Oversight, and Accountability	State established a School Reform Commission (SRC) as a replacement for the local school district board	State establishes Recovery School District (RSD) to takeover and oversee low-performing schools			State appointed Receiver to position of districtwide authority, who selected a new central office team		State created the District and School Transformation Department and identified low performing schools for one of the four turnaround interventions
System Management Leadership & Resource Allocation	SRC selects CMOs to manage schools	RSD selects CMO to manage schools	District maintains management, but replaces some principals	District maintains management, but are required to implement several state initiatives	Receiver and central office team selects outside operators to manage lowest-performing schools, shares responsibilities with operators	Districts maintain management, but some replace principals	Districts maintained management, but most principals were replaced
Site Management & Operations School Personnel & Resources	CMO oversees school operations with some constraints from the SRC.	CMO oversees school operations	District and school leaders maintain authority over operations, but replace some teachers	Districts and schools maintain authority over operations, but are required to implement several state initiatives	Receiver replaces teachers; Receiver and Operators manage schools	Districts and schools maintain authority over operations, but are required to implement several state initiatives	Districts and schools maintain authority over operations, but receive leadership (school & district) coaches and professional development
Teaching & Learning Curriculum & Instruction; Classroom-Level Management	CMO establishes curriculum and sets expectations for teaching and learning	RSD & CMO establish curriculum and sets expectations for teaching and learning		Schools maintain authority over teaching and learning, but are required to implement several state initiatives	Receiver and Operators set expectations for teaching and learning	Schools maintain authority over teaching and learning, but are required to implement several state initiatives	Schools maintain authority over teaching and learning, but receive instructional coaching and targeted professional development
Impact	No overall effects	0.1 to 0.2 SD increase in ELA; 0.1-0.4 SD increase in Math	0.14 SD increase in ELA; no effects in Math	0.14 to 0.38 SD increase in ELA and Math	0.02-0.03 SD increase in ELA; 0.17-0.19 SD increase in Math	Up to 0.35 SD decrease in ELA; no effects in Math	0.02 to 0.03 positive effects on student achievement in all turnaround schools (Henry & Guthrie 2015); no local average treatment effects in middle and elementary schools (Henry &

							Guthrie, 2015; Heissel & Ladd, 2015)

Overall, while these studies have been insightful for whether particular turnaround models can change school performance, they generally have not simultaneously compared various turnaround approaches and have not examined whether reforms under the governance of the district with district management can be effective. In this paper, we examine the effectiveness of governance of schools undergoing turnaround reforms both within and outside of the district. We also examine effectiveness of different management approaches, including a district-managed transformation approach which changed leadership, other personnel, school operations, and instructional methods. We analyze this approach relative to governance outside of the district through the state- or CMO- managed schools, which represents a restart approach. Ultimately, the analysis provides insights into the question of whether it is necessary to remove schools from district governance to effectively turnaround low performing schools.

III. Background on Turnaround in Tennessee

Inspired by the potential of RttT funding, Tennessee passed legislation called First to the Top in January 2010, which created the ASD (Public Chapter No. 2, 2010). With this legislation in hand, the state applied for RttT funding, and, in March of that year, Tennessee was awarded $500 million that called for the State Commissioner of Education to identify the state's lowest-achieving five percent of Title I schools. These schools, known as *priority schools*, would then face a number of possible interventions. Among these possible interventions, none has been bolder and, consequently, more controversial than the ASD – a restart strategy in which a new state school district removes schools from their home districts and either directly manages these schools (ASD-Run) or contracts management responsibilities to a CMO partner (CMO-Run). The stated goal of

the ASD is to move the academic performance of schools taken over from the bottom five percent of schools to the top quartile of schools in Tennessee within five years.

As initially conceived by the original First to the Top legislation, once a school is selected for the ASD, the school would remain in the ASD for at least five years. The school would return to the home district conditional on the performance of both the school and the home district (ESEA Flexibility Request, 2012). While the application did not dismiss the possibility of the state solely operating ASD schools acting as pseudo CMO, the emphasis was on a hybrid model where the state takes over and partners with CMOs to manage the ASD schools. To achieve the goal of raising student achievement to the top quartile of schools in the state, both the ASD and the CMOs are given autonomy to hire talented education professionals with emphasis on teachers (Race to the Top Application for Initial Funding, 2010). It is important to point out that while CMOs are tapped to run schools, the schools remain neighborhood schools rather than schools of choice in which parents and students must opt-in to attend.

In 2012-13, the ASD took over the first cohort of six schools with three schools run by CMOs and three run directly by the ASD. In 2013-14, the ASD added a second cohort of 11 schools, eight run by CMOs and three run directly by ASD. Finally, in the 2014-15 school year, a third cohort of eight schools were added, all CMO-run, while two schools opened in the 2013-14 school year were merged in with other ASD schools. By the 2014-15 school year, 23 schools were operating under the auspices of ASD with five managed directly by the ASD and 18 managed by CMOs.

As an alternative to the state-takeover approach, some priority schools remained under the governance of the district through district innovation zones (iZones). In the 2012-13 school year, thirteen schools opened under iZone followed by eleven schools in 2013-14 and four schools in

2014-15. These schools remained under the governance of their respective districts, but their management changed and these schools experienced new leadership because principals were replaced, which is consistent with the transformation approach. These schools were also given greater autonomy and while in many cases, the schools retained over half of their teachers, they experienced high teacher exit rates compared to other priority schools. On average, 55 percent of teachers exited these schools in their first year of operation, a lower exit rate than schools managed by CMOs, which replaced almost all teachers.

To attract and retain high quality teachers in iZone schools, the district offered substantial raises to teachers who remained in or transferred to these schools (Kebede, 2016). For instance, using statewide teacher level data provided by the state Tennessee Department of Education, including salary data, we found that while the rest of the district teachers in Memphis had a 5% increase in pay in the first year schools operated as iZone schools, teachers who stayed in and transferred into an iZone school received a 14% and 19% increase in pay, respectively. In addition, given the emphasis on attracting or retaining highly effective teachers and, under the ASD model giving them substantial autonomy, we delve into the issue more deeply by describing the relative effectiveness of the teachers who stayed, left, and entered the ASD, iZone, and other Priority schools, Table 2 compares the two measures of teachers' effectiveness, average value-added rating (TVAAS) of teachers and the ratio of teachers with value-added scores above expectations to those with value-added scores below expectations for teachers that stayed, moved or left, or entered ASD schools (both managed directly the ASD and by CMOs), iZone schools, and non-ASD, non-iZone priority schools.[2]

[2] Tennessee's Value-Added Assessment System (TVAAS) is a measure of student growth on state exams. TVAAS score range from one (below expectations) to five (above expectations).

Table 2. Teacher TVAAS Scores for Stayers, Movers & Leavers, and Incoming Teachers in ASD, iZone, and Priority (non-ASD & non-iZone) Schools in Tennessee (averaged over all years of operation within each group): 2012-13, 2013-14 & 2014-15

	Average TVAAS Scores			Ratio of Teachers with Above Expectations Scores (5+4) to Below Expectation Scores(1+2)		
	Stayers	Movers and Leavers	All Incoming	Stayers	Movers and Leavers	All Incoming
ASD	2.77	3.09	3.27	0.60	1.19	1.55
	(87)	(185)	(67)	(64)	(140)	(51)
ASD Cohort 1	2.94	3.27	3.65	0.76	1.46	2.75
	(48)	(92)	(37)	(37)	(69)	(30)
Achievement	3.08	3.25	4.00	0.90	1.31	8.00
	(25)	(52)	(22)	(19)	(37)	(18)
Charter	2.78	3.30	3.13	0.64	1.67	1.00
	(23)	(40)	(15)	(18)	(32)	(12)
ASD Cohort 2	3.44	3.03	2.95	1.60	1.10	1.00
	(18)	(58)	(22)	(13)	(44)	(16)
Achievement	3.00	3.00	3.11	1.00	0.88	1.00
	(3)	(20)	(9)	(2)	(15)	(6)
Charter	3.53	3.05	2.85	1.75	1.23	1.00
	(15)	(38)	(13)	(11)	(29)	(10)
ASD Cohort 3*						
Charter	1.81	2.71	2.38	0	0.80	0.25
	(21)	(35)	(8)	(14)	(27)	(5)
iZone	3.36	2.76	3.34	1.68	0.71	1.47
	(485)	(313)	(235)	(359)	(237)	(185)
Cohort 1	3.43	2.90	3.18	1.83	0.85	1.26
	(301)	(152)	(113)	(221)	(109)	(86)
Cohort 2	3.35	2.78	3.69	1.03	0.66	2.00
	(158)	(130)	(105)	(187)	(116)	(93)
Cohort 3	2.62	2.00	2.29	0.44	0.31	0.43
	(26)	(31)	(17)	(85)	(47)	(33)
Other Priority**	2.95	2.60	2.83	0.95	0.57	0.80
	(917)	(279)	(163)	(716)	(217)	(126)
Other TN Schools	3.43	3.19	3.19	1.88	1.29	1.30
	(68163)	(9295)	(6058)	(51371)	(6952)	(4562)

Note: Number in parentheses are total number of TVAAS scores observed.
*ASD Cohort 3 only contains charter-managed schools.
** Averages for Other Priority and Other TN Schools come from three years: 2012-13, 2013-14, and 2014-15.

Overall, the table shows that the ASD recruited teachers with slightly higher average scores than the teachers who left those schools, though those retained (stayers) scored below average. However, the ASD incoming teachers had a slightly higher ratio of those exceeding expectations to those who did not meet expectations than the iZone schools. On the other hand, iZones retained and recruited teachers with higher scores than the teachers who left and retained larger ratios of above expectations to below expectations teachers than the ASD schools. With the exception of stayers in Cohort 2, across the ASD cohorts, it appears that the average scores of retained and incoming teachers decline. While the initial cohort of ASD-run schools seems to have attracted very high performing teachers, it also appears that the teachers who exited ASD schools have scored higher on both measures than the teachers that were retained. It may also be worth pointing out that iZone Cohort 3 schools did not attract incoming teachers that were as high performing as the earlier cohorts. Finally, other Priority schools in Tennessee retained very slightly higher performing teachers than ASD and slightly lower performing than iZone, while those teachers exiting Priority schools were lower-performing than those who exited ASD and iZone schools. Priority schools did not attract as high performing teachers as compared to ASD and iZone schools. While management of mobility of teachers who exhibit differential effectiveness is not the only means in which the management of schools can affect outcomes, it is obviously an important means and these patterns may help interpret any effects on achievement that are found.

IV. Data

We use a statewide student-level longitudinal administrative dataset provided to us by the Tennessee Department of Education and compiled by the Tennessee Consortium on Research, Evaluation, and Development. The study time period spans from the 2010-11 to the 2014-15

school years and includes a unique student identifier with the school(s) students attend, the respective grades, student demographic characteristics, and test scores in English, mathematics, and science.[3]

Test scores are included for two different types of exams. Students in third to eighth grades complete the Tennessee Comprehensive Assessment Program (TCAP) test each year in these subjects. TCAP test scores are standardized statewide by subject, grade, and year. At the secondary level, students complete End of Course exams (EOC) for English I, English II, English III, Algebra I, Algebra II, Biology, and Chemistry upon completion of the course, regardless of grade level. EOC test scores are therefore standardized statewide by subject, year, and semester. EOC scores are only included if the student is taking the exam for the first time. English III exam scores were only available for years 2011-12 and after. Chemistry exam scores were only available for 2014-15. For each year, only one test score is included for each student. In cases in which students take TCAP and EOC in the same year, TCAP scores take precedence.

V. Identification of Causal Effects

In the analysis, we make two comparisons. The first is to compare the performance of iZone schools and schools under the auspices of ASD to the business-as-usual schools.[4] iZone schools are governed and managed by their home districts and while these schools receive significant reforms, these do not undergo changes in governance, while ASD schools are removed

[3] For certain validity checks, we include data back to 2007-08.

[4] It should be noted that when Tennessee's First to the Top drafted (as well as the NCLB waiver), the original intent was that iZone and the ASD reforms were meant to be implemented in concert with each other with some schools receiving one treatment and other schools receiving the alternative treatment. These reforms were not designed in a way to create a "horse race" between the two types of reforms. Rather, they were designed in a way to complement each other, and maybe inspire competitive pressure for each reform to be effective. Therefore, from a local perspective, it may not make sense to make a formal comparison between the two. However, from a national perspective, it is important to know which reform is more effective as other states and district may choose among these approaches.

from the governance of the local districts, including both Memphis and Nashville. Therefore, this analysis allows a comparison between district-managed, less intrusive iZone reforms to non-district-managed, more fundamental ASD reforms. In the second comparison, we explore the management of schools further by comparing the performance of schools under three different types of management—district managed iZone schools, the state-managed ASD-run schools, and CMO-run schools.

An ideal approach to the analysis would assign schools randomly to the various treatments and a business as usual control group. However, such an approach is not practical given the constraints of the policy. Therefore, we use what we believe to be the next best approach – a quasi-experimental design using a difference-in-differences (DD) approach. This approach examines pre and post achievement gains of the treatment groups relative to the pre and post achievement gains of similarly low performing comparison groups, which controls for secular trends in low performing schools. The basic assumption of this approach, called the "parallel assumption," is that, conditional on covariates, the average change in outcomes among the treatment would have been the same as the comparison group absent the policy change. While this assumption is not directly testable, a number of validity checks can be conducted that examine whether the DD effect estimates provide credible causal effect estimates (Angrist and Pischke, 2009). Following our findings, we conduct three validity checks for which we have data. First, the pre-intervention trends in the outcomes of interest for the schools in the treated and comparison schools are compared to see if they are parallel prior to the intervention, which would strengthen the credibility of the assumption of parallel changes after treatment absent effects of reforms. Next, the possibility of "anticipatory" or announcement effect is examined. Since the schools to be taken over by ASD or an iZone are named in the prior academic year, the school's performance may drop due to

withdrawal of effort by school personnel, who will have to compete to retain their positions within the school and may reallocate time and effort to seeking other positions. If effort rebounds the following year, the first difference could be exaggerated and appear to be greater than the change in the comparison schools due to anticipatory reactions to turnaround. Finally, we implement a validity test in the spirit of a Granger causality check (Angrist & Pischke, 2009). This check examines differences in outcomes in each of the five years preceding the initiation of turnaround for the ASD and iZone schools. If no differences are observed between either the ASD or iZone schools in years prior to the implementation of turnaround, but differences are found after the intervention, the attribution of the post-intervention differences to the specific turnaround model is strengthened.

Below, we explain the models in greater detail, but before we do, we describe the analytic sample for our study. As noted above, the DD approach requires us to observe the performance of schools both before and after treatment. Therefore, we exclude any school where we did not have pre-treatment test scores for students. In addition, we also exclude any school that did not have tested grades in the relevant school years (e.g., grades K-2). In Table 3, we highlight the total number of schools for each treatment (i.e., Priority, iZone, and ASD, also disaggregated by ASD-run and CMO-run) and the number of schools included in our analysis by year. As the table suggests, the number of ASD and iZone schools have grown over time. In addition, there are still a significant number of priority schools, 28, that have not come under the auspices of an iZone or the ASD.

Table 3. Number of Schools by Reform Approach

Year	Total Priority	Non-iZone, Non-ASD Priority	iZone	ASD Schools in Operation			ASD Schools Included in Analysis		
				ASD-Run	CMO-Run	Total	ASD-Run	CMO-Run	Total
2012-13	82	65	11	3	3	6	3	3	6
2013-14	84[5]	45	22	6	11	17	5	6	11
2014-15	77[6]	28	26	5	18	23	5	11	16

Estimating the Effects of iZone and ASD Turnaround Models

As previously mentioned, some Priority schools were taken over by the ASD, some joined district iZones, and the remaining underwent no systematic reforms other than the requirement to prepare school improvement plans overseen by their local district. To estimate the impact of iZone and ASD reforms, we compare iZone and ASD schools to the last group—Priority schools that were not subject to interventions through the ASD or iZones. To estimate the overall impact of the iZone and ASD schools, we use equation 1:

$$y_{ist} = B_0 + T_s B_1 + A_t B_2 + T_s A_t B_3 + B_4 y_{ist-1} + X_{it} B_j + S_{st} B_k + d_s + g_{it} + e_{ist} \quad \text{(Equation 1)}$$

where the dependent variable *y* represents the test score for student *i* in school *s* in year *t*. We run separate models for each of three dependent variables – reading, math, and science test scores. *T*

[5] The increase in the total number of Priority schools from 2012-13 to 2013-14 comes from the addition of four new ASD schools, the splitting of one school into two separate schools by the ASD, and the closure of three Priority schools.

[6] The decrease in the total number of Priority schools from 2013-14 to 2014-15 comes from the addition of two new ASD schools, the creation of a second school at a former school the ASD took over in 2012-13, the merging of two ASD schools into other ASD schools, and the closure of eight other Priority schools.

is vector of two binary variables indicating whether school *s* was ever in one of the two treatments (i.e., ASD, iZone) between the 2012-13 and 2014-15 school years.

A_t is a vector of binary variables indicating whether treatment occurred in year *t*. Priority school status began in 2012-13. Therefore, for our comparison schools, Priority schools who have not come under the auspices of the ASD or iZones, this vector takes a value of 0 in all years. For the iZone and ASD, because they took over schools through a phase-in process, assigning the values of 1 or 0 for these schools is complicated. For both ASD and iZone schools, the first possible academic year of treatment was 2012-13. Therefore, the binary variable indicating iZone status has a value of 1 for the first cohort of iZone schools in years 2012-13 through 2014-15 and 0 in years prior to 2012-13. Similarly, the binary variable indicating ASD status has a value of 1 for the first cohort of ASD schools in years 2012-13 through 2014-15 and 0 in years prior to 2012-13. However, because new ASD and iZone schools were phased in 2013-14 and in 2014-15, the binary dummy variables are modified for these schools such that it distinguishes between actual reform years and years prior to reform. For both the iZone and ASD, cohort 2 treatment started in the 2013-14. Therefore, for cohort 2 iZone and ASD schools, the binary variable indicating iZone or ASD status has a value of 0 in years prior to 2013-14 and 1 in 2013-14 and 2014-15. Finally, for both the iZone and ASD, cohort 3 treatment started in 2014-15. Therefore, the binary variable indicating iZone or ASD status has a value of 1 in 2014-15 and 0 in prior years. We should note that while iZone phase-ins were complete school phase-ins, many ASD schools only phased in particular grades at a time. Therefore, again, the binary variable for ASD status is modified to reflect the respective grade phase-ins as well as school phase-ins.

In equation 1, we also include a lagged test score control variable, y_{ist-1}, to allow for a value-added interpretation of the dependent variable.[7] X_{it} is a vector of student characteristics for student i in year t, which includes gender, race, free and reduced price lunch status, special education status, and English language learner status. The inclusion of these student-level characteristics improves precision in the analysis. S_{st} is a vector of school level characteristics for school s in year t, which includes the school's percentage of minority students and the school's percentage of free and reduced price lunch students[8]. Inclusion of school-level characteristics improves precision and controls for differences in school-level characteristics, including compositional changes that could be associated with student outcomes. d_s allows for a school fixed effect, and g_{it} allows for a grade level fixed effect. e_{ist} is an error term. The school fixed effect controls for any lingering time invariant school level characteristics not completely controlled for through the comparison of pre and post achievement gains of treatment and comparison groups through the DD approach. Finally, standard errors are clustered at the school level to account for lack of independence of students within schools.

In equation 1, the coefficients of greatest interest are the coefficients of the vector of interactions between T_s and A_t, which represents an overall iZone and ASD effect for the three post-turnaround years. While the interaction variables give us the overall effect for iZone and ASD schools compared to business-as-usual schools, we also compare the performance of the two treatments. To do this, we test to see if the coefficient estimates of the overall effects for iZone and ASD are statistically different from one another using an F-test.

[7] For TCAP scores, lagged test scores were TCAP scores from the previous year. For EOC scores, test scores from the previous year were not always available as EOCs are taken when students complete the course rather than in a specific grade. Therefore, we use 8th grade TCAP scores as the lagged test score.

[8] Given the transient nature of some students, school level characteristics were calculated based on the enrollment on the first day of the state testing window of each year.

To examine the issue of governance and management further, we conduct a second comparison by examining the governance and management of schools in more discrete categories. In the second comparison, we further refine the comparisons to include an examination of both types of schools under the auspices of ASD and iZone schools to the business-as-usual schools. Therefore, in a modified analysis using equation 1, we have three dummy variables included in the vector $T_{s.}$ –a dummy variable indicating whether a school is ever in the iZone, a dummy variable indicating whether a school is ever an ASD-run school, and a dummy variable indicating whether a school is ever a CMO-run school, Similarly, we modify A_t, to indicate years in which a particular school is part of the iZone, ASD-run, or CMO-run treatment. To compare the performance of three types of turnaround, we again use F-tests to examine whether the coefficients of interest are statistically different from one another.

Comparison Group Balance.

While much of the effect in the DD analysis is driven by the first difference in the treatment group, the second difference plays an essential role in generating plausible causal estimates. As a result, it is important to examine whether the comparison schools represent a strong counterfactual group for each treatment. Therefore, we do pairwise comparisons of the observable student characteristics between each treatment group and the comparison group. These comparisons are analogous to randomized design studies that do "balance checks" of observable characteristics of treatment and control groups in order to provide insight into whether the researchers have evidence that the treatment and control subjects have been randomly assigned (Hoxby et al., 2009; Abdulkadiroglu et al., 2009; Cullen and Jacob, 2009; Engberg et al, 2014; Bifulco, 2012; Zimmer and Engberg, in press). The results, shown in Table 4, suggest that the treatment and control schools are similar on student observable characteristics for the analysis of the iZone and ASD

schools with only one significant difference (in bold) in the percent minority between priority non-ASD, non-iZone schools and iZone schools. Even in this case, the magnitude of the difference is negligible (a difference of 2 percentage points). While finding no substantive differences among observable characteristics cannot exclude the possibility of unobservable differences in populations, the lack of finding significant differences provides some confidence in the appropriateness of the comparison groups.

Table 4. Comparisons of Two Treatment Groups and Business-As-Usual Comparison Group

School Characteristic	Priority – Non-ASD, Non-iZone	Ever ASD	Priority – Non-ASD, Non-iZone	Ever iZone
Proportion Male	0.51	0.52	0.51	0.51
Proportion Minority	0.99	0.98	**0.99**	**0.97**
Proportion FRPL	0.88	0.88	0.88	0.92
Proportion SpEd	0.15	0.15	0.15	0.17
Proportion ELL	0.02	0.02	0.02	0.02
Avg. Reading Score	-0.95	-1.02	-0.95	-1.00
Avg. Math Score	-0.94	-0.94	-0.94	-0.94
Avg. Science Score	-1.11	-1.12	-1.11	-1.16

VI. Results

In Table 5, we estimate the effects of schools under the auspices of ASD and iZone schools (Columns 1, 3 and 5) as well as the iZone, ASD-run, and CMO-run schools (Columns 2, 4, 6). In the table (as well as subsequent tables), we present the results as effect sizes (i.e., estimates are shown as proportions of a standard deviations) as test scores are standardized as previously discussed. To give context to these effect sizes, the average difference between the cutoffs for

basic and proficient achievement levels on the TCAP in the 2013-14 school year was 1.28 standardized units in reading, 1.19 in math, and 1.24 in science.[9]

With these magnitudes in context, the overall effects, reported in columns 1, 3, and 5, for the iZone schools across all subjects can be viewed as positive, statistically significant, and substantively meaningful. In contrast, we do not observe any statistically significant overall effect for ASD schools. When comparing the performance of iZone and ASD schools, we observe a statistically significant and substantively meaningful difference across the two types of schools across all subjects. In columns 2, 4, and 6, when examining the results for ASD-run and CMO-run schools, we do observe small, positive and statistically significant effects in math and science for ASD-Run schools. In four of six cases, we observe a larger and statistically significant gain in achievement test scores for iZone schools than ASD-run (reading and math) or CMO-run (math and science) schools. In examining differences among CMO- and ASD-run schools, we do not find any statistically significant differences (although the difference in the estimates for science effect is very close with a p value of 0.06).

Table 5. Results – Overall Effects and Effects by Management

	Reading		**Math**		**Science**	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
ASD Overall	0.04		0.01		-0.02	
	(0.03)		(0.07)		(0.07)	
iZone	0.14***		0.24***		0.22***	
	(0.02)		(0.04)		(0.05)	
CMO-Run		0.06		-0.01		-0.09
		(0.04)		(0.10)		(0.10)
ASD-Run		0.02		0.07*		0.12*
		(0.03)		(0.03)		(0.05)

[9] Students can score on one of four levels of each TCAP assessment – below basic, basic, proficient, or advanced. The cutoff for basic is between below basic and basic; the cutoff for proficient is between basic and proficient.

iZone		0.14***		0.24***		0.22***
		(0.02)		(0.04)		(0.05)
Student-Level Demographics and Lagged Test Scores	Yes	Yes	Yes	Yes	Yes	Yes
School-Level Demographics	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Grade fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
p-value of F-Tests						
ASD Overall vs. iZone	0.01***		0.00***		0.01***	
CMO-Run vs. iZone		0.07		0.01***		0.01***
ASD-run vs. iZone		0.00***		0.00***		0.13
ASD-Run vs. CMO-Run		0.27		0.38		0.06
R Squared	0.49	0.49	0.38	0.38	0.39	0.39
Observations	90134	90134	81751	81751	76892	76892

Notes: * significant at 5 percent; ** significant at 1 percent; *** significant at 0.1 percent

Overall, these results are promising for the iZone schools, which are district-managed, less fundamental reforms relative to the ASD. However, it is well known that reforms can take time as research suggests that it takes three to five years for reforms to take hold (Berends, Bodilly, and Kirby, 2002). Therefore, it is worth examining the effects by cohort and academic year since we have three years of post-intervention data for only the first cohorts of ASD and iZone schools. To carry out this analysis, in equation 1, we modified A_t to include each year of treatment (YR_{2013}, YR_{2014}, and YR_{2015}) and T_s each cohort ($ASDC_1$, $iZoneC_1$, $ASDC_2$, $iZoneC_2$, and $ASDC_3$, $iZoneC_3$). These two sets of dummy variables are then interacted together (e.g., YR_{2013}*$iZoneC_1$). In this modified equation 1, rather than an overall or cumulative effect, the interpretation of these coefficients would be different. Since, in essence, we control for the effect for each prior year by cohort (i.e., we control for the effect of cohort 1 in the first year), the effect in the second year for cohort 1 would be considered an effect over and above the cohort 1 effect in 2013. A similar logic would apply to the third year effect for cohort 1. Also, similar logic would be employed to interpret the effect for cohort 2 schools, although these schools have only been in place two years, so it

would only have a first and second year effect. For cohort 3, we will only estimate a first year effect.

To further explore the effects by cohort and by year, we conduct a third variant of Equation 1 in which we further break down these distinctions into the two different management structures offered by the ASD – ASD direct-run and CMO-run. This provides effects by cohort by year by management structure, which allows us to examine whether there are differential effects by state-management versus CMOs.

The results for the two sets of analyses are shown in Table 6. It is important to note that as we break down the effects by cohort, by academic years, by ASD management structures, we decrease power to detect statistically significant results relative to the prior analyses due to limited sample sizes. Focusing first on the iZone and ASD schools by cohort by year in columns 1, 3, and 5, we observe fairly consistent results for iZone schools across years—many of the effects can be deemed as substantively meaningful ranging from effect size of 0.12 to 0.30 of a standard deviation. For the ASD schools, the story is more complex with cohorts in most years having no effect, while other results suggest a positive and statistically significant effect for particular cohorts, in particular years and subjects, and still other results suggest significant negative effect in particular cohorts in particular years and subjects. It is also notable that the positive and negative estimates are generally large in magnitude including a negative math effect of 0.50 standard deviations for year 1 of cohort 2 ASD schools while schools in year 3 in cohort 1 experienced a gain of 0.25 standard deviations in science. Similarly, when breaking down the analysis for ASD schools by CMO-run and ASD-run schools in columns 2, 4, and 6, we again see inconsistent results as we observe mainly statistically insignificant effects as well as four positive and five negative

estimates that are large and statistically significant with more positive and significant effects for the ASD-managed schools.

Table 6. Results – Overall Effects by Cohort by Year and Effects by Management by Cohort by Year

	Reading		Math		Science	
	(1)	(2)	(3)	(4)	(5)	(6)
ASD Cohort 1, Year 1	0.00		0.08		0.22*	
	(0.04)		(0.06)		(0.10)	
ASD Cohort 1, Year 2	0.04		0.06		0.00	
	(0.08)		(0.09)		(0.13)	
ASD Cohort 1, Year 3	0.04		0.04		0.25*	
	(0.04)		(0.15)		(0.12)	
ASD Cohort 2, Year 1	-0.06*		-0.50***		-0.22*	
	(0.12)		(0.08)		(0.16)	
ASD Cohort 2, Year 2	0.04		0.18		0.15	
	(0.06)		(0.07)		(0.12)	
ASD Cohort 3, Year 1	0.01		-0.01		-0.26*	
	(0.04)		(0.24)		(0.12)	
iZone Cohort 1, Year 1	0.12***		0.22***		0.21**	
	(0.03)		(0.04)		(0.07)	
iZone Cohort 1, Year 2	0.13**		0.29***		0.23**	
	(0.05)		(0.05)		(0.07)	
iZone Cohort 1, Year 3	0.12**		0.16*		0.25**	
	(0.04)		(0.07)		(0.08)	
iZone Cohort 2, Year 1	0.18***		0.34***		0.22	
	(0.05)		(0.08)		(0.12)	
iZone Cohort 2, Year 2	0.13***		0.18*		0.30**	
	(0.04)		(0.08)		(0.09)	
iZone Cohort 3, Year 1	-0.04		0.10		0.14***	
	(0.04)		(0.06)		(0.04)	
ASD CMO-Run Cohort 1, Year 1		-0.06		0.07		0.34**
		(0.05)		(0.06)		(0.10)
ASD CMO-Run Cohort 1, Year 2		0.14		0.21		0.10
		(0.11)		(0.11)		(0.12)
ASD CMO-Run Cohort 1, Year 3		0.06		-0.16		0.13
		(0.06)		(0.14)		(0.14)

ASD CMO-Run Cohort 2, Year 1		-0.20		-0.48**		-0.48***
		(0.12)		(0.18)		(0.13)
ASD CMO-Run Cohort 2, Year 2		0.020		0.08		-0.04
		(0.07)		(0.10)		(0.10)
ASD CMO-Run Cohort 3, Year 1		0.01		-0.02		-0.27*
		(0.04)		(0.24)		(0.12)
ASD ASD-Run Cohort 1, Year 1		0.00		0.10		0.19
		(0.04)		(0.07)		(0.12)
ASD ASD-Run Cohort 1, Year 2		-0.05		-0.09		-0.10
		(0.06)		(0.05)		(0.18)
ASD ASD-Run Cohort 1, Year 3		0.02		0.30**		0.41***
		(0.03)		(0.11)		(0.08)
ASD ASD-Run Cohort 2, Year 1		0.09		-0.36***		-0.09*
		(0.11)		(0.04)		(0.04)
ASD ASD-Run Cohort 2, Year 2		0.01		0.37***		0.22
		(0.06)		(0.06)		(0.19)
iZone Cohort 1, Year 1		0.12***		0.22***		0.21**
		(0.03)		(0.04)		(0.07)
iZone Cohort 1, Year 2		0.13**		0.29***		0.22**
		(0.05)		(0.05)		(0.07)
iZone Cohort 1, Year 3		0.12**		0.16*		0.25**
		(0.04)		(0.07)		(0.08)
iZone Cohort 2, Year 1		0.18***		0.34***		0.22
		(0.05)		(0.08)		(0.12)
iZone Cohort 2, Year 2		0.13***		0.18*		0.30**
		(0.04)		(0.08)		(0.09)
iZone Cohort 3, Year 1		-0.04		0.09		0.14***
		(0.04)		(0.06)		(0.04)
Student-Level Demographics and Lagged Test Scores	Yes	Yes	Yes	Yes	Yes	Yes
School-Level Demographics	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Grade fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R Squared	0.49	0.49	0.38	0.38	0.40	0.40
Observations	90134	90134	81751	81751	76892	76892

Notes: * significant at 5 percent; ** significant at 1 percent; *** significant at 0.1 percent

Therefore, we generally conclude that the results for schools under the auspices of the ASD, as a whole and disaggregated by management structure, have been somewhat inconsistent but mainly are not sufficiently precise to conclude that they are different than zero, i.e. not different than the comparison Priority schools. We also do not observe a consistent pattern of these schools improving over time. This is contrast with iZone schools which have shown a more consistent pattern of positive effects overall and across time.

VII. Validity Checks

As we listed above, we test the validity of the parallel trends assumption in three ways: (1) examine pre-intervention trends; (2) test for an anticipatory reaction to turnaround; (3) "Granger" test for differences in the five years prior to turnaround. In addition to these checks, we also test for intervention effects on the makeup of these schools (school-level covariates used as adjustments in the DD models) that may signal a change in the desirability of enrolling in these schools. Finally, we provide a check on efforts to cream-skim or push out certain groups of students such that student mobility may bias the effect estimates. Critics of charter schools have often claimed that charter schools try to improve the academic profile of their schools as well as reduce costs by recruiting high-ability students and pushing out low-ability students (Ravitch, 2010).[10]

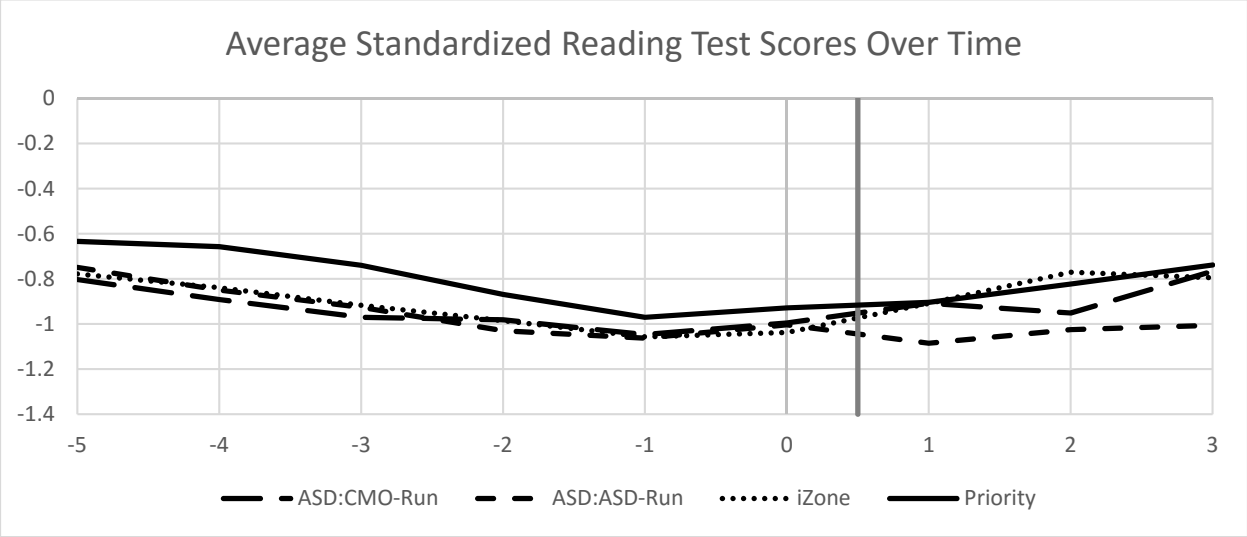
First, we examine the trends in the outcomes of interest with particular interest in the trends prior to implementation of turnaround. If outcome trends were parallel prior to turnaround it supports that the trends were likely to be similar after treatment implementation, except for the response to turnaround. Unfortunately, the current state achievement tests were only employed

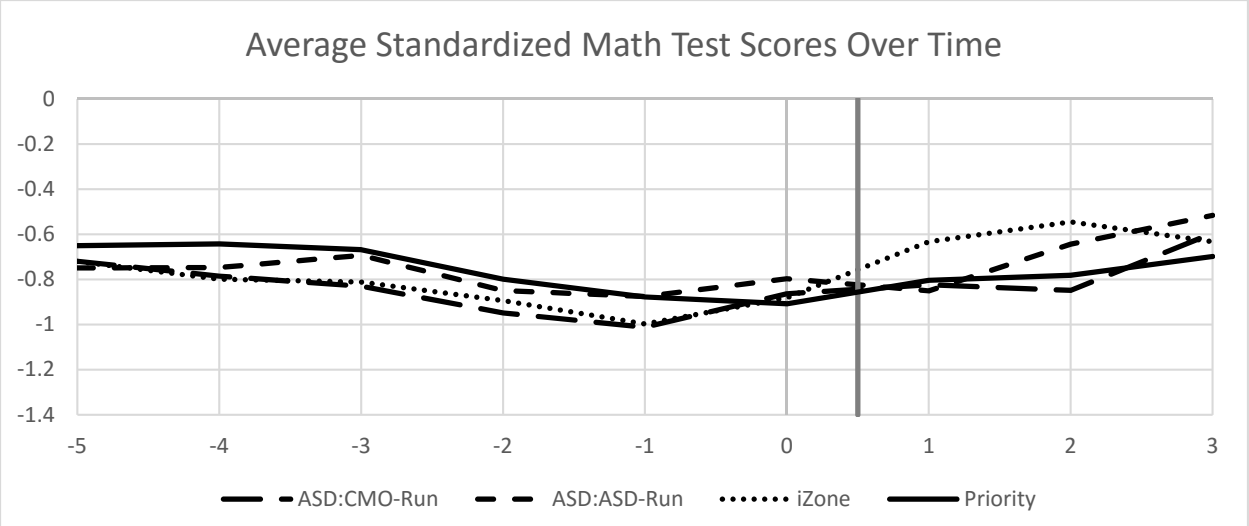
[10] A few studies have examined the cream skimming and pushout question for charter schools (Booker et al., 2005; Zimmer et al., 2011; Zimmer and Guarino, 2013; Winters, 2015; Nichols-Barrer et al., 2012). Across these studies, the researcher have generally found little evidence of cream skimming or pushing out low performing students.

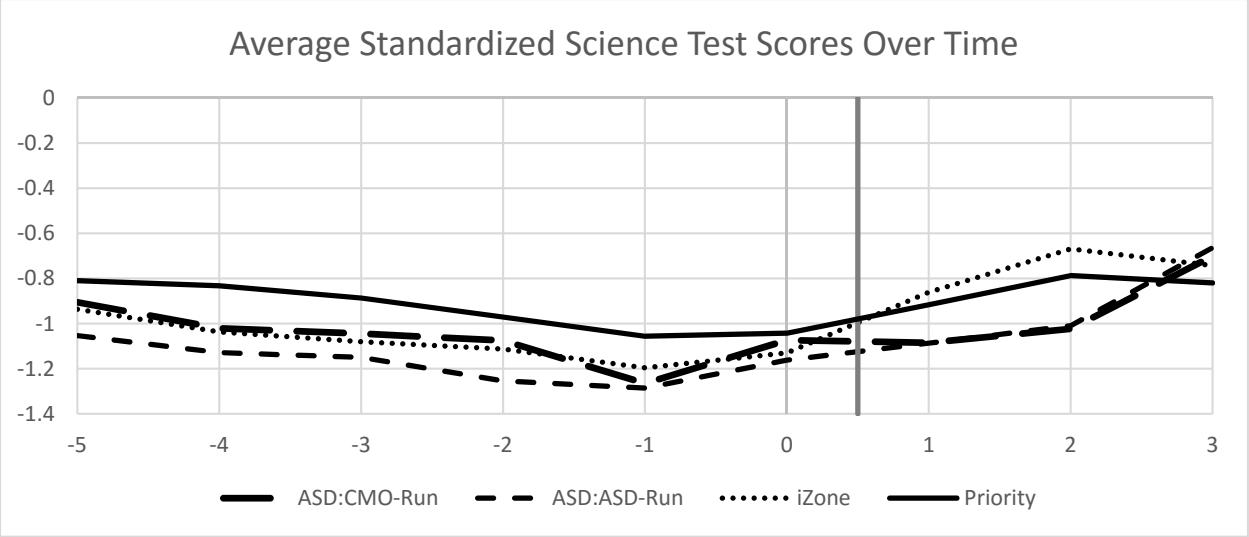
three years prior to the year treatment started as the state adopted a new state test in the 2009-10 school year. However, we do have student-level data dating back to the 2006-07 school year using the prior state accountability test. Because these tests employ different standards, it may not be appropriate to use these tests as outcome measures in our primary analyses. However, we argue that these tests can be useful in examining pretreatment trends as a validity check.[11] To implement this check, we again standardized both the previous and current standardized tests as previously described so that the tests are on a common metric. Because of the phase-in process of the ASD and iZone schools and to maintain a large sample size, we center the years such that Year 0 is the year prior to treatment and Year 1 is the first year of treatment. The results for the trend analyses are shown in Figure 1. The pretreatment trends of the treatment groups and control schools are relatively similar, with slight deviations in year 0 for math (all treatment groups crossed the comparison group between year -1 and year 0 and some crossovers among the treatment groups in year -1 in science but no crossovers with the comparison group and any treatment group). Overall, the pre-treatment trends provide strong support for the parallel assumption for the reading results, but may raise slight concerns for the math and science results. However, the magnitude of the change in trends for iZone schools pale in comparison to the effects we observe during treatment.

[11] We conducted an analysis in which we include the student-level data back to 2006-07 school year standardizing by year, by grade across the two different test. The results of the primary analysis is robust to the inclusion of the additional years of data.

Figure 1. Pretreatment Trends




Average Standardized Math Test Scores Over Time
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
-5
-4
-3
-2
-1
0
1
2
3
ASD:CMO-Run
ASD:ASD-Run
iZone
Priority


Average Standardized Science Test Scores Over Time
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
-5
-4
-3
-2
-1
0
1
2
3
ASD:CMO-Run
ASD:ASD-Run
iZone
Priority

Second, we tested for an anticipatory or announcement effect by artificially assuming that a treatment begins a year before treatment it actually began and using the same model implemented for estimating the effects (Eqn. 1). If the anticipation of turnaround causes test scores to drop in the year before turnaround begins, a rebound in scores could masquerade as a treatment effect in the first year of treatment. It appears that the possibility of an "announcement" effect has been largely ignored by most of the contemporaneous papers examining similar school reform policies, which could potentially have an upward bias of their estimates. In examining the results in Table 7, we do not observe any statistically significant effect—either positive or negative. These results should minimize concerns about an announcement effect and provide some support for the DD approach.

Table 7. Anticipatory or "Announcement" Validity Check: Estimates of Effects in the Year Prior to the Implementation of Turnaround

	Reading		**Math**		**Science**	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
ASD Overall	0.07		0.04		0.04	
	(0.04)		(0.09)		(0.08)	
iZone	0.03		0.09		-0.05	
	(0.03)		(0.09)		(0.04)	
CMO-Run		0.09		0.03		0.04
		(0.05)		(0.10)		(0.09)
ASD-Run		0.02		0.07		0.01
		(0.06)		(0.11)		(0.12)
iZone		0.03		0.09		-0.05
		(0.04)		(0.09)		(0.04)
R Squared	0.48	0.48	0.36	0.36	0.35	0.35
Observations	38488	38488	35162	35162	32706	32706

Notes: * significant at 5 percent; ** significant at 1 percent; *** significant at 0.1 percent

While the check for an anticipatory reaction using one year of prior achievement gains undermines the possibility of an announcement effect, which could bias the effect estimates upward, a Granger test was conducted to assess the possibility that differences could have existed in prior years, which may reappear after turnaround occurred. We added data for ASD, iZone and comparison schools for the period five years prior to the initiation of treatment and included a maximum of three years in which any of these schools received treatment. Because we added data described above from a different state testing regime, we conducted this test separately from the test of anticipatory effects. The results which appear in Table 8, show that no statistical differences were found between ASD and the comparison schools or iZone and the comparison schools for the five years before they entered treatment for the reading or math tests, and in the case of ASD for the science tests, which strengthens the credibility of the DD estimates as causal effects. For iZone, the science test score gains are approximately 10 percent of a standard deviation below those for the comparison schools two years and one year prior to the implementation of turnaround. The interpretation of the significant effects on science in five of the six iZone comparisons presented in cohort by year analysis may need to be tempered in light of the Granger test and the pre-intervention trend examination.

Table 8: Differences in Achievement Gains in either ASD or iZone Schools and the Other Priority Schools from Five Years Prior to Turnaround through Three Years After

	Reading	**Math**	**Science**
ASD_{-5}	0.017	0.026	-0.068
	(0.027)	(0.057)	(0.047)
ASD_{-4}	0.022	0.091	0.012
	(0.041)	(0.061)	(0.055)
ASD_{-3}	-0.009	-0.051	-0.020
	(0.051)	(0.083)	(0.072)
ASD_{-2}	-0.002	-0.012	-0.129
	(0.031)	(0.069)	(0.068)
ASD_{-1}	0.030	0.054	-0.019
	(0.052)	(0.073)	(0.062)
ASD_{1}	0.007	-0.035	-0.084
	(0.048)	(0.097)	(0.116)
ASD_{2}	0.046	0.173	-0.001
	(0.052)	(0.068)	(0.086)
ASD_{3}	0.072	0.097	0.178
	(0.041)	(0.157)	(0.123)
$iZone_{-5}$	-0.007	-0.025	-0.076
	(0.041)	(0.047)	(0.043)
$iZone_{-4}$	0.021	0.048	-0.033
	(0.050)	(0.058)	(0.047)
$iZone_{-3}$	0.006	0.020	-0.033
	(0.042)	(0.060)	(0.050)
$iZone_{-2}$	-0.062	-0.082	-0.108*
	(0.042)	(0.066)	(0.048)
$iZone_{-1}$	-0.072	0.031	-0.128*
	(0.046)	(0.068)	(0.062)
$iZone_{1}$	0.041	0.17*	0.067
	(0.052)	(0.081)	(0.066)
$iZone_{2}$	0.091	0.258***	0.169*
	(0.047)	(0.066)	(0.079)
$iZone_{3}$	-0.009	0.122	0.092
	(0.057)	(0.082)	(0.067)
R Squared	0.45	0.39	0.37
Observations	147,162	133,593	128,523

Notes: * significant at 5 percent; ** significant at 1 percent; *** significant at 0.1 percent

We next examine whether the "treatment" had any effect on the makeup of the schools by replacing the outcome of interest (i.e. test scores) in the DD approach with observable student characteristics aggregated to the school level and examine whether we observe any change in these observable characteristics. Here, we guard against the possibility that the reforms or the act of these schools being identified as part of these treatment may affect the desirability for students to enroll in these schools. In Table 9, we do not observe any statistically significant effects, which again, provides support for the DD approach.

Table 9. Testing for the Effect of Treatment on School-Level Characteristics

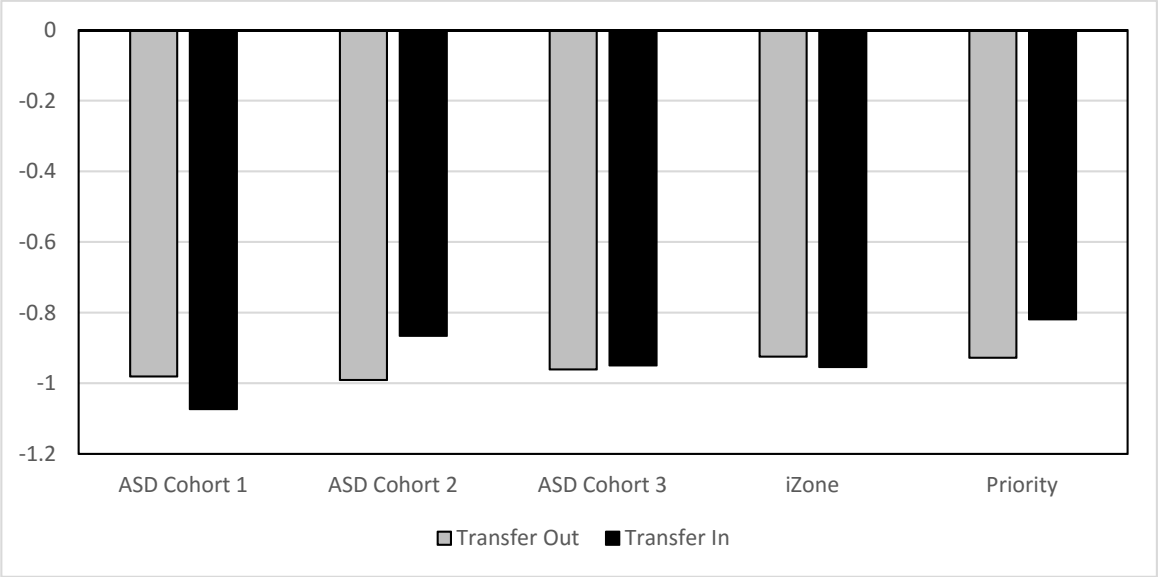
	ASD	**iZone**
Male	0.01	0.00
	(0.01)	(0.01)
Minority	0.01	0.00
	(0.00)	(0.00)
FRPL	-0.01	0.01
	(0.04)	(0.01)
ELL	-0.01	0.00
	(0.01)	(0.00)
Special Ed	-0.01	0.00
	(0.01)	(0.00)

Notes: * significant at 5 percent; ** significant at 1 percent; *** significant at 0.1 percent

Of greatest interest is whether these schools potentially changed the quality of the students by recruiting high-achieving or pushing out low-achieving students. In Figure 2, we examine the relative performance of students either moving in or moving out of priority, ASD, and iZone schools between years. Generally, the average standardized reading score of students transferring into each school type between the 2013-14 and 2014-15 school years is about the same as those

transferring out with slight deviations in Cohort 1 and 2 ASD schools (and the patterns are different across these two cohorts). Only cohort 2 schools raises any concern and the pattern is very similar to priority schools, which suggest that the results from DD analysis should not be biased.

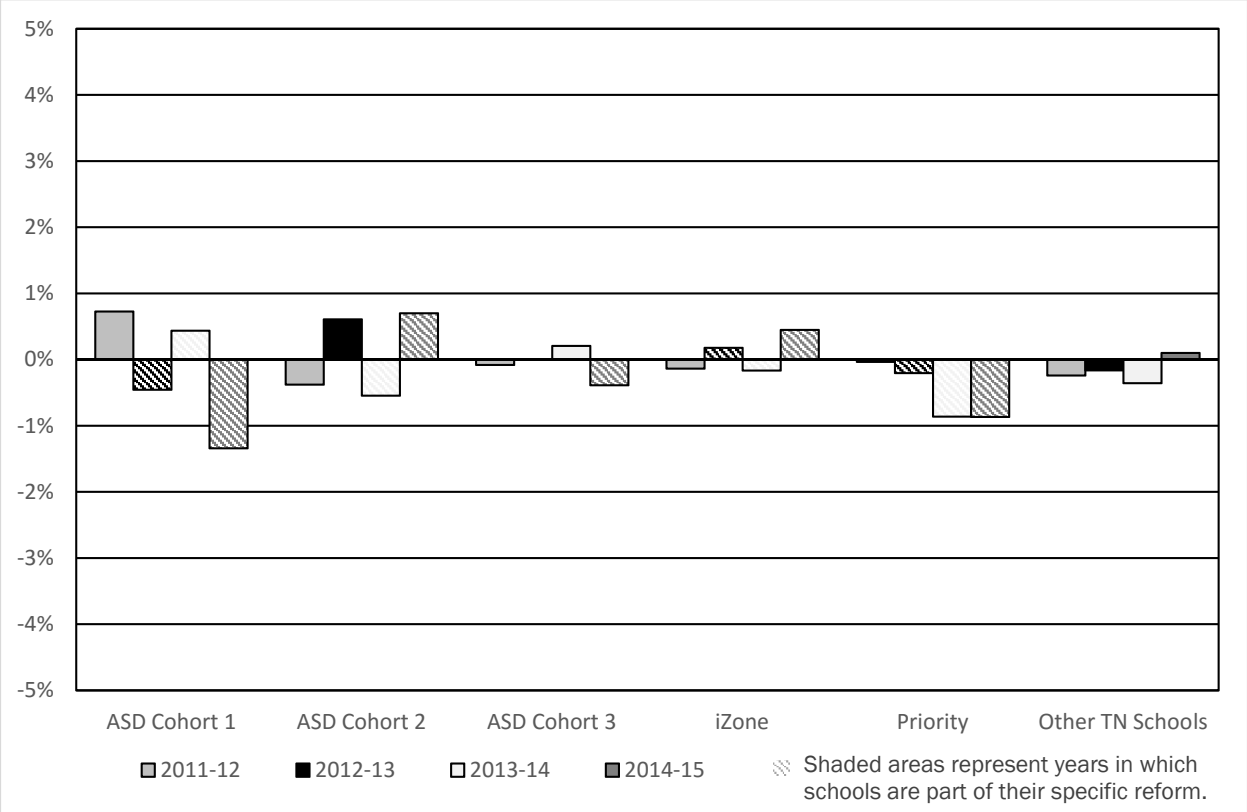
Figure 2. Standardized Reading Scores for Between Year Movers: 2013-2014 to 2014-2015



In terms of gaining understanding of whether any of the above gaps in prior year test scores would make any difference in the academic profile of individual schools, we conduct what we believe is a unique analysis in Figure 3. In the figure, we focus on proficiency of students (since that is what is publicly reported for schools and the measure in which schools are held accountable) and take the number of incoming students that are proficient minus the number of outgoing proficient students and divide this net sum by the total number of students tested. The final value is the net gain or loss of proficient students. Unlike the analysis presented in Figure 2, this analysis takes into account prior achievement scores and the number of the students entering and exiting a

school, which gives a sense of the net impact these moves can have on the school proficiency levels. In other words, the analysis provides insights into whether the school is improving (or diminishing) their academic profile through student transfers. A positive value would be consistent with acting on the distorted incentive of improving school performance through encouraging high-performing students to transfer into the school and/or encouraging low-performing students to transfer out of the school. The figure suggests no major effects from students transferring on the rate of proficiency as there are no cases in which the proficiency level changes (either positive or negative) by more than 2 percent. Overall, the analysis provides no evidence that the schools were strategically recruiting or pushing out students in hopes of improving their academic profile.

Figure 3. Average Effect of Mobile Students on Reading Proficiency Rates Across All Students



VIII. Conclusions

In a speech in 2009, then Secretary Arne Duncan suggested that there were approximately 5,000 chronically under-performing schools in the United States and to transform these schools, we need to institute major interventions, not "tinkering" (Dee, 2012). As a result, the U.S. Department of Education invested in more fundamental and potentially disruptive reforms through RttT and SIG to encourage states and districts to adopt turnaround polices to improve the performance of low performing schools. These reforms focused less on instructional practices and more on governance and management of schools and personnel replacement. Part of the theory for these reforms is that low performing schools do not have the capacity or will to enact fundamental changes to schools, including major changes in staffing (Chubb and Moe, 1990). These reforms often led states to take over low performing schools and either partner with CMOs or mandate that districts provide schools with significant autonomy and put in place significant management, staffing and operational changes. While a number of studies are beginning to examine these reforms, these studies have not compared the effectiveness of the reforms controlled by non-district operators (such as the state or CMOs) with reforms under the auspices of district—e.g., iZones.

In this paper, we examine both types of reforms. We find little evidence for the improvement of schools when removed from district governance and managed by CMOs and evidence of only modest improvement for schools managed by the state. However, we do observe significant improvement for schools that remained under the auspices of the district, but were given greater autonomy. We should note, however, that prior research suggests that reforms take time to take hold and that many of the schools managed either by the ASD or CMOs have been under new management for less than three years. Therefore, it may be premature to draw definitive

conclusions about the effectiveness of management of schools either through the state or CMOs. Nevertheless, the results provide promise for turnaround practices in which district schools are given greater autonomy and use additional resources for recruiting and retaining effective teachers and implement other reforms required under the federal transformation model. While many who advocated for these reforms would argue that the positive results experienced by iZone schools would not have been possible without the pressure created by state takeover and the use of CMOs (Glazer and Egan, 2016), the analysis, at the very least, suggests that it is not necessary for these schools to be managed outside of the districts to experience significant improvement.

Given these results, a number of states currently considering an "ASD-like" approach (including Georgia, North Carolina, and South Carolina) should consider whether it is necessary for schools to be managed by outside providers to experience significant improvement. If these states consider an "iZone alternative" (for instance, given these results, policymakers in North Carolina are also considering an "iZone-like approach"), then it would be helpful to know the mechanisms these schools employed to gain these results. While we lack the data to definitively answer this question, one strategy these schools utilized was an increase in teachers' pay to retain and attract high quality teachers. This does raise the question of whether the iZone approach is scalable if there is a fixed pool of high quality teachers. Evidence consistent with this concern is found in Table 2 which shows that the later cohorts of iZone and ASD schools appear to be keeping and recruiting less effective teachers. In addition, it does raise the question of what effect the approach could have on schools losing the high quality teachers. Both issues are beyond the scope of the paper but the effects of teacher recruitment, retention, and leaving on both turnaround schools and the schools these teachers are recruited from should be the focus of future research.

References

Abdulkadiroglu, A., Angrist, J, Cohodes, S., Dynarski, S., Fullerton, J., Kane, T. & Pathak, P. (2009). Informing the Debate: Comparing Boston's Charter, Pilot and Traditional Schools. The Boston Foundation.

Berends, M., Bodilly,S., & Kirby, S. (2002). "Facing the Challenges of Whole-School Reform: New American Schools After a Decade," Santa Monica, Calif.: RAND Corporation, MR-1498-EDU, 2002. As of January 2, 2007: http://www.rand.org/pubs/monograph_reports/MR1498/

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.

Bifulco, R. (2012). "Can Nonexperimental Estimates Replicate Estimates Based on Random Assignment in Evaluations of School Choice? A Within-Study Comparison. *Journal of Policy Analysis and Management*, 31(3): 729-751.

Bifulco, R., & Ladd, H.F. (2006). The Impact of Charter Schools on Student Achievement: Evidence from North Carolina. *Journal of Education Finance and Policy*, 1(1), 50-90.

Booker, K., Sass, T., Gill, B., & Zimmer, R. (2011). The Effects of Charter High Schools on Educational Attainment. *Journal of Labor Economics*, 29(2), 377-415.

Booker, T. K., Gilpatric, S.M., Gronberg, T.J, & Jansen, D.W. (2007). The Impact of Charter School Student Attendance on Student Performance. *Journal of Public Economics*, 91(5-6), 849-876.

Booker, K., Zimmer, R., & Buddin, R. (2005). The effect of charter schools on school peer composition. RAND Working Paper: WR-306-EDU. Retrieved January 3, 2013, from http://www.ncspe.org/publications_files/RAND_WR306.pdf

Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research*, *73*, 125–230.

Cullen, J.B, & Jacob, B.A, & Levitt, S. (2006). "The Effect of School Choice on Participants: Evidence from Randomized Lotteries" *Econometrica*, 74(5), 1191-1230.

Chubb, J.E.& Moe, T.M. (1990). *Politics, Markets and America's Schools*. Washington, D.C.: The Brookings Institution.

de la Torre, M., Allensworth, E., Jagesic, S., Sebastian, J., Salmonowicz, M., Meyers, C., & Gerdeman, R. D. (2013), Turning around low-performing schools in Chicago. Chicago, IL: The University of Chicago's Consortium on Chicago School Research.

Dougherty, S.M. & Weiner, J. (2015). The Rhode to Turnaround: The Impact of Waivers to No Child Left Behind on School Performance. Presented at the Association of Public Policy and Management, Miami, FL.

Dee. T. (2012). "School Turarounds: Evidence from the 2009 Stimulus." NBER Working Paper 17990, Retrieved from: http://www.nber.org/papers/w17990.pdf

Engberg, J, Epple, D., Imbrogno, J., Sieg, H., & Zimmer, R. (2014) "Identification and Estimating of Treatment Effects when Program Participation is Partially Determined by Lotteries: The Case of Magnet Programs", *Journal of Labor Economics,* Vol. 32(1): 2763.

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

Gill, B., Zimmer, R. Christman, J.B., & Blanc, S. (2007).State Takeover, Restructuring, Private Management, and Student Achievement in Philadelphia Santa Monica, CA: RAND Corporation and Research for Action

Glazer, J.L. & Egan, Cori. (2016) "The Tennessee Achievement School District: Race, History, and the Dilemma of Public Engagement" Tennessee Consortium of Research, Evaluation, and Development. Retrieved March 6, 2016 at: http://www.tnconsortium.org/data/files/gallery/ContentGallery/ASD_Dilemma_of_Engagement.pdf

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

Hanushek, E.A., Kain, J.F., Rivkin, S.G., & Branch, G.F. (2007). Charter School quality and parental decision making with school choice. *Journal of Public Economics,* 91, 823-848.

Hassel, E. A., Hassel, B. C., Arkin, M. D., Kowal, J. K., & Steiner, L. M. (2010). *School Restructuring: What Works When? A guide for education leaders.* Naperville, IL: Learning Point Associates.

Heissel, J., & Ladd, H. (2014). The effects of school turnaround in North Carolina: A regression discontinuity approach. Association for Public Policy Analysis and Management Fall Conference Paper.

Hess, G. A. (2003). Reconstitution—three years later monitoring the effect of sanctions on Chicago high schools. *Education and Urban Society, 35*(3), 300-327.

Henry, G., Zimmer, R., Attridge, J., Kho, A., & Viano, S. (2014). "Teacher and Student Migration in and out of Tennessee's Achievement School District," Tennessee Consortium, Retrieved Sept. 28, 2015 http://www.tnconsortium.org/data/files/gallery/ContentGallery/TCREDASDReport_Zimmer_020915_Revised_022715.pdf

Herman, R., Dawson, P., Dee, T., Greene, J., Maynard, R., Redding, S., & Darwin, M. (2008). *Turning around chronically low-performing schools: a practice guide* (NCEE 20084020). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/ practiceguides.

Hoxby, C.M. & Murarka, S., (2007). *Charter Schools in New York City: Who Enrolls and how They Affect Their Students' Achievement*, Cambridge, Mass: National Bureau of Economic Research. Retrieved November 5, 2008, from: http://www.nber.org/~schools/charterschoolseval/nyc_charter_schools_technical_report_july2007.pdf

Hoxby, C., Kang, J., & Murarka, S. (2009). Technical Report: *How New York City's charter schools affect achievement,* New York City Charter Schools Evaluation Project. Retrieved July 5, 2013 at: http://users.nber.org/~schools/charterschoolseval/how_nyc_charter_schools_affect_achievem ent_technical_report_2009.pdf

Kebede, L. F. (2016, May 2). iZone chief Sharon Griffin on fixing Memphis' most challenging schools. *Chalkbeat*. Retrieved from http://www.chalkbeat.org/posts/tn/2016/05/02/izone-chief-sharon-griffin-on-fixing-memphis-most-challenging-schools/#.VykYnmO5q-R.

MacIver, M. A., and D. J. MacIver. (2006). "Which Bets Paid Off? Early Findings on the Impact of Private Management and K–8 Conversion Reforms on the Achievement of Philadelphia Students." Review of Policy Research 23 (5): 1077–9

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.

Malen, B., & Rice, J. K. (2004). A framework for assessing the impact of education reforms on school capacity: Insights from studies of high-stakes accountability initiatives. *Educational Policy, 18*(5), 631-660.

Malen, B., & Rice, J. K. (2015). School reconstitution as a turnaround strategy: An analysis of the evidence (Working Paper).

Nichols-Barrer, I., Gill, B., Gleason, P., & Tuttle, C. (2012). *Student Selection, Attrition, and Replacement in KIPP Middle Schools,* Mathematica Policy Research, Retrieved January 3, 2013, from http://mathematica-mpr.com/publications/PDFs/education/KIPP_middle_schools_wp.pdf .

Papay, J. (2015). The Effects of School Turnaround Strategies in Massachusetts. Presented at the Association of Public Policy and Management, Miami, FL.

Peterson, P. & Chingos, M. (2007). "Impact of For-Profit and Non-Profit Management on Student Achievement: The Philadelphia Experiment" Program on Education Policy and Governance, Kennedy School of Government, Harvard University, Available at: http://inpathways.net/PEPG07-07_Peterson_Chingos_Philadelphia.pdf

Player, D. & Katz, V. (2013). "School Improvement in Ohio and Missouri: An Evaluation of the School Turnaround Specialist Program" Center on Education Policy and Workforce Competitiveness, Retrieved at: http://curry.virginia.edu/uploads/resourceLibrary/10_Player_SchoolTurnaround.pdf

Race to the Top Application for Initial Funding (2010). Submitted to the U.S. Department of Education. Available at: http://www2.ed.gov/programs/racetothetop/phase1-applications/tennessee.pdf

Rice, J. K., & Croninger, R G. (2005). Resource generation, reallocation, or depletion: An analysis of the impact of reconstitution on school capacity. *Leadership and Policy in Schools, 4*(2), 73-103.

Rice, J. K.,, & Malen, B. (2003). The human costs of education reform: The case of school reconstitution. *Educational Administration Quarterly, 39*(5), 635-666.

Ruble, W. (2015). The Effect of Contracting Out Low Performing Schools on Student Performance. Available at: https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnx3aGl0bmV5cnVibGVicm9zc3xneDo1YzhlNjZkNTI4ODM0NjAx

Sass, T.R., (2006). Charter Schools and Student Achievement in Florida. *Education Finance and Policy*, 1(1), 91-122

Sass, T.R. Zimmer, R., Gill, B., & Booker, K. (2016). "Charter High Schools' Effects on Long-Term Attainment and Earnings." *Journal of Policy Analysis and Management*.

Strunk, K.O., Marsh, J. A., Hashim, A. K., Mecenas-Bush, S. C. & Weinsten, T. (in press) Innovation and a Return to the Status Quo: A Mixed-Methods Study of School Reconstitution. Education Evaluation and Policy Analysis

U.S. Department of Education Office of Elementary and Secondary Education (2010). Guidance on fiscal year 2010 School Improvement Grants under section 1003(G) of the Elementary and Secondary Education Act of 1965. Washington, DC: U.S. DOE.

Zimmer, R. & Guarnio, C. (2013). Is There Empirical Evidence That Charter Schools "Push Out" Low-Performing Students? Educational Evaluation and Policy Analysis, 35, 461–480

Zimmer, R., Gill, B., Booker, K., Lavertu, S., & Witte, J. (2012). Examining charter school achievement in seven states. *Economics of Education Review,* 31(2), 213-224

Zimmer, R., Gill, B., Booker, K., Lavertu, S., & Witte, J. (2011). Do charter schools "cream skim" students and increase racial-ethnic segregation? In Berends, M., Cannata, M., & Goldring, E.(Eds.) *School Choice and School Improvement,* Cambridge, MA: Harvard Education Press, 215-232.

Zimmer, R., & Engberg, J. (in press) Can broad inferences be drawn from lottery analyses of school choice programs? An exploration of appropriate sensitivity analyses. *School Choice Journal*.