



credo

CENTER FOR RESEARCH ON EDUCATION OUTCOMES

Evaluation Methodology of
“Scaling the Charter Restart Model”
2017

Evaluation Methodology of “Scaling the New Orleans Charter Restart Model”

December 2017

Margaret E. Raymond, Ph.D. – Project Director
Lauren Bierbaum, Ph.D. – Senior Research Associate
Sofoklis Goulas, Ph.D. – Senior Research Associate
Lindsay Bell, M.S.W. – Research Associate
Meg Cotter Mazzola – Research Associate
Will Snow, M.A. – Research Associate



credo

CENTER FOR RESEARCH ON EDUCATION OUTCOMES

© 2017 CREDO

Center for Research on Education Outcomes
Stanford University
Stanford, CA
<http://credo.stanford.edu>

CREDO, the Center for Research on Education Outcomes at Stanford University, was established to improve empirical evidence about education reform and student performance at the primary and secondary levels. CREDO at Stanford University supports education organizations and policymakers in using reliable research and program evaluation to assess the performance of education initiatives. CREDO's valuable insight helps educators and policymakers strengthen their focus on the results from innovative programs, curricula, policies and accountability practices.

Acknowledgements

Data for this report were provided by the Louisiana and Tennessee Departments of Education, RSD, NSNO, and ASD. We thank them, the Charter Management Organizations, and schools for their cooperation and active support of the evaluation in all its phases.

CREDO also acknowledges the support and guidance of the Investing in Innovation program staff of the US Department of Education for this research. We especially appreciate the extensive collaboration with the USDE program officer for the project, Brian Lekander.

Disclaimers

The views expressed herein do not necessarily represent the positions or policies of the organizations noted above. No official endorsement of any product, commodity, service or enterprise mentioned in this publication is intended or should be inferred. The analysis and conclusions contained herein are exclusively those of the authors, are not endorsed by any of CREDO's supporting organizations, their governing boards, or the state governments, state education departments or school districts that participated in this study.

The findings and conclusions in this document are those of the authors, who are responsible for its content, and do not necessarily represent the views of the US Department of Education. Any error is attributable to the authors.

Table of Contents

Introduction	1
Organizational Capacity Study	2
Implementation Evaluation	2
Sample	2
Protocol	2
Qualitative Analysis	3
PMO	4
Explanation of original ten elements	4
Explanation of CRM-specific elements.....	6
Explanation of Qualitative Scoring	6
Impact Evaluation	7
Consolidating Student Data from Two States	7
Matched Data	7
Selection of Comparison Observations.....	8
Groups of Students under Study.....	10
Demographic Comparison of students targeted by the intervention with matched students	10
Comparison of Starting Scores of Matched and Unmatched Students	13
Model for Analysis of the Academic Impact of CRM	15
Comparison of Starting Positions of Each Type of Turnaround	17
Regression Output for Fresh Start Vs Full Turnaround Comparison	18
Testing the difference in Effect sizes of Fresh Start and Full Turnaround CRM schools	21
Comparison of Starting Positions of Opt-outs and Persisters.....	21
Mixed Methods Model	22

Table of Figures

Figure 1: CREDO VCR Methodology	9
Figure 2: Comparison of Starting Math Scores of Matched and Unmatched Students across NOLA and TN	13
Figure 3: Comparison of Starting Math Scores of Matched and Unmatched Students in NOLA	14
Figure 4: Comparison of Starting Math Scores of Matched and Unmatched Students in TN	14
Figure 5: Comparison of Starting Reading Scores of Matched and Unmatched Students across NOLA and TN	14
Figure 6: Comparison of Starting Reading Scores of Matched and Unmatched Students in NOLA	15
Figure 7: Comparison of Starting Reading Scores of Matched and Unmatched Students in TN	15

Table of Tables

Table 1: Student Match rates by Location.....	9
Table 2: Demographic Composition of All ITT and CRM Students in the Evaluation	11
Table 3: Demographic Composition of ITT Students from Closing Schools in the Study.....	11
Table 4: Demographic Composition of CRM Students in the Study	12
Table 5: Comparison of CRM and Closing Schools-Overall	16
Table 6: Comparison of CRM and Closing Schools-New Orleans	16
Table 7: Comparison of CRM and Closing Schools-Tennessee	17
Table 8: Comparison of Starting Position for Persisters and New Entrants-Overall.....	18
Table 9: Comparison of Starting Position for Persisters and New Entrants-New Orleans.....	18
Table 10: Comparison of Starting Position for Persisters and New Entrants-Tennessee	18
Table 11: The effect of the CRM by type of Turnaround.....	18
Table 12: The effect of the CRM by type of Turnaround in Reading	19
Table 13: Comparison of Fresh Start and Full Turnaround CRM Schools-Overall.....	21
Table 14: Comparison of Fresh Start and Full Turnaround CRM Schools-New Orleans.....	21
Table 15: Comparison of Fresh Start and Full Turnaround CRM Schools-Tennessee	21
Table 16: Baseline achievement of opt-outs and persisters at the first year of operation.....	22
Table 17: Operational Factors.....	22

List of Acronyms & Definitions

CMOs	Charter School Management Organizations – To qualify as a CMO, an organization must oversee the operation of at least three charter schools. Further, the CMO must be the charter holder for all of the schools operated by the CMO.
CREDO	Center for Research on Education Outcomes
EOC	End-of-Course Exam
ELA	English Language Arts
ELLs	English Language Learners
CSGF	Charter School Growth Fund
NAEP	National Assessment of Educational Progress
TPS	Traditional Public School
VCR	Virtual Control Record
CRM	Charter Restart Model
Achievement	Academic performance at a fixed point in time, consisting of content knowledge and cognitive skills. Achievement is measured against learning standards in absolute terms. Students are compared to each other in relative terms as positive or negative based on whether a student above or below the average student.
Growth	The year-to-year change in academic performance relative to one’s peers. The median student is used to benchmark the progress of other students.

Evaluation Methodology of the Charter Restart Model 2017

Introduction

The *Scaling the New Orleans Charter Restart Model* project was a five-year collaboration between the Recovery School District of the State of Louisiana and New Schools for New Orleans, a nonprofit charter school support organization funded by the Investing in Innovations (i3) program of the Office of Innovation and Improvement of the United States Department of Education. An independent evaluation of the collaboration was a requirement of the grant to the Recovery School District (RSD) and New Schools for New Orleans (NSNO) which served as the agency of record for the project. With the support of the US Department of Education, the local program partners and funders, the original five-year evaluation design was extended for an additional two years to permit longer follow-up of the program's impact on student academic performance. The evaluation was limited to the specific components of the Charter Restart Model which operated within a larger set of policies and practices in New Orleans. It did not extend to the full array of reform efforts that have occurred. Instead, its focus was on the fidelity to the project design and implementation of its various components over the course of the five-year grant period.

The evaluation consisted of three related studies. An Implementation Analysis tracked the transitions of schools as they moved through the charter restart process and reemerged as schools in a charter management organization (CMO) portfolio. The focus of that study was a process evaluation consisting of repeated interviews and observations at the school level to identify the milestones and challenges that new teams of educators and leaders faced when restarting a school. The Implementation Analysis also introduced the use of the Performance Management Organization (PMO) rubric, a tool used to assess implementation quality among the funded schools. The second study is an Impact Analysis focused on the academic progress and other outcomes of the students whose schools underwent transition via the Charter Restart Model. Careful design of methodology and analysis yields insight into the paths students took and the impact their choices had on their learning. The third and final study of the evaluation is an Organizational Capacity Analysis of the key program partners of the Charter Restart Model: RSD, NSNO and the CMOs tapped to lead the transitions of schools. The third study also included the Achievement School District in Tennessee and the CMOs in Memphis and Nashville that served as the site for replicating the Charter Restart Model in a new setting as a preliminary test of the scalability of the model.

Organizational Capacity Study

The Organizational Capacity study drew on four years of repeated semi-structured interviews with leaders and staff of RSD, ASD, NSNO, as well as the CMO leaders of the CRM schools. These interviews, conducted once yearly, addressed perceptions, expectations, and actions of the various systems-level players regarding the CRM's installation and maturation.

Organizational Capacity interviews were analyzed for common themes across time, respondents, and organizations. These findings were then assessed against the McKinsey & Company Organizational Capacity Assessment Tool (OCAT), in order to determine the degree to which the CRM, as designed, represented adequate capacity to achieve specified goals; and to determine the degree to which systems-level partners (RSD, ASD, NSNO) were equipped to implement the CRM with fidelity.

For a full discussion of the Organizational Capacity study approach, see the [Organizational Capacity Report](#).

Implementation Evaluation

CREDO conducted the Implementation Study using semi-structured interviews, extensive site observations, and online surveys.

Sample

Respondents for the Implementation Evaluation included CRM principals, teachers, board members, parent coordinators, special education coordinators, and school business managers. Board members, parent coordinators, SPED coordinators, and business managers were each interviewed a single time, in their schools' second year of operation. Interviews of school principals and teachers occurred twice (Fall and Spring) in every year for all schools in the evaluation. In total, we interviewed 36 principals across 21 schools (due to frequent leadership turnover and two interim maternity leave placements). CRM Principals were 58% female and 59% non-white (56% black, 3% Latinx). We conducted 304 CRM teacher interviews across the five years of qualitative data collection. Teacher, board member, parent coordinator, SPED coordinator, and business manager interviews were recorded anonymously, so we do not have specific demographics for these respondents. Similarly, online surveys were administered to all teachers and principals in CRM schools once yearly for each year of the study, but responses were recorded anonymously.

Protocol

CREDO conducted a single site visit to Closing and Flagship schools in the semester preceding the opening of the paired CRM school.¹ These site visits each consisted of: an interview with the school principal; interviews with two classroom teachers; a site observation; and two classroom observations. Interview protocols were semi-structured, containing rating scales, forced choice responses, and open-ended questions. Closing school protocols elicited respondents' experiences and opinions regarding their schools, their students, the reasons for their schools' chronic under-performance, and the turnaround process.

¹ KIPP East Community School did not have a Closing school visit. Wilson did not have Closing or Flagship school visits.

Flagship school interviews explored respondents' experiences and opinions regarding their schools, their students, their CMOs, and their expected relationship to their CRM school.

Site observation protocols did not differ across Closing, Flagship, and CRM schools. These observation protocols included observations of schools' physical plants, surrounding neighborhoods, and facilities. School sites were rated on both quality of facilities and availability of particular infrastructure (playground equipment, green space, sports fields/courts, safe drop-off/pickup areas, etc.). Site observations also recorded information regarding school climate and culture. In addition to site ratings, we also conducted classroom observations in two classrooms during every site visit. Classroom observations recorded both physical environment and pedagogical practice. In total, CREDO conducted 126 site observations and 252 classroom observations over the course of the Implementation study. CREDO researchers undertook extensive efforts to ensure interrater reliability across observations.

The semi-structured interview protocols contained rating scales, forced choice responses, and open-ended questions. Interview protocols elicited respondents' experiences and opinions regarding their schools, their students, their CMOs, and the project of school turnaround more broadly. All interview responses were recorded by hand on paper in the field, then transcribed to electronic media (MS Excel or MSWord).

In order to supplement interview findings, we also conducted online surveys once per year with CRM teachers and principals, generating 331 responses total. Surveys were administered using Google Forms. However, for each survey administration, response rates were low, under 50% each year.

Qualitative Analysis

Qualitative data were analyzed using both deductive and inductive coding schemes. Our deductive coding scheme was typical of deductive coding in evaluation work, stemming from a set of research-driven hypotheses regarding the key thematic areas of importance in school improvement. Our predetermined codes represented the four major areas of inquiry in the evaluation: School Mission and Philosophy, Human Capital, Performance Management, and Commitment to Continuous Improvement.

Additionally, we allowed for inductive coding from the onset of data collection in order to capture emergent structures within the data set. Inductive coding produced two key additions to our analytic framework: First, inductive coding resulted in the identification of two additional areas of inquiry: Governance and Systems, which encapsulated issues unique to charter school organizational and accountability structures; and Exogenous Shocks, which encapsulated the changes and challenges CRM schools encountered in the volatile post-Katrina environment (in New Orleans), in the wake of the installation of a state takeover mechanism (in Tennessee), and due to community poverty in schools' respective locales. Second, inductive coding allowed for the generation of sub-codes within our six primary codes (two inductive and four deductive).

Coding occurred with annual data sets as well as within the full longitudinal dataset. Early coding processes used Microsoft Excel but as the dataset grew, analyses transitioned to Atlas.ti. Individual interview or observational items were autocoded in alignment with the four original deductive codes; responses were also all manually coded to increase precision of code structures and to apply sub-codes. CREDO employed

three coders, all with previous qualitative research experience. Coders were trained specifically in this study for interrater reliability on observational scales and interrater agreement on narrative responses.

Once the full longitudinal data set was coded, findings by code were then mobilized to answer a set of [nine research questions](#). These nine research questions evaluated *Scaling the New Orleans Charter Restart Model* against the CRM's three overarching goals: to build capacity for incubation and expansion of charter operators; to create permanent infrastructure to sustain the CRM; and to scale the CRM by codifying and replicating the model.

Finally, qualitative findings were also used to create index variables around 19 operational factors. These index variables were then included in regression and correlational analyses in order to test the relationships between operational factors observed in the qualitative Implementation Study with quantitative student learning outcomes from the Student Impact study.

PMO

The Performance Management Organization (PMO) rubric is a composite index of implementation quality. CREDO developed the PMO rubric in 2008 to map discrete organizational practices against a framework of proven elements of performance management. The original Performance Management Organization rubric score examined charter school operations in ten elements divided into 23 sub-elements and rated each school in a consistent manner. In addition to this original rubric, CREDO developed four additional elements and 14 sub-elements in order to test operational domains identified in previous work as worthy of amplification; and/or domains which are particularly germane to the Charter Restart Model.

Interview and observational protocols used in school site visits throughout the course of the study included items designed specifically to capture point-in-time information of school effort in these 14 areas. This information, organized by functional area, was then rated on a scale from zero to three: *Absent (0)*, *Developing (1)*, *Accomplished (2)*, *Exemplary (3)*. Scores were assigned by CREDO researchers with extensive expertise in school quality and improvement. Researchers underwent interrater reliability training at the onset of the study, and had periodic checks for interrater agreement in each successive year. See below for a more detailed explanation of PMO scoring.

Explanation of original ten elements

1. Even before school opens, there is an early focus on academic growth and achievement. We would expect charter schools to be sensitive to the outcomes they create for their students, based on their charters' explicit focus on accountability for results. Integrating a specific focus on achievement and growth into the structure and function of the school from its inception provides a critical organizing principle for all the operating decisions that must be made over the life cycle of the school. Making student results the crystal-clear focus of the school's endeavors gives everyone a common base for evaluating choices. School leadership and staff use common definitions of achievement and growth and are clear about what targets apply in each time period.

2. The school leader is able to create a common vision in which the school's raison d'être is to support the success of all students. As the chief executive officer of the school, the school leader acts as general and quarterback. While it is necessary for the school leader to have a clear idea about results, it is not sufficient

for only the school leader to hold this vision. S/he must also be able to create that same clarity in all other adults in the school.

3. Leadership in the school is stable and consistently focused on academic results. In the early years of the charter school, design concepts meet reality, often requiring extraordinary time and energy. Because all aspects of the school are new, there is an inevitable period of flux and care is required to ensure that the frenzy of start-up resolves into smooth operations. Stability in school and board leadership helps provide a rudder to temper some of the uncertainty, but persistence alone is not enough to assure high performance. Leadership also needs to supply a steady and persistent signal about the job at hand.

4. Teachers in the charter school are fully vested in the vision of student academic results and show ownership of their crucial role in that endeavor. To create a high-performing organization, every teacher must believe that all children regardless of background can achieve academically. Further, teachers must also believe in their competencies to create those results in the students they teach.

5. There is a consistent and positive attitude among educators in the school about the value of interim assessments as performance signals of student progress. One of the most prevalent features of high performing organizations is the use of regular information about the organization and its progress towards achieving desired goals. In the education setting, the necessary signals are provided by periodic assessment with formative tools. Their value, however, depends to a great degree on a collective endorsement by school leaders and teachers.

6. There is a deliberate choice in the use of performance signals through the selection of interim assessment tools. The choice of interim assessments is monitored, and quality evaluated over time; adjustments in use or changes to different interim assessment tools are made to improve the quality of the performance signal that schools receive. Even where school staff agree on the importance of interim assessments, considerable choice exists among assessment tools. The degree to which adults agree on a particular instrument, the degree to which they gauge the fit between an instrument and student abilities, and the capacity to change tools when the need arises all play into the superstructure of performance management in schools.

7. Assessment results trigger diagnostic responses; educators use assessment results as a stimulus to regular reflection and interaction within the school. Having a clear signal of results through interim and summative assessments is certainly important, but school personnel must also respond to the signal. In performance-focused organizations, assessment results are regularly employed as a convening device to guide action. Such action can take several possible forms. For example, principals might regularly schedule meetings with teachers to drill down into assessment results and see where specific lessons need revision. Alternatively, results of assessments might inform the topics for professional development, or provide the foundation for improvement bonuses for new teachers. The choice of processes or form of action may differ from school to school, but in each case the assessment results trigger routines that are designed to determine what changes are needed in existing practices and programs.

8. The results of interim assessments are used to refine strategies of instruction. Unless the results of interim assessments are routinely exemplary, schools that are PMOs can make mid-course adjustments towards improvement, directed by the insights provided by interim assessments. Even when the interim

assessment results are good, they affect instruction by reinforcing those pedagogical approaches that have been used.

9. Resources (time and budget) are allocated to effective professional development to enhance the capacity of educators to successfully create learning gains in their students. Investment in human capital development for those in pivotal roles is one of the fundamental tenets of high performing organizations. Teacher quality is the single most important school factor for high academic achievement and growth; it is up to the school leadership to assure that they master essential knowledge and skills needed to create successful outcomes for students.

10. The organization is structured as a continuous improvement organization. Many high performing groups create an organizational commitment to high achievement. A few create the ethos of continuous improvement. Without it, organizations tend to either under- or over-react to any information that they are not achieving the desired level of results. Some take a resigned stance, including blaming the result on poor inputs. Others reach for drastic revisions, thereby destabilizing the organization; in extreme cases, the schools implode. Schools that integrate continuous improvement into their operations embrace the notion that successive approximations are not only acceptable, but they are also the surest way to reaching the desired outcomes.

Explanation of CRM-specific elements

10. There is a clear and focused behavior management plan in place and it is used consistently across the school. High expectations and consistent implementation of behavior management strategies help create an organized classroom environment that is conducive to teaching and learning.

11. The school provides a wide array of engaging opportunities for students to become further involved in the school. Activities outside of the classroom provide an outlet for students to apply their academic and social skills in a new context. Participation in non-classroom activities will increase student's sense of engagement to the school.

12. Branding plays a significant role in providing organizations with a unique identity. The first impression on the community, parents, and students is driven by branding. It is the reflection of the school's mission, values, practice and professionalism.

13. Financial Accountability is essential for organizations to be successful. A dedicated personnel position for financial management and proper financial oversight from the governing board of directors are pivotal functions of a high performing organization.

Explanation of Qualitative Scoring

The PMO rubric uses a four point scale (0 – Absent; 1 – Developing; 2 – Accomplished; 3 – Exemplary) to rate each school in each element or sub-element. We use a four-point scale intentionally — raters are required to make a fundamental judgment about which side of the midpoint a school performs. Further, the small number of categories helps raters apply the rubric with greater confidence. Despite the limited scoring set,

the rubric has performed well. The ratings for each element or sub-element in the model were grounded in concrete and objective criteria to make the assignment of a school's data unambiguous. To the extent supported by the data, the progress of each school over the years of the evaluation was mapped and then an overall element score (from 0 – 3) was identified. The validity of the rubric was checked by multiple analysts coding the data to each category. The results for each school reveal the path they have followed to date in developing the school capacity in each element area.

In each of the 14 areas, ratings were compiled, then multiplied by 100 (to eliminate decimals). The final score transformed to a scale of zero to 100%. Each of the 14 scores maps to the rubric and is scored independently; strong scores in one area cannot offset weak scores in another. Once each school received scores in all 14 areas, those 14 scores were averaged to create the school's final PMO score.

PMO scores represented an assessment of implementation quality, and as such were reported independently [here](#) in the full online report. PMO scores were also tested for [association with student impact](#).

Impact Evaluation

Consolidating Student Data from Two States

To create a student dataset for the impact study, CREDO worked with the state departments of education in Louisiana and Tennessee. Each state administered its own student testing program for school accountability purposes; the tests differed by grade, by subject and by year. Accordingly, it was necessary for CREDO to standardize the individual student test values to make them comparable. CREDO did this by creating a bell curve for each test – by subject, grade, state, and year – where the average student score on the test became central value and all other scores were distributed around it. The transformation placed each student's performance in relation to all other equivalent tested students, making it ready for comparison with other students. By comparing each student's performance relative to the other students from one year to that same student's relative performance in the next year, CREDO could estimate if the student was growing academically at a rate which was faster, similar or slower than the rate of the peers.

CREDO was able to combine growth results from multiple grades, states, and years. Even though the average academic achievement in state A might differ from the average academic achievement in state B, a change in the academic achievement (growth) of 0.05 standard deviation in state A and a .05 standard deviation change in achievement in state B both represented the same level of improvement relative to their peers in the student's home states. This is one of the reason measures of academic growth are superior to simple measures of academic achievement, the level of which can vary greatly from state to state.

Matched Data

When examining the impact of CRM on student academic progress, we used the virtual control record (VCR) method developed by CREDO (Davis & Raymond, 2012). The first step in conducting a VCR analysis is to create a matched data set. The matched data set consists of treated students (in the case students in closing schools in their last year of operation, students attending a CRM school, students from the closing school that were ineligible to enroll in a CRM school, students from the closing school that chose not to enroll in a

CRM school, students that attended a CRM school in the past but left²) and demographically identical students in the control group.

Selection of Comparison Observations

A fair analysis of the impact of CRM requires a comparison group which matches the demographic and academic profile of students in CRM schools to the fullest extent possible. As in previous CREDO studies, this study employed the virtual control record (VCR) method of analysis developed by CREDO. The VCR approach created a “virtual twin” for each CRM student who was represented in the data. If matched, this virtual twin would differ from the CRM student only in that one student attended a CRM school. The VCR matching protocol has been assessed against other possible study designs and judged to be reliable and valuable by peer reviewers (Fortson, Verbitsky-Savitz, Kopa, & Gleason, 2012, Ackerman, & Egalite, 2017)

Using the VCR approach, a virtual twin was constructed for each CRM student by drawing on the available records of students with identical traits and identical or very similar baseline test scores but who were enrolled in a continuing low-performing school in same area. Factors included in the matching criteria were:

- Grade level
- Gender
- Race/Ethnicity
- Free or reduced-price lunch eligibility
- English Language learner status
- Special education status
- Test score on state achievement tests in the year prior to first enrollment in the CRM school.

Figure 1 shows the matching process used by CREDO to create the virtual twins linked to each CRM student. In the first step, CREDO identified all low-performing schools that were permitted to remain open by state and year. These schools in the same state and identified as low-performing for the final year of operation for a closed school were referred to as “comparison schools” for that particular CRM school. Once a school was identified as a comparison school for a particular CRM school, all the students in that low-performing school became potential matches for students in that particular closed school. All of the student records from all of a CRM’s comparison schools were pooled and became the source of records for creating the virtual twin match.

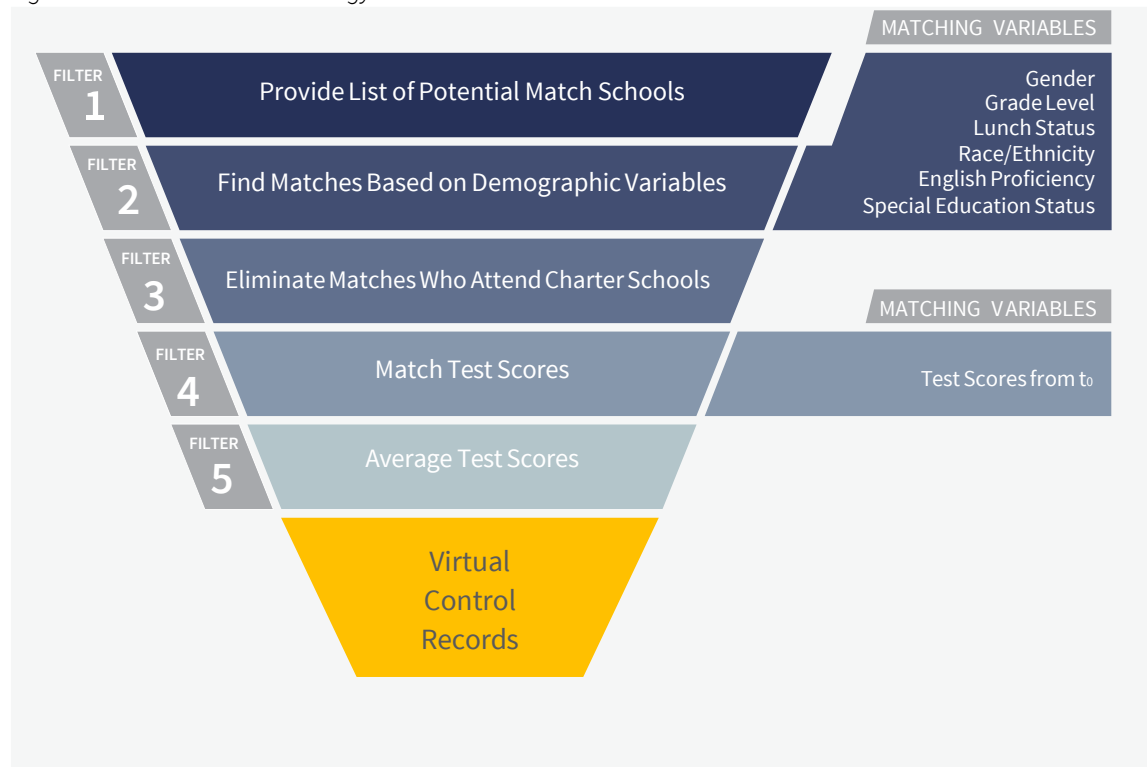
The VCR matching method then eliminated any of the students from the match pool whose demographic characteristics did not match exactly to the individual CRM student. As part of the match process, we also dropped from the match pool any students who were enrolled in a closed or alternative school in subsequent comparison years and who did not have test scores one year before the final year of operation for the CRM school.

Using the records of students at comparison schools in the final year of operation for the CRM school, CREDO randomly selected up to seven comparison students with identical values on the matching variables in Figure 1, including identical or very similar test scores in the final year of operation for the CRM school. Students with similar test scores were used only when there were not enough comparison students with exact score matches. The values for the selected comparison students were then averaged to create

² A detailed description of the affected student categories can be found [here](#).

values for the virtual twins. As all other observable characteristics were identical, the only observable characteristic that differed between the CRM student and the VCR was attendance in a CRM school. The test score achieved right before enrollment for the first time in a CRM school represented the impact on academic achievement of both the observable and the unobservable student characteristics up to the time of the match. Since we matched on observable characteristics and the test score in the final year of operation for the closed school, we concluded that any differences in test scores during enrollment in a CRM school were primarily attributable to CRM.

Figure 1: CREDO VCR Methodology



The control pool for CRM students in New Orleans, LA was comprised of all public school students in New Orleans that never enrolled in a CRM school (i.e., students from any RSD charter or traditional public school in New Orleans), referred to as “All NOLA”. This comparison enables a view into where the Closing and CRM school impacts fit into the overall New Orleans education market. The control pools for CRM in Memphis and Nashville comprised of students in public school at the bottom five percent of the school performance distributions in Memphis and Nashville, respectively. The proportions of case students for whom matches were created are displayed in Table 2 below. As we can see, our overall match rate is quite high, over 80 percent. This allows for a valid estimation of the impact of the CRM on student population.

Table 1: Student Match rates by Location

Location	Mathematics	Reading
New Orleans	78%	79%
Tennessee	89%	89%
Overall	82%	83%

Groups of Students under Study

Below we describe the different student categories affected by CRM:

- **Persisters:** These students attended both the Closing school and its CRM counterpart or newly enrolled students in the CRM who continued to be enrolled for more than one school year.
- **New Entrants:** The students who attended a CRM school but did not attend the corresponding Closing school. These students are New only in their first year of enrollment.
- **Opt-out:** Students who attended the Closing school, had the option of attending the new CRM school, and chose not to enroll there.
- **Flux:** These are students who attended the Closing school and for whom some special accommodation was made when they were displaced by the CRM school due to a grade configuration mismatch.
- **Ineligible:** Students who attended the Closing school but were not able to attend the CRM school due to a grade configuration mismatch and for whom no special accommodation was made.
- **Aged out:** The students who attended the Closing school in its highest grade level and were therefore going to enroll in a different school the following year regardless of the CRM project.

We estimated the impact of the CRM schools on two groups of students. The first is the intent-to-treat (ITT) group, i.e., students who were enrolled in a Closing school in its final year of operation before the CRM school opened. Our Intent-to-Treat (ITT) estimate consisted of the weighted average effects of Persisters, Opt-outs, Flux, and Ineligible students. Recall that some of the ITT group did not continue their enrollment in the CRM school (for a variety of reasons), yet this was the group that was the focus of the restart efforts. The second group of students comprises those who were actually enrolled in a CRM school, some of whom came from the Closing school and some who have different enrollment histories. The ITT group overlaps to a degree with the group of students that actually attended a CRM school because both groups include the Persister students (Persisters attended both the Closing school and the CRM school).

Demographic Comparison of students targeted by the intervention with matched students

Table 2 describes the demographic composition of all students included in our study, both the Intent to Treat (ITT) and the CRM groups. Students targeted in our study are predominantly black and eligible for free/reduced-price lunch. Matched students differ demographically from the CRM student population in important ways. The final study population (CRM students and their matched VCRs) had higher proportions of Black students and students in poverty than the entire CRM population. At the same time, the study population has fewer Asian or Pacific Islander, Special Education and English Language Learner (ELL) students than the CRM population. These patterns are evident in both New Orleans and Tennessee.

Table 2: Demographic Composition of All ITT and CRM Students in the Evaluation

	New Orleans		Tennessee		Overall	
	Target Students	Matched Students	Target Students	Matched Students	Target Students	Matched Students
Number of Tested Students	8254	6458	5520	4875	13774	11333
Match Rate		78%		88%		82%
Students in Poverty	84%	88%	92%	93%	87%	90%
Special Ed Students	11%	8%	6%	4%	9%	6%
ELL Students	4%	1%	1%	0%	3%	1%
White Students	0%	0%	2%	1%	1%	0%
Black Students	92%	98%	95%	98%	93%	98%
Hispanic Students	3%	1%	3%	1%	3%	1%
Native American Students	0%	0%	0%	0%	0%	0%
Asian or Pacific Islander Students	4%	0%	0%	0%	3%	0%
Multi-Racial Students	0%	0%	0%	0%	0%	0%

Table 3 below describes the demographic composition of students from Closing schools included in our study. Similar to Table 2, Students in the Closing schools are predominantly black and eligible for free/reduced-price lunch. The matched students attending the Closing schools are demographically similar to the targeted students, that is, predominantly black and eligible for free/reduced-price lunch. Again, student subgroups that are more representative of the underlying student population in the Closing schools, such as black students or students eligible for free/reduced-price lunch, have a higher the match rate compared to that of less representative students, such as special education students or English-language learners.

Table 3: Demographic Composition of ITT Students from Closing Schools in the Study

Student Group	New Orleans		Tennessee		Overall	
	Target Students	Matched Students	Target Students	Matched Students	Target Students	Matched Students
Number of Tested Students	1201	980	1455	1371	2656	2351
Match Rate	82%		94%		89%	
Students in Poverty	88%	90%	91%	92%	90%	91%
Special Ed Students	8%	7%	4%	3%	6%	5%
ELL Students	1%	*	1%	**	1%	*
Black Students	95%	99%	96%	98%	95%	98%
Hispanic Students	2%	*	2%	1%	2%	1%

Table 4 below describes the demographic composition of students in the CRM schools (i.e. Persisters and New Entrants) included in our study. Similar to Table 2, students in the CRM schools are predominantly black and eligible for free/reduced-price lunch. The matched students in New Orleans attending the CRM schools compared to the targeted students are demographically slightly more likely to be black and more likely to be eligible for free/reduced-price lunch. The matched students in Tennessee attending the CRM schools are demographically similar to the targeted students. Student subgroups that are more representative of the underlying student population in the CRM schools, such as black students or students eligible for free/reduced-price lunch, have a higher match rate compared to that of less representative students, such as special education students or English-language learners.

Table 4: Demographic Composition of CRM Students in the Study

Student Group	New Orleans		Tennessee		Overall	
	Target Students	Matched Students	Target Students	Matched Students	Target Students	Matched Students
Number of Tested Students	4984	3789	1483	1330	6467	5119
Match Rate	76%		90%		79%	

Students in Poverty	82%	87%	95%	92%	85%	89%
Special Ed Students	12%	9%	10%	6%	11%	8%
ELL Students	5%	1%	2%	1%	4%	1%
Black Students	89%	98%	97%	95%	90%	98%
Hispanic Students	4%	2%	2%	1%	3%	2%

Comparison of Starting Scores of Matched and Unmatched Students

Although the VCR method used in this study provided matches for 82 percent of the targeted students, it is important to examine how different unmatched students may be compared to those included in our analysis. Whether unmatched students are similar or not to matched students determines the extent to which our results can be generalized to the universe of students targeted by the intervention. The ability to extrapolate findings from a particular sample to the broader population is referred to as external validity. In the case of this analysis, CREDO's sample encompasses a large proportion of the entire population of targeted students. At the same time, we see that the test scores of matched students are significantly higher in statistical terms than those of unmatched students in both math and reading in which they were matched. This is because targeted students at the very low and high end of the test score distribution have less likely to find matches in the match pool. It is important to note that both matched and unmatched students are found to be dramatically below the state average in terms of starting endowments in both math and reading. The fact that our data represent 82 percent of the students targeted by the intervention makes us confident that our estimated are highly aligned with the actual population values, although we are uncertain to what extent our results apply to students without matches.

Figure 2: Comparison of Starting Math Scores of Matched and Unmatched Students across NOLA and TN
Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
No	2,268	-.6321484	.0272109	1.295881	-.6855093	-.5787874
Yes	11,333	-.5565053	.0084218	.8965525	-.5730134	-.5399972
combined	13,601	-.569119	.0083596	.974926	-.585505	-.552733
diff		-.0756431	.022418		-.1195855	-.0317006
diff = mean(No) - mean(Yes)					t = -3.3742	
Ho: diff = 0					degrees of freedom = 13599	
Ha: diff < 0		Ha: diff != 0		Ha: diff > 0		
Pr(T < t) = 0.0004		Pr(T > t) = 0.0007		Pr(T > t) = 0.9996		

$$\Delta A_{i,s,t} = \theta A_{i,t-1} + \beta X_{i,t} + \rho Y_t + \sigma S_s + \gamma C_{i,t} + \varepsilon_{i,s,t}$$

where the dependent variable is

$$\Delta A_{i,s,t} = A_{i,t} - A_{i,t-1}$$

And $A_{i,s,t}$ is the state-by-test z-score for a student i in State s in period t ; $A_{i,s,t-1}$ is the state-by-test z-score for student i in State s in year before t ; $X_{i,s,t}$ is a set of control variables for student characteristics; Y is a year fixed effect; S is a state fixed effect, C is an vector of indicator variables capturing all student categories affected by CRM, the most notable being the category of students attending a CRM school in t_1 ; and ε is the error term. The following student categories affected by CRM were included in vector C : Persisters, New Entrants, Opt-outs, Flux, and Ineligible. An additional category was included in vector $C_{i,t}$, that of students attending a closing school at time t_0 , the year prior to the opening of the corresponding CRM school. The students we examined in this Impact Evaluation fell into two aggregate groups. First, Closing school students were all the students who attended the Closing school in its final school year, i.e., Persisters, Opt-out, Flux, Ineligible, and Aged Out. These were the students that the CRM was designed to assist, even though some students did not attend a restart school. We refer to this group as Intent to Treat group and studied their subsequent academic experience. We have estimated a CRM effect for the Intent to Treat group by taking the weighted average of the CRM effects each of the included student categories experienced. The second group consisted of students who attended the restart schools – Persisters and New Entrants. This aggregated category, CRM attendees (Persisters and New Entrants), was the focus of the majority of the Impact Analysis. The CRM effect for CRM attendees was estimated by an indicator variable capturing CRM participation included in vector $C_{i,t}$. A Wald test was conducted between the effect of CRM attendees and that of students in the closing school at time t_0 . This test allowed us to evaluate CRM school in comparison to the schools they replaced and the results can be found in the Tables 5-7.

Table 5: Comparison of CRM and Closing Schools-Overall

	Math			Reading		
	Effect	SE	p-value	Effect	SE	p-value
CRM Schools	-0.02	0.04	0.64	0.01	0.03	0.86
Closing Schools	-0.15	0.05	0.00	-0.11	0.04	0.01
Difference p-value			0.03			0.01

Table 6: Comparison of CRM and Closing Schools-New Orleans

	Math			Reading		
	Effect	SE	p-value	Effect	SE	p-value
CRM Schools	-0.01	0.05	0.78	0.00	0.05	1.00

Closing Schools	-0.26	0.07	0.00	-0.17	0.07	0.02
Difference p-value			0.00			0.04

Table 7: Comparison of CRM and Closing Schools-Tennessee

	Math			Reading		
	Effect	SE	p-value	Effect	SE	p-value
CRM Schools	-0.04	0.04	0.31	-0.00	0.02	1.00
Closing Schools	-0.08	0.05	0.10	-0.07	0.03	0.04
Difference p-value			0.57			0.07

In addition to the baseline model above, we explored differential CRM effects across the following dimensions. We investigated differential CRM effects exploiting additional interactions beyond a simple binary to indicate CRM participation.

- We explored differential CRM effects across student characteristics using double interactions between the CRM participation influence and student characteristics. For example, to identify the impact of CRM on different racial groups, we estimated models that broke the CRM variable into “CRM_black,” “CRM_hispanic,” etc.
- To break down the impact of CRM by school level, the CRM variable was split into “CRM_high_school,” “CRM_middle_school,” etc.
- We also investigated the differential impact of attending a Fresh Start CRM school or a Full Turnaround CRM school on the growth of CRM students. To that end, we estimated models that split the CRM variable into “CRM_fresh,” and “CRM_full.”
- An important part of our evaluation of the CRM intervention was to identify student growth that is attributable to each CRM school. To do this, the CRM variable was split into “CRM_school1,” “CRM_school2”, etc.
- We were also interested in investigating to which extent CRM participation varied across year of CRM school operation. To identify the CRM effect on student growth for each year of operation we replaced the simple binary for CRM participation with a series of binary variables, one to capture CRM effect for each year of CRM school operation.
- We investigated whether CRM effect varied by the time students spent in a CRM school. To break down the impact of CRM by years of enrollment, the CRM variable was split into “one_year_in_CRM”, “two_years_in_CRM”, “three_years_in_CRM”, etc.

[Comparison of Starting Positions of Each Type of Turnaround](#)

Policy discussions benefit from an understanding of whether different turnaround approaches yield different improvements in test scores. To answer that question, we first compared the incoming students served by the two types of restarts. Examining potential differences in students’ educational preparation as they enter a CRM school is one important consideration in parsing CRM schools’ performance. We conducted a t-test on the difference between the average state standardized achievement scores in the year before CRM enrollment of students attending Fresh Start and Full Turnaround CRM schools. To ensure fair comparisons,

we also distinguished whether prior achievement was the result of attending the Closing school. Full school turnarounds in their first year of operation had some students carry over from the Closing schools (Persisters) and some students who enrolled from other settings (New Entrants). Because fresh start schools opened with only their entry grade, all their students are New Entrants. Implementation study findings suggested that these distinctions may be important: non-academic influences from the Closing school were reported to carry over into the restart environment. The results are shown in Tables 19-21 in the [Impact Report](#), which are also reported below in Tables 8-10.

Table 8: Comparison of Starting Position for Persisters and New Entrants-Overall

	Math			Reading		
	Fresh	Full	Significant Difference	Fresh	Full	Significant Difference
Persisters		-0.77			-0.69	
New Entrants	-0.44	-0.30	*	-0.70	-0.37	**
All attendees	-0.45	-0.63	**	-0.70	-0.60	

Note: †, *, ** indicate significance at the 10, 5, and 1 percent level.

Table 9: Comparison of Starting Position for Persisters and New Entrants-New Orleans

	Math			Reading		
	Fresh	Full	Significant Difference	Fresh	Full	Significant Difference
Persisters		-0.78			-0.65	
New Entrants	-0.08	-0.23	†	-0.51	-0.26	
All attendees	-0.13	-0.61	**	-0.51	-0.54	

Note: †, *, ** indicate significance at the 10, 5, and 1 percent level.

Table 10: Comparison of Starting Position for Persisters and New Entrants-Tennessee

	Math			Reading		
	Fresh	Full	Difference p-value	Fresh	Full	Difference p-value
Persisters		-0.69			-1.22	
New Entrants	-0.56	-0.94	**	-0.70	-1.24	*
All attendees	-0.56	-0.78	†	-0.70	-1.23	**

Note: †, *, ** indicate significance at the 10, 5, and 1 percent level.

Regression Output for Fresh Start Vs Full Turnaround Comparison

In Tables 8 and 9 we report the regression output for the impact analysis of CRM in two types of school turnaround; Fresh Start and Full Turnaround.

Table 11: The effect of the CRM by type of Turnaround in Math

Variable	Overall		New Orleans		Tennessee	
	Coefficient	SE	Coefficient	SE	Coefficient	SE

Starting score	-0.383	0.009	-0.401	0.015	-0.347	0.010
Fresh Start CRM	0.086	0.053	0.140	0.109	0.054	0.038
Full Turnaround CRM	-0.048	0.045	-0.038	0.057	-0.121	0.039
CRM Previous	-0.025	0.060	0.021	0.070	-0.243	0.059
CRM Ineligible	0.016	0.072	0.001	0.077		
CRM Opted Out	0.040	0.053	0.103	0.067	-0.105	0.048
CRM Flux	-0.103	0.089	0.041	0.095	-0.151	0.075
CRM Closing	-0.103	0.066	-0.0221	0.088	-0.075	0.047
Female	0.040	0.009	0.027	0.013	0.055	0.013
Black	-0.061	0.041	0.068	0.142	-0.063	0.041
Hispanic	0.136	0.068	0.274	0.171	0.086	0.060
Asian/Pacific Islander	0.480	0.078	0.668	0.159	-0.045	0.082
Native American	0.327	0.152			0.219	0.154
Multi-ethnic	-0.135	0.061	-0.002	0.161		
Is Special Ed	-0.260	0.026	-0.256	0.030	-0.249	0.044
Is English Learner	-0.239	0.054	-0.281	0.079	-0.128	0.104
Is in Poverty	-0.057	0.013	-0.053	0.018	-0.074	0.018
Repeated grade	0.267	0.040	0.324	0.053	0.204	0.061
Grade_03	-0.084	0.096	-0.076	0.100	-0.090	0.122
Grade_04	0.006	0.033	0.026	0.043	0.008	0.052
Grade_06	0.038	0.037	0.190	0.032	-0.124	0.045
Grade_07	0.057	0.028	0.128	0.033	-0.051	0.037
Grade_08	0.199	0.037	0.387	0.036	-0.015	0.032
Grade_09	0.163	0.047	0.156	0.061	0.180	0.047
Grade_10	-0.028	0.038	0.048	0.047	-0.393	0.164
Grade_11	0.133	0.081	0.173	0.089		
Grade_12	0.057	0.086	0.125	0.082		
Year_2010	-0.028	0.038	-0.197	0.097		
Year_2011	0.007	0.086	0.030	0.097	0.0670	0.046
Year_2012	-0.032	0.084	-0.060	0.095	0.065	0.055
Year_2013	0.077	0.084	0.030	0.094	0.170	0.048
Year_2014	-0.073	0.083	-0.092	0.092		
Growth_period_02	0.014	0.019	-0.009	0.023	0.060	0.032
Growth_period_03	0.066	0.029	-0.023	0.034	0.126	0.037
Growth_period_04	0.053	0.028	-0.001	0.031	0.089	0.044
Growth_period_05	-0.077	0.045	-0.121	0.018		
NOLA	0.154	0.021				
Constant	-0.179	0.095	-0.196	0.186	-0.140	0.056
Observations	22634		12884		9750	
R-squared	0.261		0.307		0.237	

Table 12: The effect of the CRM by type of Turnaround in Reading

Variable	Overall		New Orleans		Tennessee	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
Starting score	-0.309	0.010	-0.364	0.011	-0.248	0.009
Fresh Start CRM	0.863	0.038	0.208	0.053	0.0169	0.040
Full Turnaround CRM	-0.014	0.038	-0.020	0.051	-0.013	0.029
CRM Previous	-0.025	0.049	-0.022	0.062	0.024	0.044
CRM Ineligible	-0.066	0.064	-0.069	0.070		
CRM Opted Out	-0.023	0.045	0.029	0.060	-0.072	0.038
CRM Flux	-0.045	0.051	-0.036	0.098	-0.048	0.033
CRM Closing	-0.092	0.051	-0.151	0.083	-0.067	0.032
Female	0.066	0.008	0.079	0.011	0.047	0.010
Black	-0.141	0.036	0.159	0.252	-0.123	0.038
Hispanic	0.038	0.057	0.407	0.257	-0.031	0.057
Asian/Pacific Islander	-0.073	0.066	0.280	0.255	-0.224	0.148
Multi-ethnic	-0.149	0.243	0.177	0.350		
Is Special Ed	-0.220	0.024	-0.270	0.023	-0.157	0.049
Is English Learner	-0.064	0.049	-0.127	0.068	-0.069	0.072
Is in Poverty	-0.062	0.013	0.061	0.017	-0.066	0.022
Repeated grade	0.333	0.053	0.103	0.069	0.106	0.054
Grade_03	0.136	0.141	0.176	0.170	-0.212	0.189
Grade_04	0.049	0.031	0.078	0.043	-0.031	0.030
Grade_06	0.134	0.027	0.228	0.031	0.016	0.032
Grade_07	0.096	0.023	0.129	0.030	0.030	0.027
Grade_08	0.148	0.029	0.239	0.042	0.029	0.028
Grade_09	0.217	0.034	-0.210	0.243	0.128	0.035
Grade_10	0.075	0.034	0.070	0.047	0.122	0.048
Grade_11	-0.034	0.042	-0.005	0.052	0.015	0.151
Grade_12	0.321	0.199	0.305	0.205		
Year_2010	-0.120	0.085	-0.174	0.101		
Year_2011	-0.068	0.082	-0.029	0.094	-0.025	0.027
Year_2012	-0.042	0.081	-0.061	0.094	0.056	0.026
Year_2013	0.061	0.081	0.092	0.094	0.114	0.030
Year_2014	-0.022	0.080	-0.017	0.091		
Growth_period_02	0.008	0.015	-0.014	0.023	0.025	0.016
Growth_period_03	0.020	0.019	-0.019	0.027	0.036	0.021
Growth_period_04	0.008	0.021	-0.039	0.028	0.017	0.038
Growth_period_05	-0.109	0.048	-0.136	0.052		
NOLA	0.139	0.020				
Constant	-0.110	0.096	-0.330	0.266	-0.058	0.047
Observations	22178		12134		10044	
R-squared	0.214		0.261		0.172	

Testing the difference in Effect sizes of Fresh Start and Full Turnaround CRM schools

We used a Wald test to test the statistical significance of the difference between the effect of Fresh Start CRM schools and that of Full Turnaround CRM schools. The results of the Wald test can be found in Tables 13-15.

Table 13: Comparison of Fresh Start and Full Turnaround CRM Schools-Overall

	Math			Reading		
	Effect	SE	p-value	Effect	SE	p-value
Fresh Start Charter Schools	0.09	0.05	0.11	0.09	0.04	0.02
Full Turnaround Charter Schools	-0.05	0.04	0.29	-0.01	0.04	0.71
Difference in Effects			0.05			0.07
p-value						

Table 14: Comparison of Fresh Start and Full Turnaround CRM Schools-New Orleans

	Math			Reading		
	Effect	SE	p-value	Effect	SE	p-value
Fresh Start Charter Schools	0.14	0.11	0.20	0.21	0.05	0.00
Full Turnaround Charter Schools	-0.04	0.06	0.51	-0.02	0.05	0.70
Difference in Effects			0.14			0.00
p-value						

Table 15: Comparison of Fresh Start and Full Turnaround CRM Schools-Tennessee

	Math			Reading		
	Effect	SE	p-value	Effect	SE	p-value
Fresh Start Charter Schools	0.05	0.04	0.16	0.02	0.04	0.67
Full Turnaround Charter Schools	-0.12	0.04	0.00	-0.01	0.03	0.64
Difference in Effects			0.00			0.59
p-value						

Comparison of Starting Positions of Opt-outs and Persisters

Although the CRM intervention intended to treat students in the closing schools, some students elected not to enroll in a newly-formed CRM school. This type of students fall in the opt-out category. To understand the extent to which our analysis on the closing-school students who enrolled in the CRM school can be extended to those who opted out we compare the achievement of persisters to that of opt-outs in the year prior to the

CRM's first operation. A t-test allowed to infer the statistical significance of those differences. Persisters and opt-outs across New Orleans and Tennessee were not found to have statistically different average achievement in either subject. In New Orleans, alone, opt-outs were found have statistically higher average baseline achievement than persisters in New Orleans.

Table 16: Baseline achievement of opt-outs and persisters at the first year of operation.

	Math			Reading		
	Opt-outs	Persisters	Difference p-value	Opt-outs	Persisters	Difference p-value
Overall	-0.80	-0.77	0.64	-0.77	-0.69	0.14
New Orleans	-0.67	-0.78	0.11	-0.50	-0.65	0.04
Tennessee	-0.91	-0.69	0.14	-1.01	-1.22	0.23

Mixed Methods Model

Our CRM evaluation involved the careful examination of both student- and school-level observations. The observations from different layers of school operation allowed us to exploit variation in qualitative observations to interpret academic performance across schools. To that direction, we combined school-level measures of academic growth with quantified observations of operational drivers at the school level in order to investigate to which effect these factors are related with student outcomes. A detailed description of the factors investigated can be found in the [Impact Report](#). The operational factors studied are nested within three broad categories as shown in Table 17.

Table 17: Operational Factors

CMO Endowments	Site & Location	Human Capital
Charter School Growth Fund grants	Physical Location Move	Principal Turnover
Local CMO Strategic Plan	Facility Index	Principal Performance Contract
Academic Committee on Board	Co-Location with Closing School	Principal CMO Review
Board Expertise	Grade Match with Flagship School	
Board Activity	Parent/Community Engagement named as key challenge	
Board Engagement		
Board Training		
Governance Index		
Lagged Average Achievement of Flagship School in Reading		
Lagged Average Achievement of Flagship School in Mathematics		

In addition to the operational drivers, we have also examined whether the variation in academic performance across school could be explained by the variation in implementation quality, proxied by PMO, an index described earlier in the Implementation Evaluation. A t-test between the average value of each

operational driver and the average PMO across turnaround approach allowed us to investigate the prevalence/strength of each operational driver, as well as the level of implementation quality in Fresh Start and Full Turnaround CRM schools.

The association between the PMO and the academic performance of each school was investigated with the following model.

$$\overline{\Delta A}_{st} = \alpha PMO_s + \beta \overline{X}_{st} + \eta_{st}$$

Where $\overline{\Delta A}_{st}$ is the average academic growth at school s in year t . \overline{X}_{st} captures school-level averages of student characteristics (race/ethnicities, free or reduced-price lunch eligibility, special education status, English language proficiency, grade retention) at school s and time t . The parameter of interest was α . We estimated the following model to investigate to which extent the variation in academic performance across schools can be explained by the variation in operational drivers or the PMO.

$$\overline{\Delta A}_{st} = \alpha F_s + \beta \overline{X}_{st} + \eta_{st}$$

To reduce the number of parameters for estimation, we fixed the coefficients of \overline{X}_{st} at the level of the coefficient of $X_{i,t}$ estimated from the Impact model above. F_s is a variable capturing each operational factor investigated in school s . The factors whose influence on student growth was investigated varied either across schools or across schools and year. In the latter case, the variable F_s was replaced with vector F_{st} capturing iteratively each operational factor varying across schools s and time t .