The Determinants and Plans of North Carolina Restart Schools

Lilly Winfree Ariella Hirsch

University of North Carolina, Chapel Hill

January 2019

Abstract

This paper studies the student characteristics of North Carolina Restart Schools and the expected efficacy of the programs they will be enacting. We compile data on demographic composition and end of grade exam performance and analyze elementary schools' Restart applications, which detail the changes schools wish to make. We find that Restart schools differ significantly from schools that do not apply for this initiative and that both racial composition and the percent of students performing at the cusp of grade level proficiency are correlated with Restart implementation. Many schools also propose to extend the amount of time teachers and students spend in the classroom, and similar past proposals have had varying levels of success when used in other schools. Our findings help describe the larger trends within a program that currently is not well understood due to its relatively short history.

This research was supported by the Matthew Guest Family Fund and made possible by the UNC Economics Department. We would like to thank Dr. Jane Cooley Fruehwirth and Dr. Catherine Zimmer for their invaluable help and support.

I. Introduction

Since North Carolina received nearly $400 million in federal grant money through the Race to the Top initiative, 103 chronically underperforming public elementary, middle, and high schools have been approved to become NC Restart schools. These schools are given greater autonomy in an attempt to improve their students' outcomes. NC Restart schools were first in the 2016-2017 school year, and so far, no school's application has been rejected. Given its sudden uptick in prevalence and the universal acceptance of the applicants we wish to understand the schools that choose to apply to this program and what distinguishes them from the other underperforming schools in the state. The recent implementation of this program makes an analysis of its effectiveness impossible, however, using the changes schools have stated they will make within their application, we survey the empirical literature to assess the potential efficacy of these changes and how we expect them to affect students attending NC Restart schools.

II. Institutional Background

The recent uptick in North Carolina Restart Schools can be understood by looking at the progression of federal intervention in underperforming schools. In 2002, George W. Bush signed the No Child Left Behind Act. This act mandated schools receiving federal funds to adopt ambitious academic achievement standards and to measure the extent to which they are accomplished with yearly exams for all students between third and eighth grade. Measures of achievement were also required to be further dissected to look at student performance in different subgroups including racial background, economic status, disability status and others. These disaggregated scores have allowed for the further identification and investigation of

underperforming schools (Powers, 2013). Legislation passed later by Barack Obama further outlined programs to be implemented to improve underperforming schools.

In 2009 the American Reinvestment and Recovery Act was passed in an attempt to reinvigorate the United States' economy. This bill included over three billion dollars in grant money for schools through the Race to the Top (RTT) program (McGuinn, 2011, Powers, et.al. 2013). RTT fostered state competition for its grant funds. States applying for these funds submitted plans that focused on intervening in low-performing schools. It was particularly emphasized that schools have a history of intervention and a thorough plan for further improving their schools, as well as have a focus on improving teacher, principal, and administrator effectiveness (Powers, 2013).

RTT proposed four primary models for intervention- Turnaround, Transformation, Restart, and Closure. The Turnaround Model dictates the principal be replaced and no more than 50% of staff be rehired. It also adds imbedded professional development and provides financial and professional incentives for remaining staff. It adds learning time and grants flexibility to the school's leadership (McGuinn, National, 2010). A more scaled back iteration of the Turnaround Model is the Transformation Model. Transformation schools still replace their principal and have greater flexibility with hiring and - more significantly – firing, however they are not required to remove at least half of their staff. Transformation also allows for increased professional development, learning time, and incentives for staff as well as operating flexibility (National, 2010). Both Turnaround and Transformation schools make significant changes to structure of the school and facilitate the replacement, retention, and training of teachers and administrators while maintaining the school's autonomy. Restart and school Closures take a much more drastic approach to improving underperforming schools.

The federal Restart Model involves a school closing and reopening under the control of a charter school operator or charter management organization selected by the LEA (local education agency). All students previously attending the school are welcome to re-enroll. The last option, Closure, simply calls for the school to close and for former students to be enrolled in nearby higher achieving schools (National, 2010, Powers, 2013). With these two options, the previous structure of the school is completely abandoned. Schools applying to RTT funds were tasked with assessing these models and creating implementation plans for them.

In August of 2010, North Carolina was awarded nearly $400 million as a winner of Phase II of the Race to the Top initiative (North Carolina, 2010, Powers, 2013). Since then, 103 schools in the state have applied to become NC Restart Schools, which actually follow a model more similar to the federal Transformation model. Schools that have been designated as underperforming for at least two years may apply to the NC Restart program. To date there have been no applications rejected. NC Restart schools are given new budgetary and hiring flexibility in order to extend teaching and planning time, adjust class sizes, provide financial incentives for teachers and administrators, and several other plans outlined in the applications (North Carolina, 2010).

III. Data Collection

We retrieved our data from the online databases provided by the North Carolina Department of Education. We collected information on the composition of each school, including race and sex make up, crime and violence rates, and numbers of free and reduced lunch recipients. We then included the performance of schools on end of grade (EOG) exams, which are given to all students grade 3-8, and scored on a scale of one through 5, with anything above a two considered grade level proficient. EOG scores for each school are given, as well as disaggregated test scores

that include the performance for separate groups of students. We refined this data to only include elementary and middle schools, the ones using EOG testing. We then included percentages of students at each level of EOG performance, and the percent of students in each racial category, economically disadvantaged status, and sex at each level of EOG performance. With this rich data, we were able to find not only what schools look like, but how different types of students at those schools perform. The most recent years of data available was for the 2016-2017 school year, and we additionally included several years prior up to the 2013-2014 school year.

All applications for Restart schools are available through the online meeting minutes provided for the North Carolina Board of Education's meetings. After reading each application, we created binary variables to indicate which programs schools were planning on implementing and were then able to assess the prevalence of these programs.

IV. Results

We estimate that 92 NC schools have applied for Restart status as of Summer 2018; this number indicates that approximately five percent NC schools will transition to the Restart model within the next couple of years.

Table I summarizes a number of T-tests that seek to identify significant differences between Restart and non-Restart schools. These tests show that Restart schools have significantly more racial minority students, specifically Black and Hispanic students, than non-Restart schools. We find that Restart schools differ significantly from non-Restart schools in student achievement on standardized tests. This finding holds true when looking at racially aggregated student achievement data and when looking at data that show achievement by race for low performing Black and Hispanic students. The only test indicating no statistical

Table I: Differences Between Restart and Non-Restart Schools

T-Test	Group	Observations	Mean	P-value
Percent Black by Restart Status	0	7006	0.254	
	1	365	0.471	0.000
Percent White by Restart Status	0	7006	0.508	
	1	365	0.232	0.000
Violent Acts by Restart Status	0	7030	6.690	
	1	366	6.820	0.939
Level 1 by Restart Status	0	7124	23.107	
	1	368	35.720	0.000
Level 2 by Restart Status	0	7124	21.062	
	1	368	25.437	0.000
Level 3 by Restart Status	0	7124	10.139	
	1	368	9.943	0.157
Level 4 by Restart Status	0	7124	32.727	
	1	368	23.282	0.000
Level 5 by Restart Status	0	7124	12.894	
	1	368	5.387	0.000
GLP by Restart Status	0	5345	56.400	
	1	276	38.664	0.000
Black % Level 1 by Restart Status	0	2493	23.674	
	1	37	29.689	0.002
Black % Level 2 by Restart Status	0	2418	23.537	
	1	37	26.003	0.027
Hisp % Level 1 by Restart Status	0	4195	21.316	
	1	104	27.811	0.000
Hisp % Level 2 by Restart Status	0	4038	22.843	
	1	103	24.851	0.004

Group 0 = non-Restart schools; Group 1 = Restart schools

difference shows that Restart and non-Restart schools experience comparable rates of crime and violence.

We now look at average marginal effects from probit models to understand the effects of race and student achievement levels on schools becoming Restart schools. We selected the probit model as a small percentage of NC schools have applied for the Restart program and linear models do not predict well at the tails of data.

Table II: Marginal Effects of Student Characteristics on School Restart Status

	Delta Method Marginal Effect	Standard Error	z	P>|z|	[95% Confidence Interval]
Regression 1 *					
percent black	0.161	0.010	15.430	0.000	0.141 , 0.182
percent hispanic	0.166	0.016	10.670	0.000	0.136 , 0.197
percent other	-0.375	0.029	-12.960	0.000	-0.431 , -0.318
Regression 2 *					
Percent Level 1	0.236	0.018	13.000	0.000	0.201 , 0.272
Percent Level 2	0.716	0.061	11.820	0.000	0.597 , 0.835
Regression 3*					
Percent Level 1	0.170	0.020	8.470	0.000	0.131 , 0.209
Percent Level 2	0.489	0.062	7.830	0.000	0.367 , 0.612
Percent Black	0.077	0.011	6.760	0.000	0.055 , 0.099
Percent Hispanic	0.111	0.015	7.460	0.000	0.082 , 0.140
Percent Other	-0.230	0.037	-6.210	0.000	-0.302 , -0.157
Regression 4*					
Percent Hispanic Level 1	0.009	0.035	0.260	0.794	-0.059 , 0.077
Percent Hispanic Level 2	0.039	0.043	0.930	0.353	-0.044 , 0.123
Percent Level 1	0.137	0.042	3.260	0.001	0.055 , 0.219
Percent Level 2	0.184	0.080	2.310	0.021	0.028 , 0.340
Percent Black	0.053	0.015	3.430	0.001	0.023 , 0.083
Percent Hispanic	0.054	0.017	3.170	0.001	0.021 , 0.087
Percent Other	-0.271	0.098	-2.760	0.006	-0.463 , -0.078
Regression 5*					
Percent Black Level 1	0.008	0.049	0.170	0.864	-0.087 , 0.104
Percent Black Level 2	0.004	0.067	0.060	0.954	-0.127 , 0.134
Percent Level 1	0.058	0.061	0.950	0.341	-0.061 , 0.177
Percent Level 2	0.256	0.090	2.830	0.005	0.079 , 0.433
Percent Black	0.048	0.016	2.940	0.003	0.016 , 0.080
Percent Hispanic	0.049	0.020	2.510	0.012	0.011 , 0.087
Percent Other	-0.194	0.116	-1.670	0.095	-0.421 , 0.034

*All regressions were run using a probit model.

Seen in Table II, Regression 1 looks at the likelihood of schools transitioning to the Restart model conditional on the percentage of Black students, the percentage of Hispanic

students and the percentage of "other" students" (including multi-racial, American Indian, Pacific Islander). This model seeks to understand whether the racial composition of a school is correlated with which schools have applied for Restart status. We initially see that race is a significant predictor of which schools apply to the Restart Program. We then look at the marginal effects of race on predicting Restart status and see that the percent of Black students and the percent of Hispanic students attending a school have similar effects on the likelihood a school becomes a Restart School. We estimate that a 10 percent increase in Black students at a school increases the likelihood that school becomes Restart by 1.6 percent, while a 10 percent increase in Hispanic students at a school increases the likelihood that school becomes Restart by 1.7 percent.

Seen in Table II, Regression 2 looks at the likelihood of schools transitioning to the Restart model conditional on the percentage of students achieving at Levels 1 and 2 on standardized tests. In the above regression we test whether a student body performing at Level 1 or Level 2 on standardized tests is more highly correlated with becoming a Restart School. We find that both performance at Level 1 and Level 2 are significant predictors of application to the Restart program. We then look at the marginal effects and find that performance at Level 2 is more highly correlated with Restart status. Specifically, we see that a 10 percent increase of students achieving Level 2 on standardized tests is correlated with a 7.2 percent increase in the likelihood a school becomes a Restart School. Likewise, a 10 percent increase of students achieving Level 1 is correlated with only a 2.4 percent increase in likelihood of Restart status.

Seen in Table II, Regression 3 looks at the likelihood schools transition to the Restart model given the racial composition and student testing achievement at Levels 1 and 2. In the above model we test whether, conditional on race, achievement at Level 1 or Level 2 on

standardized tests is more strongly correlated with Restart status. We find that conditional on race, both Level 1 and Level 2 achievement remain significantly correlated with Restart status. The marginal effects of this model complement the results from our previous model. We find that a 10 percent increase in Level 2 achievement is correlated with a 4.9 percent increase in the likelihood of becoming a Restart school, and a 10 percent increase in Level 1 achievement in correlated with a 1.7 percent increase in the likelihood of becoming of Restart school.

Seen in Table II, in Regression 4 we look at whether the performance of Hispanic students is a predictor of schools transitioning to the Restart model above and beyond aggregate performance of students. We include variables for Hispanic student achievement at Levels 1 and 2, overall student achievement at Levels 1 and 2, and racial composition. We find that the percentage of Hispanic students performing at Level 1 and Level 2 on standardized tests is not significantly correlated with schools becoming Restart schools. We continue to see a significant relationship between student achievement at Level 1 and Level 2 and schools transitioning to the Restart model. This model shows that percent of Black and Hispanic students is still significantly correlated with Restart status, though the performance of Hispanic students has no additional explanatory power regarding Restart status over the performance of the student body.

Seen in Table II, in Regression 5 we seek to understand whether achievement of Black students at Level 1 and Level 2 is a predictor of schools becoming Restart schools above and beyond aggregate student performance. We include variables for Black student achievement at Levels 1 and 2, overall student achievement at Levels 1 and 2, and racial composition. We find that the percentage of Black students achieving and Levels 1 and 2 is not significantly correlated with transition to the Restart model. This model shows that percent of Black and Hispanic students is still significantly correlated with Restart status, though the performance of Black

students has no additional explanatory power regarding Restart status over the performance of the student body.

Table III: Programs Planned for Implementation in Restart Schools

Variable	Observations	Mean	Standard Deviation	Min	Max
Certification Flexibility	92	0.533	0.502	0	1
Teacher Bonuses	92	0.402	0.493	0	1
Curriculum Flexibility	92	0.348	0.479	0	1
Extra school time for students*	79	0.759	0.430	0	1
Extra school time for staff*	70	0.929	0.259	0	1
Teacher coaching	92	0.641	0.482	0	1
Professional Development	92	0.946	0.228	0	1
Time for teaching teams	92	0.717	0.453	0	1
Pay for performance initiatives	92	0.391	0.491	0	1
Class size flexibility	92	0.391	0.491	0	1
Extneded school calendar*	89	0.933	0.252	0	1
Teacher recruitment flexibility	92	0.739	0.442	0	1
Teacher evaluation flexibility	92	0.239	0.429	0	1

*Lower total observations for select programs are a result of our difficulties interpreting some Restart schools applications

Table III summarizes the rough plans of the almost 100 schools that have applied for Restart status using data from their applications to the NC State Board of Education. Notably, 92.9 percent of Restart school applications mentioned increasing school time for staff; 94.6 percent of applications mentioned changing professional development requirements; and 93.3 percent of applications mentioned extending the school day. Furthermore, 76 percent of applications noted increasing school time for students; 71.2 percent of applications noted increasing teacher times with teaching teams; and 74 percent of applications mentioned using teacher recruitment flexibility under Restart status. Schools also indicated plans to use teacher

certification flexibility during their hiring process, pay for performance models, class size flexibility, and flexibility in administering teacher evaluations. We now review data from studies looking at the impact of individual interventions on student outcomes.

V. Evidence of Efficacy

North Carolina Restart schools are planning to implement a various combinations of programs in an attempt to improve student outcomes in low performing schools. Each of the different programs Restart schools are planning to implement have been tested in other environments, and as such, we are able to use results from other programs in an attempt to predict how these programs will impact Restart schools. A key difference between the Restart initiative and past programs is the number of programs Restart schools will be implementing during a short time period. Therefore, it will be important to investigate how different programs interact with each other and impact student outcomes. Further, it will be important to collect data on which programs Restart schools actually implement, as the programs each school indicated on their applications are simply program possibilities that the schools were considering.

In our research on Restart schools, we use a list of Restart schools in NC from Kelly Hinchcliffe, a reporter for WRAL (K. Hinchcliffe, personal communication, May 10, 2018). We were able to review the Restart applications of 92 schools in North Carolina. Common programs listed for implementation on these applications include the following: flexibility with teacher certification or teacher licensing, funding flexibility to allow for bonuses and stipends including pay-for-performance, curriculum flexibility, extra school time for teachers and/or students, funding flexibility to allow for additional mentor or coaching programs, funding flexibility to allow for professional development, more time for collaboration between teacher or teacher team

meetings, flexibility with class size or class ratio laws, programs to increase teacher recruitment and/or retention, and flexibility with teacher evaluation.

Class Size

Of the applications we reviewed, 39.13 percent mentioned use of flexibility with class size or ratio laws. Chingos (2012) studied a class size reduction law in Florida that required each class to be below a certain size limit with the goal of increasing student outcomes, as measured by standardized test scores. The Florida law provided funding to each school district regardless of whether the district needed to reduce class sizes in order to comply with the policy. This disbursement of funding meant that some districts received funding they needed to use in order to reduce class sizes, while other districts received funding that could be used at their discretion. Chingos found that Florida's class size reduction law had "little, if any, effect on student achievement" with regards to standardized test scores (Chingos 2012). The Florida policy did, however, appear to have an effect on non-cognitive outcomes such as "student absenteeism in elementary schools and incidents of crime and violence in middle schools" (Chingos 2012). The cost of the Florida program was significant, 20 billion annually for the initial eight years, as are the costs of class size reduction programs generally, and given the absence of positive impact on test scores, the money allocated to class size reduction may be used more effectively in other programs to improve student outcomes (Chingos 2012).

Cho, Glewwe, and Whitler (2012) study another class size reduction program, this one in Minnesota elementary schools. In contrast to Chingos (2012)'s findings, they find that a reduction in class size does improve performance on standardized tests. However, they find that "a decrease of ten students would increase test scores by only 0.04-0.05 standard deviations (of the distribution of test scores)" (Cho Glewwe &Whitler 2012). This result demonstrates that

even a substantial decrease in the number of students in a classroom has a very small impact on student performance on standardized testing. Therefore, despite finding a positive impact of class size reduction on student outcomes, Cho Glewwe and Whitler's findings still support the idea that there are more efficient and effective ways to improve student outcomes.

Given that a number of NC Restart schools have asked for flexibility with class size laws, which are typically written to limit the number of students in a classroom, we posit that Restart schools asking for this flexibility are looking to increase the number of students in a classroom. Increasing the number of students per classroom would allow schools to decrease the number of classes and divert resources used to support these classrooms elsewhere. As Chingos (2012) finds no positive impact on student achievement as a result of the Florida policy, and given that they also review other research on the impact of class size reduction policies and find considerably variable results as to the impact of class size reduction policies on achievement, increasing class sizes in order to reallocate limited resources may be a logical way for Restart schools to find funding within their budget for other programs (Chingos 2012). The diversion of funds from keeping class sizes within legal limits in Restart schools may allow for programs that improve student outcomes more holistically and efficiently given the budget limitations schools face.

Florida's class-size reduction policy raises other questions about fundamental assumptions behind the Restart schools program. The Restart program provides schools with increased financial discretion while assuming that schools will allocate funds more efficiently to maximize their utility, thereby maximizing student outcomes. The design and outcomes of the Florida program, however, call into question this assumption. Given that the Florida policy resulted in some districts receiving money that needed to be spent on reducing class sizes while

other districts received money to be spent at their discretion, the student outcomes between students in these two groups of districts can be compared to students in non-Restart and Restart schools (with the students in non-Restart schools compared to students in districts that had to reduce class sizes). Chingos found that "the test scores of districts required to reduce class size and districts provided similar funding but not required to reduce class size were not significantly different" (Chingos 2012). This result implies that districts that had more flexible spending failed to use this flexibility to improve student outcomes. In the case of Florida's class-reduction policy, there was no significant difference in student outcomes measured by standardized test scores in schools with and without spending flexibility. These results question a central tenant behind Restart schools – that greater discretion in spending will lead to efficient resource allocation and improved student outcomes.

Teacher retention and recruitment, bonuses, pay for performance

The Restart school applications we reviewed commonly mentioned difficulties in teacher retention and recruitment and cited high level of teacher turnover. 73.91% of the applications mentioned the goal of improving teacher retention or recruitment and 40.23% of the applications mentioned using funding flexibility to allow for teacher bonuses and stipends, often for the purpose of attracting and retaining teachers. 39.13 percent of the applications specifically mentioned pay for performance policies.

Guarino et al. (2011) look at how school demographics, policies, teacher characteristics, and their interactions affect "mobility behaviors of public school teachers" in North Carolina (Guarino, Brown & Wyse 2011). Their study takes place under a pay-for-performance program in North Carolina that rewarded schools that met or exceeded their growth targets. They find that teacher movements between schools and school districts largely results in schools that serve

disadvantaged students losing teachers that have "more desirable observable characteristics" (Guarino et al. 2011). Furthermore, they observe that "wealthy schools with low minority populations have an advantage in all aspects of recruitment and retention of teachers" (Guarino et al. 2011). Additionally, they find that pay-for-performance policies such as the one in North Carolina "appear to exacerbate the inequities in the distribution of teacher qualifications" (Guarino et al. 2011). These policies can incentivize teachers with mobility, often the better teachers, to leave underperforming schools that do not receive awards to go to better schools that have received awards. Simultaneously, however, there is evidence that pay-for-performance can decrease "mobility at schools with large poor and minority populations that receive the bonus awards" (Guarino et al. 2011). As such, pay-for-performance policies designed like North Carolina's that disburse awards on a school level can help retain teachers at low performing schools that receive the awards, but they also harm the lowest performing schools that do not receive the award. In their conclusion, Guarino et al. (2011) suggest the possibility of modifying pay-for-performance policies to award individual teachers based on their class's performance as an alternative to school level awards.

Jackson et al. (2014) also discuss pay-for-performance policies and bring up a number or concerns about this type of policy. They discuss the wide variety of inputs that impact student performance in addition to the teacher effect. This situation can make pay-for-performance policies seemingly unfair or discouraging for teachers that feel their students will not achieve the standards necessary for the award. In addition, a problem arises when pay-for-performance policies are implemented without the necessary support to help teacher become more effective. In other words, "merit pay may not be effective at improving teacher performance if individual teachers do not know what to do to improve their teaching performance" (Jackson, Rockoff &

Staiger 2014). In this case, pay-for-performance policies may simply be awarding teachers that are already effective rather than truly assisting teachers to increase their effectiveness, and thereby, student achievement. Jackson et al. (2014) review a number of pay-for-performance programs in other countries including Israel, England, and India. Each of the international programs they review seem to have more positive effects than the domestic programs they review. Further, each of these international programs differs significantly from the domestic programs in that the international programs each gave awards on a teacher level basis rather than a school level or average basis (Jackson et al. 2014).

These findings from Guarino et al. (2011) and Jackson et al. (2014) provide important insight into problems facing Restart schools, especially as many Restart schools serve majority disadvantaged populations and would like to focus in part on attracting and retaining effective teachers. It is plausible that the Restart schools that included a pay-for-performance program on their applications would be comparable to what Guarino et al (2011). suggest in their conclusion or to the international programs reviewed by Jackson et al. (2014). These schools would likely reward specific teachers at the school whose students meet a designated benchmark of performance. This teacher level pay-for-performance policy may succeed in retaining effective teachers at low performing schools by rewarding them with a bonus while also disincentivizing teachers from leaving a school that does not receive school level bonuses by providing teacher level bonuses. Furthermore, many Restart schools that are considering implementation of a pay-for-performance program are also looking to include professional development programming in their Restart model. The inclusion of professional development in addition to pay-for-performance could assist teachers in becoming more effective, avoiding the unintended consequences of pay-for-performance as discussed by Jackson et al (2014).

Teacher certification and teacher characteristics

Kane et al. (2008) review the impact of different certification statuses on student achievement by comparing traditionally certified teachers, alternatively certified teachers, uncertified teachers, and Teach for America teachers. They find that there is much more variation in teacher effectiveness between teachers with the same type of certification than there is variation between the certification groups when they add controls for level of teacher experience. Significantly, Kane et al. (2008) find "no difference" between certified and uncertified teachers when looking at their "impacts on math achievement". Students of certified teachers do slightly outperform students of non-certified teachers. Given their findings, Kane et al. (2008) propose that schools may better be able to improve their teacher quality by "selectively retaining only the most effective teachers" rather than basing hiring and tenure decisions mostly off of certification status and observable characteristics.

Kane et al. (2008) also look at differences in teacher certification status based on school demographics. They find that low performing schools have a harder time attracting teachers in general. This leads to a greater distribution of uncertified teachers in underperforming schools, and often, such schools also serve concentrated populations of minority or economically disadvantaged students (Kane et al. 2008).

Kane et al. (2008)'s research contains important information for Restart schools considering hiring uncertified teachers. Their research indicates that the effectiveness of uncertified and certified teachers differs only slightly in terms of teacher impact on student outcomes. Therefore, schools that have difficulty recruiting certified teachers may benefit from hiring uncertified teachers to teach classes in subject areas in which they are qualified or

knowledgeable. Furthermore, Kane et al. (2008)'s suggestion to school administrators to evaluate and retain teachers based on their performance in the classroom can assist Restart schools in assuring their students are receiving quality instruction. Flexibility to hire non-certified teachers can help schools have the staffing flexibility necessary to dismiss teachers that have proven ineffective and hire, possibly non-certified teachers, who may prove to be more effective in the classroom.

Extended school time

Of Restart school applications, 93 percent propose increasing their school time. The effect of extended day and extended year schedules has been discussed for decades and the effect that it has on student achievement is still unclear. Patall, et al. (2010) reviewed the 15 studies published between 1985 and 2009 evaluating extended day and extended time programs. Meyer and Klaverlen (2013) provide a much-needed randomized experiment and updated empirical results.

Patall, et al. (2010) surveyed 15 studies from 1985-2009 regarding the effectiveness of extended time and extended day programs. Seven of the studies looked solely at extended year programs while only two were solely interested in the effects of extended days. The results of extended school times were typically modest gains in achievement, although there were cases, such as Bishop, Worner, and Weber (1988), that found decreases in student achievement depending on the subject taught in this extra time. However, increases in school time were tied with a larger gain in achievement among minority students and students of a low socioeconomic status. Additionally, programs were more successful the more time they added. School days that were extended by 5 minutes showed smaller gains than those extended by 2 hours and school years increased by several weeks showed more improvement than those increased by only a few

days (Patall, et al., 2010). These last two findings can help us understand why Pattal et al. (2009) and Meyer and Klaverlen (2013) have such differing results.

Meyer and Klaverlen (2013) conducted a randomized study surrounding the rollout of an extended day program for elementary schoolers in the Netherlands. Seven elementary schools were granted an extended day program that added five hours of addition, interactive learning to school time per week for eleven weeks. Students' parents were randomly assigned vouchers that allowed their child to participate in the program. All students were given an exam both before and after the program to measure their growth in math, reading comprehension and spelling. Meyer and Klaverlen (2013) found no significant increase between the difference in pre and post-test scores between those students in the control and treatment group.

The insights from Pattal et al. (2010) may explain this result. An eleven-week program meeting three times a week may have been too brief for any effects of an extended day program to be seen. Additionally, the population of the town all seven of the elementary schools was only 8% ethnic minorities, and given that increased school time has larger effects on minority students, it is possible that the students who would benefit most from this program were not included in large numbers (Meyer and Klaverlen, 2013, Pattal et al., 2010).

These studies indicate that North Carolina Restart schools may be successful in improving achievement, as a majority of their students are from majority disadvantaged backgrounds and are therefore the students that respond best to additional school time. However, another important factor is the amount of time added to the school day or year. This is a detail not specified on any of the Restart applications. A substantial increase in learning time can contribute to the improvement of these underperforming schools and make this program a valuable feature of Restart schools.

VI. Conclusion

NC public schools that have applied to implement the Restart model differ significantly from schools that are not implementing this model. Schools implementing the Restart model have higher percentages of Hispanic and Black students and higher percentages of Hispanic and Black students performing at Levels 1 and 2 on standardized tests. Testing performance data on students of all racial backgrounds also shows that Restart schools have higher percentage of students performing at Levels 1 and 2 and a lower percentage of students performing at Levels 4 and 5, as expected. Student performance at Level 2 indicates a greater likelihood of schools transitioning to the Restart model than performance at Level 1. Therefore, we conclude that the program is targeting students on the verge of grade level proficiency (Level 3 and above) rather than the neediest students. Our probit regressions demonstrate that the racial composition of a school is correlated with the likelihood that a school becomes a Restart school, above and beyond performance. Specifically, we see that schools with higher Hispanic and Black student populations are more likely to become Restart schools, even after controlling for performance levels. We also find that student achievement at Level 2 on standardized testing is more highly correlated with Restart status than student achievement at Level 1; however, both performance at Levels 1 and 2 are significantly correlated with schools applying for Restart. These finding continue to hold true when looking at student achievement conditional on race. We analyzed the applications to the Restart program from 368 NC public schools and found that schools are interested in implementing a number of changes to improve student outcomes. Schools were frequently interested in adding extra school time for students or staff, requesting flexibility to help recruit and retain teachers, and adjusting professional development requirements. Based on existing literature, we expect these policies to be mixed in their success. Future research will

need to investigate whether implementation of the Restart model has significant effects on student outcomes and document which school changes have significant impacts on student achievement.

References

Bishop, J., Worner, W., & Weber, L. (1988). Extending the school day: An evaluation study of a seven-period class schedule. Studies in Educational Evaluation, 14(3), 361-380.

Chingos, M. M. (2012). The impact of a universal class-size reduction policy: Evidence from Florida's statewide mandate. Economics of Education Review, 31(5), 543–562. https://doi.org/https://doi.org/10.1016/j.econedurev.2012.03.002

Cho, H., Glewwe, P., & Whitler, M. (2012). Do reductions in class size raise students' test scores? Evidence from population variation in Minnesota's elementary schools. Economics of Education Review, 31(3), 77–95. https://doi.org/https://doi.org/10.1016/j.econedurev.2012.01.004

Guarino, C. M., Brown, A. B., & Wyse, A. E. (2011). Can districts keep good teachers in the schools that need them most? Economics of Education Review, 30(5), 962–979. https://doi.org/https://doi.org/10.1016/j.econedurev.2011.04.001

Jackson, C. K., Rockoff, J. E., & Staiger, D. O. (n.d.). Teacher Effects and Teacher-Related Policies. Annual Review of Economics, 6, 801–825. https://doi.org/https://doi-org.libproxy.lib.unc.edu/10.1146/annurev-economics-080213-040845

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. Economics of Education Review, 27(6), 615–631. https://doi.org/https://doi.org/10.1016/j.econedurev.2007.05.005

McGuinn, P. (2011). Simulation reform: Race to the Top, competitive grants, and grants and the Obama education agenda. *Education Policy,26*(1), 136-159.

Meyer, E., & Klaveren, C. V. (2013). The effectiveness of extended day programs: Evidence from a randomized field experiment in the Netherlands. Economics of Education Review, 36, 1-11.

National Association of State Title I Directors. (2010). *School Improvement Grants.* US Department of Education. https://www2.ed.gov/programs/sif/nastid2.pdf

North Carolina State Board of Education. (2010). *NC Race to the Top.* Department of Public Instruction. http://www.ncpublicschools.org/rttt/

Patall, E. A., Cooper, H., & Allen, A. B. (2010). Extending the school day or school year: a systematic review of research (1985-2009). Review of Educational Research, 80(3), 401-436. doi:10.3102/0034654310377086

Powers, J., Gomez, L., Holloway-Libell, J., Paufler, N. & Symonds, S. (2013). Race to the top. In J. Ainsworth (Ed.), *Sociology of education: An a-to-z guide*, 1, 629-630. Thousand Oaks, CA: SAGE Publications, Inc.