Improving the Evaluation of Summer Interventions: How Testing Intervals Affect Summer Learning Loss Measurements

By

Bryan Walker
M.A., Baker University, 2008
B.S., Creighton University, 2003

Submitted to the graduate degree program in Educational Leadership and Policy Studies and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Chair: John Rury, Ph.D.

Argun Saatcioglu, Ph.D.

Thomas DeLuca, Ph.D.

Brian Donovan, Ph.D.

ChangHwan Kim, Ph.D.

Date Defended: 16 May 2019

The dissertation committee for Bryan Walker certifies that this is the approved version of the following dissertation:

Improving the Evaluation of Summer Interventions: How Testing Intervals Affect Summer Learning Loss Measurements

Chair: John Rury, Ph.D.

Date Approved: 26 July 2019

Abstract

This study considered how the timing of assessments affects summer learning loss estimates and the evaluation of interventions designed to reduce summer learning loss. By utilizing an efficient assessment called reading curriculum-based measurement (R-CBM), measurements of seasonal learning were calculated with a minimized school influence schedule that compressed out as much school time from the summer vacation period as possible to compare with a more traditional schedule. Measuring based on the minimized school influence schedule resulted in summer learning loss of 17.370 words per minute compared to learning loss 5.564 words per minute. Based on these measurements, summer learning loss was underestimated by 219%. The study also revealed that controlling for testing interval by including the testing dates does not completely account for this underestimation because the learning growth rate at the end of the year is significantly greater than the rest of the school year. The growth rate at the beginning of the year was not significantly different than the rest of the year but this may be because of the tendency of teachers to focus on relationship building and classroom management at the beginning of the year. Follow-up interviews with teacher showed the increased growth rate at the end of the year possibly occur because teachers meaningfully change instruction to maintain the attention of their students as the school year end approaches. This study also considered if the patterns in the data support the findings of previous research of a compensatory effect of schools. The results show children on free/reduced lunch learn at a slower rate than children not on free/reduced lunch during the school year, but then the growth rates are similar during the summer. The effectiveness of Summer Boost, a learning loss intervention, were also considered. While utilizing the minimized school influence model improved the outlook for Summer Boost, the resulting gains were not statistically significant.

However, this result may be because of a small sample size (the treatment group included 68 students) rather than the quality of the program.

Acknowledgements

I would like to thank my family and friends for all their support and encouragement as I have completed this dissertation including Lyle Paczkowski, Adam Walker, Ross Paczkowski, Juan Beccera, America Beccera, Isai Beccera, Shane Ross, and Michelle Ross. I would especially like to thank my parents, Paul and Vicky Walker, for pushing me to do more with my career and making sure I had the support to be successful. Thank you to my wife, Lisa Walker-Beccera, for staying strong with me. This process included many peaks and valleys, and no one stood with me through the valleys more than Lisa. Thank you to my aunt, Dr. Mary Walker, for her push to motivate me to complete the process and my brother-in-law, Abrams Beccera, for his help with organizing and editing the final paper.

Thank you to my professors in the ELPS program and the Sociology Department. My time at the University of Kansas has been life-changing because of all the new colleagues and professors with whom I have had the chance to work. Thank you to my committee members, Dr. Argun Saatcioglu, Dr. ChangHwan Kim, Dr. Thomas DeLuca, Dr. Suzanne Rice, and Dr. Brian Donovan, for their great support and feedback. A special thank you to my advisor, Dr. John Rury, for his help planning my research, interpreting results, conceptualizing the follow-up study, and preparing the final draft. It has been nothing but a pleasure working through all the degree requirements with Dr. Rury when I know that is not the experience of many graduate students around the country. I would also like to thank Dr. Emily Rauscher, who was an original member of my committee, for her help conceptualizing this study and planning the data analysis.

Thank you to all my friends and colleagues at the school district that was the focus of this study. All of this was possible only through the flexibility and willingness to try new ideas of the central administrative staff. Thank you to the interventionist team for their professionalism and

collegiality through the years, and especially for their help completing all the assessments. Thank you to the participating teachers for their willingness to share about their classroom practices during the interviews.

Dedicated to my grandparents: Russell, Fran, Richard, Pearl, and Elvira-

for their dedication to their family, service to their country, and passion for education

Table of Contents

List of Figures

List of Tables

Chapter I- Introduction

Introduction/ Statement of the Problem

There is tremendous social and economic inequality in the United States. America is advertised as a meritocracy: a place where hard work can lead to a good life. The idea that education is an important route for social advancement is one of the common narratives of this ethos. Consequently, when low social and economic status (SES) students achieve lower levels of academic success than their high SES peers, schools receive much of the blame. It seems like an easy conclusion to assume the public education system is responsible for the achievement gap when, in many ways, schools serving low SES students seem to perform poorly compared to those serving high SES children. But this assumption is not that simple. The reality is that the problem goes well beyond the education system, and there are many deeper mechanisms allowing inequality among students to replicate. Moreover, for much of the populace, working hard and seeking higher levels of education does not directly translate into social advancement.

Summer learning research has started to bring new evidence regarding the explanation of the achievement gap. Researchers have found that both low SES and high SES students learn at similar rates during the school year but differ significantly over the summer to the point that some high SES students achieve academic gains while their low SES peers fall substantially behind (e.g. Heyns 1978; Alexander, Entwisle, & Olson, 2007b). This body of research suggests, while imperfect, schools can actually serve as compensatory institutions that ensure the achievement gap is not larger than it is (Downey, 2016). Outside-of-school factors play a significant role in the differential learning outcomes that produce the achievement gap (e.g. Heyns, 1978; Downey, von Hippel, & Broh, 2004). Nevertheless, schools are too easily made a

political scapegoat, and criticizing them has been a convenient strategy to avoid conversations about challenging inequality on a more fundamental level.

Unfortunately, summer learning research has not successfully challenged the established perspective on the achievement gap because it does not readily dictate clear action (Downey, 2016). One of the major challenges of summer learning research is the timing of the assessments. The school year and testing calendars generally do not align. Assessments are not typically given on the first and last days of school, so significant amounts of school time are potentially included in estimates. Cooper, Nye, Charlton, Lindsay, and Greathouse (1996) conclude: "…greater amounts of instructional time in summer intervals probably serve to mitigate the estimated negative impact of summer break. Therefore, the effect of summer vacation would likely be more detrimental, perhaps dramatically so, if it were measured from the day school is dismissed to the day students return. (p. 259)" Researchers have made efforts to control for this extra instructional time (e.g. Burkam, Ready, Lee & LoGerfro, 2004), but have not accounted for possible differential learning rates within the school year.

Because the roots of inequality lie at more basic levels of society, schools cannot eliminate the achievement gap, but they can take strategic action to narrow it. Schools can deliberately use resources to benefit at-risk students in ways informed by summer learning research. One compelling way they can help is through summer academic programs. In the past, evidence of the effectiveness of summer programming has been discouragingly mixed because there is considerable diversity of programs in their design and academic focus (McEachin, Augustine, & McCombs, 2016). In addition, many evaluations of summer programs have utilized observational designs that have not accounted for selection bias (Cooper, Charlton, Valentine, Muhlenbruck, & Borman, 2000).

Purpose and Description of the Study

The purpose of this study was to examine how the timing of assessments affects summer learning loss measurements in reading achievement. The study took place in one school district in a relatively small Midwestern city. Study participants included all the 1st and 2nd grade students in the district. Assessment scores were collected for students as they finished the 2016-17 school year and began the next grade level. The study was designed to minimize the amount of in-school time in summer learning estimates by utilizing reading curriculum-based measurement (R-CBM) instead of traditional achievement tests. R-CBM has proven to be a reliable, valid measure of reading achievement and it has shown to be moderately predictive of scores on more comprehensive reading tests. When testing, students read a passage for one minute while a human proctor scores the number of words read correctly (Sandberg Patton & Reschly, 2013). Because the test is time efficient, it is more practical to give it nearer to summer vacation. Scores were collected using the traditional schedule, in which students are tested several weeks before the academic year ends and several weeks after school begins again in the fall. Scores were also collected using the minimized school influence schedule, testing within a few days of leaving for and returning from summer break to compress the amount of school time included. Students were assessed twice in the spring, about one month apart, and twice in the fall. The first spring score and the second fall score followed the traditional schedule. The second spring score and the first fall score followed the minimized school influence schedule (see Figure 1 below to see how the different schedules were conceptualized). The results from the two schedules were compared for significant differences in summer learning loss measurements. The data collected was also used to analyze if efforts to control for in-school time were effective and

to see if there was evidence of a compensatory effect of schools.

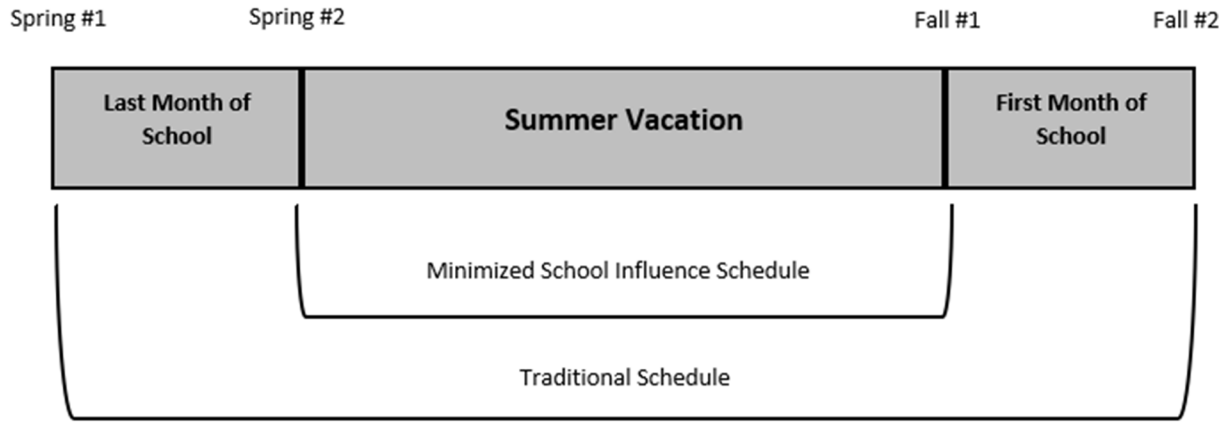

Figure 1: Diagram of Traditional and Minimized School Influence Schedules

This study also considers the effectiveness of Summer Boost, a summer academic program in the school district designed to reduce summer learning loss. The program was designed for students in kindergarten through 4th grade. Approximately 80 students from the 1st and 2nd grade sample participated in Summer Boost. Students were selected for Summer Boost based on low performance in reading achievement. Participating students attended academic programming three days a week for four hours throughout the summer (totaling 22 sessions). Busing was provided for students at no cost and students also received free breakfast and lunch. An experimental design was not possible, so this study utilizes propensity score matching to reduce selection bias. Logistic regression with nearest neighbor matching was used to match participating students with students who did not participate based on several covariates including free/reduced lunch status, ethnicity, gender, and school (Stone & Tang, 2013). The average treatment effect on the treated was calculated to measure the effectiveness of the program at reducing summer learning loss.

Significance of Study

This study is significant because it helps to address one of the biggest methodological concerns in summer learning research: testing intervals that do not align with the school year calendar. Students were assessed near the last and first day of school to remove as much school time from the summer learning measurement as possible. The study utilizes R-CBM, a type of assessment that is relatively new to summer learning research. R-CBM only takes one minute to administer and is designed to track learning rates with shorter intervals between assessments allowing consideration of how assessing at different time periods affects summer learning measurements.

Additionally, this study also provides updated summer learning data from the accountability era. The major summer learning research studies in the past were conducted in the 1970's, 80's, and 90's. The most recent important work was conducted right at the start of the accountability era and so these data do not account for differences in policy and instruction. This study can help to contribute to the slowly emerging body of more recent summer learning research and consider if historic patterns in seasonal learning have changed.

Theoretical Framework & Synthesis of Most Relevant Research

High SES students tend to attend better schools with more resources than their low SES peers. High SES students are also more likely to be assigned to higher tracks and receive more favorable treatment from teachers (Condron & Roscigno, 2003). However, seasonal learning research has shown that schools can have a compensatory effect for working class and underprivileged students as highlighted by three seminal studies. Heyns (1978) found in a study of middle school students in Atlanta that low SES students were able to keep up with higher SES students while school was in session but then fell significantly behind during the summer

months. The Beginning School Study (BSS) extended the work of Heyns, finding the achievement gap was well-established in the elementary years and that it had profound negative impacts on low SES students into adulthood (Alexander, Entwisle, & Olson, 2007a). Studies in the early 2000s, utilizing Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K) data, provided more recent, nationally representative evidence the seasonal learning pattern based on SES existed and helped to account for the inclusion of school time in summer learning estimates. The ELCS-K studies also examined summer learning across schools where earlier studies primarily examined patterns within schools (Burkam, Ready, Lee, & LoGerfo, 2004).

The implications of this research are that non-school effects and other exogenous societal forces largely account for the achievement gap. This dissertation study works from the assumption that there are school and non-school effects that impact the achievement gap, and in many ways, the non-school effects overshadow school effects. This notion means that schools can play an important role in closing the achievement gap but cannot eliminate it without support from other institutions in society. Unfortunately, schools are easy political scapegoats, and political will does not exist in the United States to make widespread changes to the economic system required to eliminate the achievement gap. While non-school mechanisms also affect the achievement gap, schools can do more to better serve low SES students.

The long summer vacation highlights how non-school effects can have a significant impact on the achievement gap. During summer vacation, families have more control over how children spend their time and the achievement gap tends to widen. Differences in summer time activities that influence learning derive from structural roadblocks and cultural mismatch. Some families have limited financial resources and less flexible schedules preventing them from accessing quality programs, while some parents lack the cultural capital to seek out quality

programs for children (Chin & Phillips, 2004; Burkam, et al., 2004). Educational institutions can help by providing quality academic programming during the summer at no cost to families to alleviate SES differences in summer learning rates.

Research Questions

- Question #1: Does assessing students with the minimized school influence schedule (spring #2 to fall #1) significantly change the measurement of summer learning loss with respect to reading fluency compared to the traditional schedule (spring #1 to fall #2) employed in other studies?
- Question #2: Does the gain rate for reading fluency achievement differ significantly for the first and last month of the school year compared to the rest of the school year?
- Question #3: Does the gain rate for reading fluency achievement during the academic year and summer differ significantly based on SES?
- Question #4: Is there a significant difference in reading fluency achievement for students who participate in the Summer Boost academic program compared to similar students who do not participate using test scores collected using the minimized school influence schedule?
- Question #5: Is there a significant difference in reading fluency achievement for students who participate in the Summer Boost academic program compared to similar students who do not participate using test scores collected using the traditional schedule?

Null Hypotheses

- $H_0$1: Assessing students with the minimized school influence schedule does not significantly change the measurement of summer learning loss with respect to reading fluency compared to the traditional schedule.
- $H_0$2a: The gain rate for reading fluency achievement does not significantly differ for the last month of the school year compared to the rest of the school year.
- $H_0$2b: The gain rate for reading fluency achievement does not significantly differ for the first month of the school year compared to the rest of the school year.
- $H_0$3a: The growth rate for reading fluency achievement during the academic year does not significantly differ based on SES.
- $H_0$3b: The growth rate for reading fluency achievement during the academic year does not significantly differ based on SES.
- $H_0$4: There is no significant difference in reading fluency achievement for students who participate in the Summer Boost academic program compared to similar students who do not participate using test scores collected using the minimized school influence schedule.
- $H_0$5: There is no significant difference in reading fluency achievement for students who participate in the Summer Boost academic program compared to similar students who do not participate using test scores collected using the traditional schedule.

Definition of Terms

Summer learning loss is the phenomenon of students losing ground academically while on summer vacation from school.

A *seasonal learning pattern* occurs when students learn at different rates throughout the year depending on the season. Typically, the seasonal learning pattern used in education research is the difference in learning rates between the academic year and summer. Seasonal learning patterns can differ by groups.

The *traditional schedule* is defined for this study as the time periods when schools typically assess students before students leave for summer vacation and when they return to school in the fall. Often schools will assess students several weeks before the school year ends and several weeks after the new school year begins.

The *minimized school influence schedule* is defined for this study as the time periods when assessments can be given at the beginning and end of the school year to reduce or eliminate the amount of school time included in summer learning measurements. While schools may have good reason for assessing using traditional model, ideally for summer learning research, the assessments would be given on the last day of school and the first day of the next school year to eliminate all in-school time.

A *summer intervention* is a program designed to eliminate or mitigate summer learning loss. A summer intervention strategically offered to students most likely to experience significant learning loss could help to close the achievement gap. For this study, the summer intervention is an academic program organized by the school district for students who are academically behind their peers.

Socioeconomic status (SES) is a measure of a student's social class that can be measured based on family income, parental education level, and type of parental employment. This study will measure SES based on family income as determined by federal and state free/reduced lunch

status guidelines. Throughout this study, *low-SES* will refer to poor and working class families, while *high-SES* will refer to middle class families.

Fluency is one of the components of successful reading that involves quickly and accurately reading words. While a seemingly simple task, fluency is actually complex and predicts success in more advanced components of reading. This is because students who cannot read fluently have to use most of their mental resources to decode words. If a student can read fluently, their brain decodes automatically leaving mental resources for deeper comprehension. This study uses an assessment of fluency because of its predictive ability and because it can be assessed quickly.

Curriculum-based measurement (CBM) is a broad category of assessments designed to measure the larger goals of the curriculum (instead of specific objectives) and to be administered efficiently so it can measure growth at close intervals.

Reading curriculum-based measurement (R-CBM) is a specific type of CBM that measures reading fluency by having a student read a leveled passage for one minute and then scoring the number of words read correctly during the minute.

Chapter II- Literature Review

Introduction

While schools often get the blame for the achievement gap, summer learning research points to non-school effects such as the family and community as the primary cause. Past research studies have identified a general, consistent pattern across studies. Students from different social class backgrounds often learn at similar rates during the school year. However, during the summer, students from working class and underprivileged families fall significantly behind middle-class peers. This pattern suggests schools actually play a compensatory role by helping to prevent the achievement gap from growing even larger.

Nevertheless, this actuality does not absolve schools of responsibility from doing more to help their most vulnerable students. There are very real differences between schools that serve middle-class and working class families that contribute to differences in achievement levels. Academic summer programs are a promising strategy schools can employ to help at risk students. These programs can be very expensive for parents when not provided with public resources. While there are mixed results on the effectiveness of summer programs altogether, well-designed programs that ensure sufficient academic time throughout the summer and employ effective teachers can play a profound role in strategically closing the achievement gap.

The Achievement Gap and Its Potential Causes

There is a significant achievement gap in school performance for low-SES and minority children in the United States when compared to children from advantaged, upper-class backgrounds. Children from high-SES backgrounds tend to attain higher levels of academic achievement than children from low-SES families. Many low-SES children enter school at a young age already behind their peers academically (Lee & Burkam, 2002) and the gap only widens as they progress through school (Farkas & Beron, 2004). Low-SES youth are much less

likely to enter a college preparatory course of study in large part because their achievement levels at the time of placement are already lower than their high-SES peers. Low-SES youth are much less likely to graduate from high school or attend a four-year college. The low academic achievement levels that contribute to this pattern are already in place early in a child's school experience (Alexander, Entwisle, & Olson, 2007a). The consequences extend into adulthood where low-SES students face diminished economic prospects and fewer opportunities for desirable positions in the workforce (Alexander, Entwisle, & Olson, 2014).

High-SES students typically attend schools with better teachers, better facilities, more extracurricular options, fewer discipline problems, and higher percentages of academically focused students (Downey, 2016). Even within large school districts, there are often wide disparities in resource allocation and academic achievement between schools along SES lines (Condron & Roscigno, 2003). However, when considering the independent effect of schools separate from such factors as family, neighborhoods, and communities, it is less clear that high-SES students attend better schools (Downey, 2016). Douglas Downey states:

> But it is difficult to know if these outcomes are a result of better schools or simply because the schools serve advantaged children in the first place. Interestingly, when we employ methods that attempt to isolate schools' contribution to learning, there is surprisingly little evidence that differences in the school experiences of high- and low-SES children explain achievement gaps. (p. 15)

The argument that schools are primarily responsible for the achievement gap is rooted in reproduction theory developed from a class-conflict frame of reference. Reproduction theorists argue that society is set up to be unequal and that schools are one of the institutions that help to reproduce inequality. There is evidence that schools are operated to correspond with the social

class of the students and thus reinforcing the student's position in society. Children from upper-class backgrounds on average attend schools with more resources at their disposal. Schools that serve lower-class students may emphasize skills that prepare them for low-income jobs such as obedience to authority, punctuality, and rote memorization instead of skills that prepare them for leadership roles (Bowles & Gintis, 2002). Even within schools, children from different social classes may encounter very different experiences. Children from upper-class backgrounds may be placed in higher ability groups and more advanced curriculum tracks than children from lower-class backgrounds. School personnel may unconsciously favor the speech patterns and style of dress of middle-class students (Downey, von Hippel, & Broh, 2004).

However, many researchers argue that institutions other than schools such as family and community play a bigger role in causing the achievement gap. The Bourdieusian perspective supports the argument that the achievement gap based on social class differences is caused by non-school effects (Downey, von Hippel, & Broh, 2004). This pattern is also consistent with a reproduction perspective. Parents from all backgrounds want their children to thrive emotionally, socially, and academically. Even so, some of them lack the means to provide children with enriching experiences to cultivate success. Parents from middle-class backgrounds typically have succeeded in school and can help their children do well in academics, while parents from lower-class backgrounds often struggled through school and thus can offer less assistance to their offspring. Parents play a key role in teaching and reinforcing skills in their children especially when they are younger (Alexander, Entwisle, & Olson, 2007b). Parents transmit habitus through parenting styles and the way they interact with authority figures around their children (Lareau, 2011). Habitus, as conceptualized by Pierre Bourdieu, is the tendencies and predispositions of a

person shaped by social experiences that influence how people interact with the world. Pallas (2016) describes the influence of habitus:

> These ways of viewing the world- specifying which social roles and behaviors are appropriate for people like them, and which are inconceivable, and not meant for them- point individuals toward differing destinations. French sociologist Pierre Bourdieu referred to such a world view as a habitus, a set of habits and ways of seeing the world that are forged through experience and operate at a subconscious level to propel people to act across a range of social settings in ways that reproduce the logic of those formative experiences. As individuals experience new environments, their habitus may change in response, but minor tweaking is much more common than a fundamental reorganization of how one sees the world. (pp. 119-120)

The habitus transferred to children along SES lines can be very different. For example, the way parents talk with their children can differ greatly, with high-SES families having a higher quantity of quality interactions than low-SES families (Hart & Risley, 1995). The habitus that middle-class parents transmit to their children closely matches the fields of school and professional work. The habitus of lower-class families is mismatched with schools and leads to decreased chances of academic success (Lareau, 2011).

It can be difficult to separate school and non-school effects when analyzing causes of the achievement gap because different factors tend to influence each other. Schools are evaluated in large part by student scores on standardized achievement tests. The target or "cut" score for passing the test is the same regardless of the background of the child. If children were assigned to schools randomly, this would be an objective way of evaluating the schools. In reality, parents select the communities they live in and the schools their children attend through a number of

means influenced by family income, parental education, proximity to work, and cultural capital (Downey, von Hippel, & Hughes, 2008). The level of a mother's education is highly predictive of the willingness of her family to consider and seek alternatives to neighborhood schools (Godwin & Kemerer, 2002). The makeup of a school's student population can vary greatly between schools and is not distributed randomly. While schools have substantial influence over some factors that affect achievement, other factors are not within the school's complete control. It can be very difficult to disentangle these school and non-school effects when searching for answers about the achievement gap (Downey, von Hippel, & Hughes, 2008). Complicating matters even more, certain families recognize the benefits of education and have shaped much of their children's non-school time to enhance school performance (Wiseman & Baker, 2004).

Summer Vacation Helps to Solve the Mystery

Summer break may be the key to getting the most complete picture of the achievement gap. The summer break that falls between the end and the beginning of any given academic school year is particularly long in the United States compared to other countries. The standard school year in the United States is about 180 days which ranks near the bottom in length compared to 25 other developed countries (Alexander, Entwisle, & Olson, 2007b). Comparing learning during summer vacation to learning during the school year constitutes a natural experiment that allows the separation of school and non-school effects. Children are with their families and neighborhoods throughout the entire year, but they are in school for only part of the year. So, learning during summer break is influenced by only non-school effects, while learning during the school year is influenced by both school and non-school effects. Schools do not typically have a significant influence on learning during summer break. The effects of non-school factors such as family and neighborhood are essentially held constant. Therefore, much

can be learned about the nature of school effects by looking at the difference between school-year and summer learning (Alexander, Entwisle, & Olson, 2007b).

Past research studies point to a consistent pattern along SES lines of the effect of summer learning loss on the achievement gap: students from low-SES families often can keep pace with their peers from higher-SES families during the school year, but then fall behind during the summer. Three seminal studies helped to establish this pattern: Barbara Heyns' Atlanta study, the Beginning School Study (BSS) in Baltimore, and the Early Childhood Longitudinal Study-Kindergarten (ECLS-K) Cohort.

While summer learning research for all students dates back to the beginning of the 20th century (Gold, 2002), the first researcher to consider how summer learning rates differ based on the background of students was Barbara Heyns (1978). She conducted a study of sixth and seventh grade students in Atlanta beginning in the early 1970s and considered how their scores would change on the word recognition section of the Metropolitan Achievement Test based on socioeconomic status and race/ethnicity. Heyns (1978) conceptualized summer vacation as a natural experiment where a researcher could isolate the effect of the school because students were in the school, family, and community environments during the school year but only in the family and community environments during the summer. Heyns found low-income students grew at a similar rate as high-income students during the school year but then lost significant ground to their high-income peers during the summer. Heyns reached two important conclusions: schooling has a significant independent effect on academic achievement and the effects of schooling for low-income and high-income students were more equal than expected. High-income students grew consistently throughout the year regardless of whether school was in session. In contrast, low-income students grew during the school year but then experienced

learning loss during the summer. Low-income families relied on schools more than high-income families to achieve academic progress. This discovery suggested that, while schools that serve high-income students may have important advantages, the global effect of schooling is positive for low-SES students.

The Beginning School Study (BSS), conducted in Baltimore starting in 1982, elaborated on Heyns' Atlanta study by considering how gaps in achievement in early school experiences carried over into high school and young adulthood. The BSS was designed to be a long-term longitudinal study that followed students as they entered first grade and continued to track them periodically for twenty-five years. The study sample consisted of 790 students scientifically selected in order to allow generalization to the citywide population. The students took the California Achievement Test twice each academic year in fall and spring which allowed researchers to create school year and summer learning estimates. In reading comprehension, low SES students entered school 0.5 grade equivalents behind their high SES peers. The gap grew to 3.0 grade equivalents by the time the students completed fifth grade (Alexander & Condliffe, 2016). The researchers found the students grew in parallel during the school year so the differences in summer learning explained almost all of this growing achievement gap (Alexander, Entwisle, & Olson, 2007b). The gap grew to 3.5 grade equivalents in ninth grade and students who experienced this lagging achievement level were much less likely to graduate high school and attend college. Alexander & Condliffe (2016) arrived at four major conclusions about the results of the BSS:

- "Children learn more, and learn more efficiently, when they are in school."
- "During the school year, lower-income children's skills improve at about the same rate as upper-income children's."

- "During the summer, lower-income children's skills do not improve, while upper-income children's skills continue to improve."
- "The summer learning shortfall over the elementary grades experienced by lower-income children has consequences that reverberate throughout their schooling into high school and beyond." (p. 28)

Research studies utilizing data from the Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K) extended the seasonal learning research from Atlanta and Baltimore in important ways. The ECLS-K was sponsored by the National Center for Education Statistics and tracked students through the elementary school years starting with students in kindergarten in 1998. The fact that ECLS-K utilized a nationally representative sample of about 20,000 students identified from about 1,000 randomly selected schools is a significant advantage (Downey, von Hippel, & Broh, 2004). The study also recorded the dates assessments were given and the beginning/end dates for school years. This information allowed researchers to control for any school time included in summer learning estimates which could be substantial (Burkam, Ready, Lee, & LoGerfo, 2004).

Cooper et al. (1996) reviewed thirty-nine previous studies on summer learning and conducted a meta-analysis of the thirteen most recent, highest quality studies at the time including data from Baltimore and Atlanta. The meta-analysis showed that the overall loss for all students was on average about one grade equivalent month with the decrease in mathematics being more significant than in reading. This conclusion likely downplays the effect of summer learning loss because typically the tests were not given on the last and first day of school, so the summer period incorporated several weeks of school instructional time. When examining results based on family income, the meta-analysis showed that all students regardless of income

experienced significant learning loss in mathematics. However, the subject of reading netted a different outcome. Comprehension scores dropped for both middle-class and lower-class students but significantly more for lower-class students. Middle-class students actually experienced significant increases in scores in reading recognition over the summer while lower-class students experienced decreases (Cooper, et al., 1996).

There have been few recent summer learning studies conducted during the accountability era but those studies that have conducted may be pointing to a change in the seasonal achievement pattern. Atteberry & McEachin's (2016) study of NWEA assessment data from a Southeastern state found, during the summer, students on average lose 25% to 30% of the academic gains they made in the previous school year in reading and math. The study found consistent patterns of summer learning loss during every summer between second and ninth grade. Atteberry and McEachin (1996) found significant variance between students in summer learning and school year learning rates. Most students achieve some level of gain during the school year, but the size of the gain is highly variable based on school demographic characteristics such as percentage of students qualifying for free/reduced lunch and percentage of non-White students. By contrast, about half of students experience summer learning loss, while the other half of students achieve gains. This variation in scores can contribute significantly to the achievement gap, as Atteberry and McEachin (2016) explain:

> In sum, the typical finding that, on average, students' learning slows, stops, or even regresses during the summer obscures a more problematic pattern: For some as-yet unknown reason, certain students actually maintain their school-year growth rates throughout the summer, while other students lose almost as much ground during the summer as they gained during the previous school year. When this happens in systematic

> ways so that the same students tend to have higher or lower learning rates summer after summer, it leads to a dramatic 'fanning out' of student outcomes as they progress through school. (p. 44-45)

Another recent study by Quinn and Le (2018) compared the earlier ECLS-K data to new data collected in 2010-2011. The analysis showed between-group achievement gaps based on SES and race/ethnicity narrowed during the kindergarten academic year and then widened over the summer. However, in contrast to previous research conclusions, the achievement gap continued to widen during the first grade school year and each academic year after. The possible explanation for this change is that inequitable school funding and re-segregation are accelerating in recent years, while the earlier important summer learning studies took place before the accountability era (Quinn & Le, 2018).

The Compensatory Effect of Schools

The patterns identified in seasonal learning research point to a conclusion that undermines the traditional narrative of the achievement gap: schools often serve as compensatory institutions, helping to prevent the achievement gap from growing larger even if they cannot eliminate it. Evidence to support this conclusion lies in the seasonal learning pattern: low-SES students tend to keep up with their higher-SES peers during the school year but then fall significantly behind during the summer. The seasonal research design allows for the comparison of all the school factors that advantage high SES students to all the factors that advantage low-SES students. This pattern provides an important opportunity to consider the overall effect of schooling even if the model is limited in its ability to identify specific mechanisms of consequence. There are indeed many individual characteristics of schools serving high-SES populations such as teacher quality, funding, and discipline patterns that are better than schools serving low-SES students. That being said, these between-school differences often are less

consequential than the overall effect of schooling. Low-SES students may need their schools, even with deficiencies, more than high-SES students to accomplish academic growth. High-SES students grow steadily throughout the calendar year, including the school year and summer. However, low-SES students may grow during the school year but then lose ground during the summer. This reasoning suggests, as flawed as the schools serving low-SES students might be, they can help compensate for the effects of non-school environments and help to prevent the achievement gap from widening further (Downey, 2016).

The seasonal learning model is not particularly well suited for identifying specific mechanisms, but Downey (2016) has extended some possible explanations regarding how schools serving low SES students accomplish this compensatory effect. First, the organization of schools by chronological age serves to ensure almost all students are exposed to relatively the same curriculum. This balancing effect is most significant at the elementary level because secondary students typically are grouped into curriculum tracks. High performing students often experience the same basic curriculum as underperforming students and are rarely exposed to material that challenges their academic abilities. To the extent to which this is true, it means low performing students are constantly being pushed to achieve higher levels while high performing students are not hard-pressed to the same extent. Second, there have been many educational policies enacted specifically to improve school outcomes for disadvantaged students including the Americans with Disabilities Act, Head Start, and Title I. These programs have helped to close funding gaps between low-SES and high-SES schools even if imbalances still exist. Even if these policies are not always creating environments for disadvantaged students to thrive, they are serving to enforce minimums and ensure that basic levels of services are available to all students. Finally, people who tend to become teachers are more equitable in their thinking and often value

helping disadvantaged students (Downey, 2016). Eighty percent of teachers are more likely to provide one-on-one help to struggling students instead of advanced students according to a national survey (Duffett, Farkas, & Loveless, 2008).

Structural and Cultural Explanations for Summer Time Use

If an overall effect of schooling is compensatory then there remains a question about how the achievement gap comes to be in the first place. If it expands more rapidly in the summer, differences in the way children and parents from different socioeconomic backgrounds use their time during the summer may help to explain variances in summer learning. Students from higher SES backgrounds participate at higher rates in enriching activities such as taking trips to museums, playing organized sports, taking swimming lessons, and using computers for educational purposes (Burkam, et al., 2004). On average, children from low-SES backgrounds spend more time watching television while children from higher-SES backgrounds spend more time talking with adults (Gershenson, 2013). However, researchers that study specific uses of time during the summer often report mixed results for activities that they theoretically believe should improve summer learning. A possible reason for this phenomenon may be that the survey data used for the research is better at measuring the frequency of the activities rather than the quality or nature of the activities (Burkam, et al., 2004). For reading achievement, a summer activity that is consistently associated with learning gains is participation in literacy activities (Burkam, et al., 2004). Children who read on their own or are read to by their parents often experience increases in summer learning (Phillips & Chin, 2004). Children who have more access to print by either having more books at home or going to the library regularly also experience summer learning gains (Burkam, et al., 2004).

There are two dominant theories in educational research to explain how social class influences academic achievement. The cultural theory is that schools are inherently middle-class

institutions and that some students and families are better prepared to succeed because of parenting practices and the cultivation of habits and worldviews that match the culture of the school. The structural theory is that society is set up to be unequal and the institutions of society (including schools) help to replicate that inequality through generations. A specific mechanism of structural theory in this instance may be that families have different levels of access to resources (including money and time) to provide learning opportunities for children that improve performance in schools.

The cultural argument is that the values and practices of parents when they interact with their children play a stronger role than the availability of time and resources. The reason why the summer time use studies had mixed results could be attributed to these activities only representing the intentional influence of social class on child development during the summer. Parents from all social backgrounds may plan activities for their children during the summer months, but these activities do not speak to the unconscious and inadvertent cultivation of habitus parents may accomplish simply by spending time and talking with their children. Upper-class parents pass along the knowledge, tastes, and mannerisms that are highly valued by schools and society. Summer vacation may provide children from privileged backgrounds greater amounts of time to spend immersed in this cultural landscape (Burkam, et al., 2004).

Lareau (2011) finds that parents from different social class backgrounds perform parenting tasks in different styles and these differences manifest themselves in many important ways in the lives of children. Middle-class parents constantly engage their children in a dialogue that helps the children learn to negotiate with adults and assert themselves. In contrast, lower-class parents tend to talk less with their children utilizing more commands instead of have ongoing conversations. Middle-class parents tend to structure their children's lives with

opportunities to interact with adult authority figures, while lower-class parents tend to leave a lot of children's out-of-school time unstructured. Middle-class parents advocate for their kids within institutions and teach their children how to do so for themselves. Lower-class parents tend to shy away from confronting institutional authorities and teach their children to react in this manner as well (Lareau, 2011).

The structural argument for why high SES-children participate in more enriching activities during summer break is that their parents have the financial resources and time to devote to their children's development. Budget limitations and rigid work schedules may prevent low-SES parents from providing such a wide range of activities for their children (Gershenson, 2013). Families with more resources are better able to provide activities and learning materials that may cost money such as computers, books, and vacations with educational value (Burkam, et al., 2004). Chin & Phillips (2004) find that the values and expectations of parents from different social classes are very similar. However, middle class parents work with far fewer constraints when acting upon those values and expectations. Most parents desire for their children to have enriching experiences, but only some have the means to provide them. Middle-class parents can control their availability to their children because they often have jobs with more flexible work hours and are better able to pay the required cost to get high-quality extended care for their children. Middle-class parents more often pick summer camps that are the best fit for their children regardless of cost or schedule. Poor and working class parents may still send their children to camps but are more likely to make decisions based on cost and transportation rather than the quality of the experience. Even when children have unstructured time, middle-class families are more likely to live in safer neighborhoods, so children can participate in more activities outside of their homes. Children from all social class backgrounds differ in their

motivation level to engage in creative, social, and academic activities, but parents from middle-class backgrounds have more resources to compensate when their children are unmotivated (Chin & Phillips, 2004).

In sum, differences in summer time use are influenced by both cultural and structural influences. Michael Boulay, in his unpublished dissertation work as described by Pallas (2016), shows how parenting practices change during the summer compared to the school year by social class. Boulay interviewed a sample of working and middle-class parents with children in 2nd grade to determine if parenting practices change seasonally. Two significant changes shared by families of all social classes were flexible bedtimes and increased screen time. The screens referred to in the study included cell phones, televisions, computers, and video games. However, two important differences are the perception of discontinuity of schedules and the focus on academic activities. Summer was a source of stress for many low-SES families because these families relied on schools to provide reliable supervision and structure. Managing child care was a constant challenge for these families, and their neighborhoods often lacked the safety necessary for their children to spend long periods of time playing outside. For many middle-class families, in contrast, summer was awaited enthusiastically as a break from the pressured schedule during the school year. While academic work was still expected at a much higher rate for middle-class children, the pace slowed down to allow for summer camps and vacations. As Pallas (2016) notes:

> In a sense, then, the shift from the school year to the summer Boulay (2015) described amplified the consequences of the child-rearing logics of concerted cultivation and the accomplishment of natural growth. Along with the faucet for school resources, the faucets for neighborhood and school resources flowed at a different rate during the

> summer, in ways that middle-class families could adjust to much more easily than those families of lesser economic means. Although many features of children's lives were less regulated during the summer, middle-class families continued to provide customized- and costly- experiences for their children, many of which could yield academic benefits. Conversely, working-class and poor children frequently were left even more to their own devices- especially devices with screens. (p. 122)

In the end, this dissertation study is strengthened by considering both structural and cultural mechanisms. Researchers from both camps even had some overlapping conclusions. For example, Lareau (2011) and Chin and Phillips (2004) agreed that families from different social class backgrounds often had similar values and expectations in regard to the importance of academic success. Where they differed was in describing how some families were better able to act on their values and expectations that resulted in success and the types of barriers less successful families encountered.

Summer Interventions as Possible Solutions to the Achievement Gap

Schools and communities can help to overcome barriers by providing access to educational resources and enriching experiences during the summer. Unfortunately, there is a significant opportunity gap in access to quality, affordable summer programs. According to survey data from the Afterschool Alliance, about one-third of children are enrolled in summer programs, and half of families that do not have access to a quality program would enroll their children if one became available. The average cost of a quality program is $288 per week which is even cost prohibitive for some middle-class families let alone families facing constant financial limitations. Yet, the opportunities for enrichment for middle-class families have grown significantly compared to low-SES families. According to the Consumer Expenditures Surveys,

from 1972 to 2006, the spending of the most affluent parents on enrichment activities increased about 250% compared to only 57% for the least well-off families (Pitcock, 2016).

However, the results of past research on how summer programs can mitigate summer learning loss have been mixed. This is in large part due to the differences in the structure of these school programs. Summer programs can look very different based on quality, length, and academic focus. In addition, the methods used to evaluate programs have varied greatly and often do not utilize experimental or quasi-experimental design (Cooper et al., 2000).

Cooper et al. (2000) conducted a meta-analysis of summer intervention research and found that only eleven of the more than three hundred studies they reviewed utilized an experimental/quasi-experimental design. The results show that summer programs had a moderate effect on student learning and that typically the achievement gain was equal to or larger than the estimate of summer learning loss (Cooper et al, 2000). McEachin, Augustine, and McCombs (2016) extended Cooper et al's (2000) work and reviewed twenty-five more recent evaluations of summer programs that utilized experimental/quasi-experimental methods. The researchers found a similar pattern: well-planned summer programs have a small-to-moderate positive effect (0.05 to 0.15 SD) on academic outcomes in math and reading. McEachin, Augustine, and McCombs (2016) summarize the results of research into the effectiveness of summer programs*:*

> Taken together, the evidence from the variety of summer programs implemented since 2000, along with Cooper et al's (2000) meta-analysis, suggest that many types of summer learning programs have the potential to reduce summer learning losses and perhaps create learning gains. However, implementing a summer program does not guarantee positive effects on students' learning. The variation in the results of these studies raises a key question: What factors make a summer learning program effective? (p. 197)

Designing Effective Summer Interventions

A growing body of research has started to identify consistent characteristics of summer programs that are most likely to achieve positive academic results (McEachin, Augustine, & McCombs, 2016). In their meta-analysis of summer literacy programs, Kim and Quinn (2013) found significant positive results for programs with a combination of seventy or more total program hours, at least four hours per day when the program is in session, and class sizes less than fourteen. Utilizing research-based instructional practices including modeling strategies and connecting reading to prior experience also result in more significant learning gains (Kim & Quinn, 2013). Teacher quality is an important consideration. When students work with high-quality teachers who teach their sending or receiving grade level, they perform better on academic assessments in the subsequent academic term. The site culture of the program can make a difference if it minimizes bullying, physical fighting, and disorderliness. Of course, high rates of student participation and attendance play a significant role in determining the effectiveness of summer programs. (McCombs, Pane, Augustine, Schwartz, Martorell, & Zakaras, 2014). Other program features associated with positive results include rooting instruction in evidence-based curricula, aligning lessons to students' needs, providing professional development to program teachers, and distributing instruction regularly throughout the summer (McEachin, Augustine, & McCombs, 2016).

Other Relevant Summer Learning Patterns for Study Design

There also exists an achievement gap by race/ethnicity that is affected by summer learning rates. Atteberry and McEachin (2016) find the Black and Latino students lost ground to White students during the summer and then gained it back at a lesser rate than White students during the school year. While race/ethnicity has an independent effect on seasonal learning rate, the significance of this effect is debatable because the effect of race/ethnicity is significantly

entangled with the effect of socioeconomic status (Atteberry & McEachin, 2016). When they enter high school, Black and Hispanic children are three years behind their White and Asian peers on average in literacy. In contrast, low-SES students enter high school with literacy skills five years behind high-SES students on average. This noticeable disproportion points to an important difference between the achievement gaps for race/ethnicity and socioeconomic status: the race/ethnicity achievement gap closed dramatically during the 1970s and 80s, while the SES achievement gap continues to widen (Reardon, Valentino, & Shores, 2012). However, since the late 1980s, there has not been much progress in closing the gap in part because in-school processes such as tracking and discipline policies have disproportionately hurt minority students (Darby & Rury, 2018). Quinn (2015) points out why it is important to consider race/ethnicity in achievement gap research regardless of the significance of the data:

> Black-White inequalities in SES-related measures are the legacy of slavery and racism (Wilson, 2009), and it is important to acknowledge the full extent of that impact. The task of explaining gaps should be kept conceptually distinct from the task of describing them; by blurring the lines between description and explanation, we risk glossing over or understating the racial element of inequality. (p. 52-53)

Studies that have tested gender have reported mixed results. Cooper et al. (1996) found that gender does not have a significant moderating effect on summer learning. However, gender does play an important role in achievement differences by subject (boys perform better in math, girls perform better in reading), so this demographic feature is important to the study as a control variable. In addition, the types of activities boys and girls prefer may be significantly different which warrants the inclusion of gender as a predictor variable for propensity score matching.

Age may play a key role in determining the rate of summer learning because the time when students enter school is a critical time to study. The achievement gap for older students is large but generally already in place coming out of elementary school. When students are younger and just starting school, they are more sensitive to the effects of their families and other non-school environments. The period when children are just entering school is thus a good time to study summer learning because they are relatively uninfluenced by school effects (Entwisle & Alexander, 1992). This timeframe is also important because cognitive growth occurs at a higher rate for young children, and it may be harder for students to catch up with their peers as they grow older (Burkam, Ready, Lee, & LoGerfo, 2004). In their meta-analysis, Cooper et al. (1996) show that all first and second graders regardless of socioeconomic status background tend to show moderate gains over summer, which then transition to significant losses for older students. Because many children tend to gain over the summer, any younger students who do decline are at a particularly strong disadvantage compared to their peers. One explanation for why summer losses start to become significant for older students is that parents may be less able to help students with more advanced academic material, especially in mathematics. (Cooper, et al., 1996).

Differences in the pattern based on academic subjects is significant. For all students, regardless of SES background, summer learning loss is greater for math related subjects than it is for reading and language skills (Cooper, et al., 1996). This pattern may occur because of differences in how mathematics and reading are treated in non-school environments. Families and communities may be better able to provide more opportunities to practice reading and improve vocabulary. Learning in the area of reading takes place in both school and non-school environments while learning in mathematics takes place primarily within school settings (Bryk &

Raudenbush, 1988). Conceptually-based skills such as math problem solving and reading comprehension are less susceptible to summer learning loss because, although they require many different experiences to develop, they do not require a lot of practice (and are less susceptible to decline). By contrast, skills that involve the memorization of procedural and factual knowledge, such as math computation and accurate spelling, require a significant amount of practice so are much more likely to regress over the summer (Cooper, et al., 1996).

Within-Year Oral Reading Fluency Patterns

This study utilizes reading curriculum-based measurement (R-CBM) because its format allows for testing closer to the summer break. R-CBM involves having students read aloud a passage for one minute while a proctor scores the number of words read correctly. R-CBM is a reliable, valid measure of overall reading ability and is somewhat predictive of performance on longer, comprehensive assessments including state assessments. It is designed to measure growth in shorter periods of time so it is particularly well suited for summer learning research (Sandberg, Patton, & Reschly, 2013). A more detailed description of R-CBM is provided in the following chapter.

Typically, schools that utilize R-CBM give the assessment at least three times a year: in the fall, winter, and spring. That being said, the assessment can be given more often (as often as once a week) so it can be utilized to measure the rate of growth during a single school year. Early studies suggested a linear measurement of growth rates was acceptable (Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993). However, later research including Nese, Biancarosa, Anderson, Lai, Alonzo, and Tindal (2012) found a discontinuous or seasonal measurement of growth was more accurate. In particular, several studies found students grew at a higher rate during the fall semester than the spring semester especially during the elementary years. In addition, the growth

rate for students in earlier grade levels was greater than for students approaching secondary school (Nese, et al., 2012).

Important Methodological Issues

Researchers have identified two major methodological issues regarding summer learning research: the interval of the assessments often does not line up with the school calendar and summer program participation is open to many sources of potential bias.

One of the major issues with summer learning research over the years has been the inconsistency between the testing window and the school year calendar. Ideally, to capture only the pure effect of the non-school environment, achievement tests should be given on the last day of school in spring and the first day of school in fall. However, often the tests used in research studies are administered several weeks or even months before school ends in the spring and after school begins in the fall. The measurement of summer learning time is contaminated by some school instructional time. The ECLS-K data included both information about the date assessments were administered and the dates for the beginning/end of the school year, so researchers have been able to use these data to analyze the issue. Burkam et al. (2004) used time lag estimates to control for interval differences, and their results showed that 80% of the summer growth in literacy, 84% of the summer growth in mathematics, and 50% of the summer growth in general knowledge were actually attributed to in-school learning in their study. In Cooper et al.'s (1996) meta-analysis, only one of the studies, the Sustaining Effects Study, reported summer academic gains, and this study contained four times as many students as the other studies combined. Consequently, it was solely responsible for the more conservative conclusions of the meta-analysis. The students in the Sustaining Effects Study were on average assessed five weeks before school ended and three weeks after school began. The rationale given for this inconsistent time gap was that teachers believed academic gains were unlikely during these

periods and that students needed the opportunity to get back into the school routine (Cooper, et al., 1996). While these justifications are valid under certain circumstances (i.e. before you choose what level of math course a student should enroll in for a semester), they are problematic for a study of summer learning. Having said that, it is likely that all studies of summer learning are affected by this methodological issue because it would be nearly impossible for schools to administer assessments to all students even on the same day, let alone the last and first days of school (Cooper, et al., 1996). When a study does not account for this lag in timing in its analysis, the researcher is assuming there is no such effect which is almost certainly not the case. The amount of school instructional time being included in summer learning estimates is significant and should be treated as such by researchers (Burkam, et al., 2004).

In addition, evaluating summer programs can be problematic because there are so many ways bias can affect the analysis. Most summer programs are voluntary so there is the potential for significant selection bias, especially when considering that families choosing to have children participate may share characteristics that differ significantly from those who choose not to participate. For example, an analysis of low-income families in Baltimore suggest there is significant variation in the importance families place on utilizing summer for academic growth. A family that places a significant value will likely achieve better outcomes for its children than families that do not regardless of summer program enrollment. The tendency for summer programs to have a limited number of seats also opens the door for the possibility of selection bias. For example, if a program has a “first come, first serve” policy for admission, the families that enroll children early will likely have different characteristics from families that enroll at the last minute. Because of their voluntary nature, summer programs are also susceptible to attrition bias because a family that ensures children finish a program are likely to have some different

characteristics from those that remove their children early that are unobserved in the analysis (Stein & Fonseca, 2016).

Chapter III- Methodology

Introduction

The review of the literature showed that summer learning patterns suggest non-school effects are more responsible for the achievement gap than schools themselves. There is not consistent evidence to support the conclusion that schools serving working class and poor families are significantly worse than schools serving middle class families based on impact after accounting for student characteristics. This means, while schools can and should do more to close the achievement gap, a more comprehensive strategy that incorporates additional resources would be necessary to close the achievement gap altogether. However, schools can utilize existing capabilities strategically in order to better serve at-risk students. The summer program that is at the center of this research study is an example of such a strategic use of resources that could hold some promise for reducing achievement gaps.

The review of the literature also highlighted an important methodological flaw in past summer research that this study is designed to correct. Students typically are not assessed on the very last day of school or on the first day of school. In fact, students are often assessed at time periods that ensure several weeks of school time have been included in summer learning estimates. Because of the way the ECLS-K data was collected, researchers have been able to control for this school time. However, there is a strong possibility that learning rates differ throughout the school year and that the time at the beginning of the year and end of the year might be significantly different than the rest of the school year.

This study is an early step in exploring the contrasting learning rates that might be occurring at different times during the school year and even during summer break. This is made possible through the use of curriculum-based measurement, a type of assessment that is meant

for quick, easy administration throughout the year and is specifically designed to identify growth. Very few summer learning studies have utilized curriculum-based measurement even though it holds much promise for exploring a different facet of the problem that may be able to lead to better, actionable data.

Research Questions

- Question #1: Does assessing students with the minimized school influence schedule (spring #2 to fall #1) significantly change the measurement of summer learning loss with respect to reading fluency compared to the traditional schedule (spring #1 to fall #2) employed in other studies?
- Question #2: Does the gain rate for reading fluency achievement differ significantly for the first and last month of the school year compared to the rest of the school year?
- Question #3: Does the gain rate for reading fluency achievement during the academic year and summer differ significantly based on SES?
- Question #4: Is there a significant difference in reading fluency achievement for students who participate in the Summer Boost academic program compared to similar students who do not participate using test scores collected using the minimized school influence schedule?
- Question #5: Is there a significant difference in reading fluency achievement for students who participate in the Summer Boost academic program compared to similar students who do not participate using test scores collected using the traditional schedule?

Null Hypotheses

- $H_0$1: Assessing students with the minimized school influence schedule does not significantly change the measurement of summer learning loss with respect to reading fluency compared to the traditional schedule.
- $H_0$2a: The gain rate for reading fluency achievement does not significantly differ for the last month of the school year compared to the rest of the school year.
- $H_0$2b: The gain rate for reading fluency achievement does not significantly differ for the first month of the school year compared to the rest of the school year.
- $H_0$3a: The growth rate for reading fluency achievement during the academic year does not significantly differ based on SES.
- $H_0$3b: The growth rate for reading fluency achievement during the academic year does not significantly differ based on SES.
- $H_0$4: There is no significant difference in reading fluency achievement for students who participate in the Summer Boost academic program compared to similar students who do not participate using test scores collected using the minimized school influence schedule.
- $H_0$5: There is no significant difference in reading fluency achievement for students who participate in the Summer Boost academic program compared to similar students who do not participate using test scores collected using the traditional schedule.

Study Participants

The participants in this study are the 1st and 2nd grade students in the Carsonville CISD school district selected during the 2016-17 school year and the study follows them as they enter

2nd and 3rd grades. This group consists of approximately 500 students. As much as possible, all the first and second graders were included in the study, because the diversity of the group's socioeconomic status and race/ethnicity characteristics help to make it comparable to other studies. This study focused on 1st and 2nd grade students because examining academic outcomes during early school experiences allows better consideration of non-school effects before students have been influenced by long periods of time in school. First grade is the earliest grade level in the school district when the R-CBM, the instrument used in this dissertation study that is described in detail later in this chapter, is given regularly throughout the school year. The study focuses on two grade levels for focus and consistency. Because it requires giving assessments at less conventional times, narrowing the focus to these levels helped to facilitate the efficient administration of assessments. The study also focused on two levels to ensure a sample large enough to get statistically significant results. One grade level of students likely would not have yielded enough power.

Approximately 80 students from this group participated in Summer Boost, a summer learning intervention provided by the school district. The participating students made up a treatment group that was then compared to a control group. The control group was created by calculating propensity scores using logistic regression for all students in the study and then matching participants to non-participants using nearest neighbor matching. The students were selected for Summer Boost using school district procedures. The district's main focus was to recruit at-risk students in order to help fill academic gaps. There were three steps to the selection process. The first step was to invite students that attended the summer program and/or after school program in the previous year with an attendance rate of 75% or more. These students were initially selected based on low achievement test scores. The second step was to invite

additional students with low achievement test scores from the current school year. They were invited to participate if they scored below the 25th percentile of national norms in Aimsweb Plus benchmark testing. The final step was to allow teachers to refer students for the program to fill the remaining spots. Some siblings of students eligible for the program were allowed to participate even if they did not meet the criteria. Behavior during the school day was also taken into account. Students either needed to have no office referrals during the 2017-18 school year or have the consent of the principal to participate. An equal number of slots were available for the students from the four elementary schools in the district. Attendance and behavior were used as selection criteria because the program had a limited number of slots and organizers wanted to serve students most likely to benefit from it. After the deadline for submission, open slots were filled from a waiting list regardless of school building. It is important to note the selection criteria includes the same assessment tool used in this study which is an important consideration for statistical modeling.

Description of School District

Carsonville CISD is a school district in a small Midwestern city that includes 4 elementary schools, 1 middle school, 1 high school, and 1 alternative school. The school district had a total enrollment of about students and employed about 250 teachers in 2016-17. The teacher/pupil ratio was 14.8 compared to the state average of 13.66 and a national average of 16.17 (National Center for Education Statistics, 2019). In 2014-15, the district had total revenues of $49 million and total expenditures of $43 million. The revenue per pupil was $12,475 compared to the state average $12,058 and national average $13,166. About two thirds of these resources come from state sources, a bit less than a third from local taxes, and less than ten percent from federal sources (National Center for Education Statistics, 2019). In 2016-17, the

district had an on-time graduation rate of 89.6% which was slightly above the state average of 86.9%. The dropout rate was 2.2% which was above the state average of 1.6%. The attendance rate in the school district was 91.8% which was below the state average of 94.7% (Kansas State Department of Education, 2019). The students in Carsonville CISD lag behind the state average in performance on the state assessment: 53.57% in mathematics and 56.16% in English/Language Arts achieve a Level 2 or higher, the minimum level of proficiency (Kansas State Department of Education, 2019). In 2016-17, 63.21% of the students were categorized as economically disadvantaged. All 4 elementary schools were Title I schools. 53.4% of the students were male and 46.6% were female. 57.4% were White, 18.2% were African-American, 10.3% were Hispanic, and 14.1% were categorized as other. 17.26% of district students were identified with disabilities (Kansas State Department of Education, 2019).

The school district is in the city of Carsonville which is close to a larger regional city. A relatively small city, Carsonville is a distinctive type of community compared to many: neither inner city, suburban, or rural although it does share some aspects of each. Carsonville includes a military base and several prisons. The population of Carsonville in 2017 was about 36,000. The per capita income was $28,743 and the unemployment rate was 4.9% (City Data, 2018).

Description of Summer Program

Summer Boost was an academic intervention program in Carsonville CISD paid for with Title I funds. The program operated 3 days a week throughout the entire months of June and July. Summer Boost was in session every Tuesday, Wednesday, and Thursday. There were 22 total days for the program throughout the summer. The program had approximately 4 hours of academic programming each day. Students participated in whole group reading, small group reading and math/science activities each day of the program. The program was provided at no

cost to families and free busing is provided for students. Students were provided breakfast and lunch each day at no cost through federal school lunch funds. The program took place at one of the four elementary schools in Carsonville. The program employed 1 administrator, 18 teachers and 5 teacher aides. The teachers were selected because of their current teaching assignments and their history participating in district intervention programs in the past. Teachers who taught the same grade levels as participating students and had the recommendation of their principal were given priority in hiring. An equal number of teachers were created from each elementary school to ensure all students were familiar with some of the teachers. The program utilized commercially available reading intervention curriculum. Staff members were provided two days of professional development before the start of the summer program to familiarize staff with the curriculum. The staff members were split into three instructional rotations and each rotation was assigned a coach to help match lessons to student needs. A field trip was organized to a local amusement park at the end of the summer for students with a high attendance rate to encourage attendance.

Instrumentation

This study utilizes an assessment called reading curriculum-based measurement (R-CBM). R-CBM is a type of assessment that involves having a student read a passage out loud for one minute while the proctor listens, calculates the number of words read correctly, and records errors. The passage is typically selected to be at the instructional level or grade level of the student (Fuchs & Deno, 1991). Passages at the same level are calibrated for readability and comparable difficulty in order to facilitate long-range goal tracking. R-CBM also utilizes standardized administration procedures that specify student directions, test duration, and scoring procedures (Fuchs & Fuchs, 1991). R-CBM is utilized in many schools for the purposes of

benchmarking, screening, program evaluation, and progress monitoring. In fact, it has been significantly researched because of its use in schools (Reschly et al., 2009) but is generally absent from the body of research on summer learning (Sandberg, Patton, & Reschly, 2013). R-CBM has several advantages compared to more traditional assessments used in summer learning research because it is designed to measure growth and is intended to be quick and easy to administer. It has proven to be a reliable and valid measure of reading ability and is predictive of performance on more time-consuming assessments while requiring only a fraction of the time to administer. R-CBM is well-suited to help address some of the recurring challenges of summer learning research it can be given quickly at different times during the year. Unfortunately, it has only been utilized in a small number of summer learning studies[1] that have been focused on very narrow populations so far (Sandberg Patton & Reschly, 2013).

One advantage of R-CBM is that it is designed to show growth. R-CBM assessment involves using multiple forms with comparable passages that can be given frequently, even as often as once a week (Reschly, et al., 2009). It is designed to be sensitive to small performance changes in relatively short time windows (Deno, 2015). R-CBM can be used to measure fluctuations in learning rates throughout the entire year including recoupment rates when students return to school after summer break (Sandberg Patton, & Reschly, 2013). By contrast, traditional achievement tests are not typically designed administered so frequently or to show changes over short periods of time. They are instead designed to compare the performance of students to their peers or to an achievement criterion (Baker & Good, 1995). However, to

[1] For example, Rosenblatt (2002) studied struggling readers in general education and Allinder & Fuchs (1994) studied students with intellectual disabilities.

effectively measure summer learning loss, an assessment needs to be given at least two times: in the spring and the fall (Alexander, Entwisle, & Olson, 2001).

Another advantage of R-CBM is it quick and easy to administer. R-CBM is very time efficient because it only requires 1 minute of time for an individual student to complete. R-CBM is easy to administer because it utilizes standardized administration procedures allowing a proctor to be trained and then be able to administer the assessment frequently without re-training. R-CBM assessments are also low cost for school districts which helps to account for their wide use (Reschly, et al., 2009). While a more traditional achievement test of reading ability may not be practical to administer on the first and last days of the school year, the R-CBM is a more realistic possibility because of the short period of time.

Despite its brevity, R-CBM has proven to be a reliable and valid measure of reading ability when evaluated based on traditional psychometric methods (Deno, 1992). This is because R-CBM is designed to measure broader curriculum goals instead of specific reading objectives. Focusing on the overall outcome of instruction instead of a series of quick tasks taught in isolation results in a stronger relationship between teaching and assessment than more traditional achievement tests (Deno & Fuchs, 1987). Curriculum-based measurements were conceptualized as an indicator of the vital signs of academic progress similar to how checking blood pressure and weight are used as general outcomes of overall health. R-CBM helps to determine how a student is doing overall by focusing on the bigger picture (Deno, 1985; Wayman, et al., 2007).

On the surface, it would seem that the R-CBM would only correlate well with decoding. However, analysis has shown that R-CBM has a strong correlation with several aspects of reading including comprehension, vocabulary, and decoding. This is powerful evidence

supporting the idea that R-CBM serves as a general outcome measure of overall reading ability (Reschly, et al, 2009; Hamilton & Shinn, 2003).

Over time, R-CBM has proven to be moderately to highly predictive of achievement of state reading assessments and nationally available comprehensive reading assessments. Yeo (2010) conducted a meta-analysis of studies analyzing the relationship between R-CBM and statewide reading tests. The meta-analysis included 27 studies that examined assessments from 14 different states. Most of the studies were completed after the passage of No Child Left Behind but 8 studies came before. Three of the studies incorporated scores from a different type of curriculum-based measurement called MAZE as well as R-CBM but most only used R-CBM. The overall correlation coefficient between R-CBM and the reading state assessments for the meta-analysis was .689. (Yeo, 2010). In their meta-analysis, Reschly et al. (2009) find a .67 correlation between R-CBM and other standardized reading assessments. Predictably, the strength of the correlation diminishes as the length of time between the R-CBM and the standardized test increases. (Reschly, et al, 2009). Reschly et al. (2009) states: "In reading, there is remarkable consistency in the relationship between R-CBM and other standardized measures of reading achievement across decades, samples, and various achievement tests."

This study utilizes R-CBM assessments provided by Aimsweb Plus. The primary reason Aimsweb Plus passages were utilized was because it was the adopted assessment program for the school district at the elementary level. However, Aimsweb Plus does come with some distinct advantages. First, Aimsweb Plus utilizes standardized administration protocols and scoring rules to ensure reliability of scoring. Aimsweb Plus also provides training so proctors are familiar with all rules and have a chance to calibrate scoring with each other. Second, Aimsweb Plus assigns

norms to raw scores based on a stratified sample the company collected to be representative of national demographics.

Variables in the Study

A significant part of this study is the comparison of growth variables calculated based on the minimum school influence schedule, traditional schedule, and the school year. For each time period, a gain variable was calculated by finding the difference between the beginning and end scores for the time period. This is the same procedure used to measure summer learning loss in many previous studies (i.e. Entwisle & Alexander, 1992; Burkam et al., 2004). Because of the substantial differences in time periods, rate of gain variables were also calculated by dividing this difference by the number of days in the time period. Participation in the Summer Boost is the independent or treatment variable. Important moderator/control variables for this study included grade level, lunch status, gender, race/ethnicity, and school. Grade level is determined by student enrollment records and important because learning rates can differ significantly by grade level. Free/reduced lunch status is a variable that measures family income based on state and federal guidelines and serves as a proxy for socioeconomic status in this study. The categories for gender are male and female. The categories for race/ethnicity are White, African-American, Hispanic, and Other. The school the student attended is also included as a control variable because the demographic makeup and neighborhoods of some of the schools are quite different.

Experimental Design

This study addressed the issue of the timing of assessments by testing students four times: about one month before the school year ends, on one of the last days of school, on one of the first days of the next school year, and about one month after the school year begins. By assessing four times, measurements can be taken to evaluate how different assessment points affect the

measurement of summer learning. The goal was to get as close to a thirty day difference between the two fall and two spring tests as possible to facilitate easier interpretation. This study utilized R-CBM, a one-minute oral reading fluency test that made measurement at these time periods more practical.

The students were assessed by teachers who serve as interventionists at each school using standardized administration protocols. The interventionist teachers were also highly trained and had significant experience at giving the R-CBM, and that helped ensure the reliability of the assessment administration. Students were assessed by teachers assigned to their own school. This group of teachers had received extensive training in the administration and scoring of the Aimsweb Plus R-CBM assessment. The teachers were trained directly by Aimsweb Plus trainers as well as by school district personnel. Annually, the teachers are trained and have the opportunity to calibrate scoring amongst each other to ensure inter-grader reliability. The teachers in this group collaborated approximately every month during the school year so they had regular opportunities to communicate and share insights. This team is made up of teachers with many years of experience administering R-CBM assessments.

This study utilized assessment scores already being collected by the school district for student improvement efforts and to track district goals so the administration would fit seamlessly into the schedule of the students. The traditional model assessments were given during the district benchmarking period. This means the students are given two passages to read and then the scores are averaged to get the final score. The minimum school influence model assessments were given outside of the benchmarking window and during the progress monitoring window. This means the students were given only one passage to read and the score they receive for this one passage was the final score. Students were given an assessment for the grade level they were

currently enrolled in at the time of assessment. The passages get slightly more difficult as the grade level increases. Proctors selected from a small set of passages that ensured the students did not have a chance to learn the passages but also were similar enough to produce reliable scores for comparison.

I took important steps to protect the rights of the human subjects in the study. The school district housed the data during the data collection phase subject to all applicable laws and school board policies. The data passed to me for data analysis did not include any direct identifiers. I also masked the name of the school district and city. These steps were necessary because, although the risk to research subjects is low, I collected private demographic data such as free/reduced lunch status and race/ethnicity. I received permission from the school superintendent to conduct the study and we agreed on procedures that protected the rights of students and their families under the Family Educational Rights and Privacy Act (FERPA). In addition, I designed the study in such a way as to be able to collect data the school district was already collecting. This not only prevents the students from having to take additional assessments, but it also allows for exploration of what is possible with data that school districts collect in future research.

Threats to validity

Because of its complex nature, several factors could potentially threaten the internal validity of this study. This study was designed to account for as many threats to validity as possible in order to increase the strength of the conclusions about summer learning and the effectiveness of the summer program.

1. Establishing causation can be difficult with studies of summer learning programs because any number of factors can influence learning rates and many of them can

be difficult to measure and account for in the research. The solution to this is to utilize a randomized experimental design with a control group. While this was not possible because of the school district policy for enrollment and expectations of participation for certain students, this study does utilize a quasi-experimental design in which students who participated in the summer program are matched with students who did not participate in the program using Propensity Score Matching.

2. Studies of summer learning programs are susceptible to selection bias because participation is typically voluntary and there are often a limited number of seats available. The families of students that participate may possess an unknown pattern of characteristics that differ from families of students that do not participate. Summer Boost was no exception because it had an enrollment cap and included selection criteria centered on academic need. This threat to validity was minimized with the quasi-experimental design matching participants with non-participants with similar demographic and achievement profiles for analysis.
3. Similarly, summer program research is also susceptible to attrition bias because students can exit a voluntary program at any time. A few students only participated in Summer Boost for part of the summer because of travel plans or because of fluctuating living arrangements with family members living out of town. A few students started the program and then dropped out. Once again, the families of these students may possess similar characteristics that differ from characteristics of families of students who finished the program. Parallel to this issue, family mobility is a variable that holds some explanatory power for how socioeconomic status

affects academic achievement yet can be very difficult to track. This study minimizes the threat to validity attrition poses by controlling for attendance. Careful attendance was kept during Summer Boost and the number of days attended sheds light on whether students only participated for part of the summer and how much academic programming the students were exposed to.

4. The time between the spring and fall assessments is also an important consideration for both the summer school evaluation and the analysis of learning rates. A lot can happen in the 2-3 months between assessments. Students can mature and family/living situations can change dramatically for some students. While assessments were given to track the progress of summer program participants, non-participating students were not assessed throughout the summer so no comparison of learning rates at shorter intervals can be made. This is accounted for by the fact students were assessed during similar time windows. The dates of assessments were also kept and accounted for in models when appropriate. In future research, R-CBM may be ideal for addressing this issue because its efficient administration would make it more possible to assess non-participating students.
5. The validity of the instrument is an important consideration: in this case, the ability of R-CBM to measure reading fluency and predict overall reading achievement. This study is limited to reading achievement because of the strong research base supporting R-CBM and its moderate correlation to more traditional reading achievement tests.

6. The reliability of the R-CBM of scores were also an important consideration. The assessments were given by many different proctors who could potentially administer and score the assessment in many different ways that might affect the comparison of scores. This study used standardized administration protocols to address this issue. The team of proctors received the same training and had the opportunity to calibrate scoring procedures. Aimsweb Plus also includes the exact same set of instructions to be used each administration and the same set of passages to administer. Aimsweb Plus also has a scoring guide with well-defined rules for scoring assessments.

Statistical Analysis

Individual information for each student in the sample was collected including how the students' scores during the four testing periods, free/reduced lunch status, race/ethnicity, and gender. Learning and growth rate variables were created for each student during three different time periods: the traditional model, the minimized school influence model, and the school year model. For the summer learning rate questions (questions 1-3), dependent sample *t*-tests at the .05 probability level were used to analyze if there were statistically significant differences between the means and standard deviations of these variables. Multi-variate linear regression was also used for question 3 to test if the relationship between oral reading fluency gain rates and free/reduced lunch status was affected by race/ethnicity and gender. STATA version 14.2 was used to complete the data analysis.

When an experimental design including random assignment is not possible, matched pairs can be used to reduce selection bias in an observational study. Propensity scores allow researchers to match subjects based on several, diverse covariates. Rosenbaum and Rubin (1983)

describe propensity scores as the "propensity towards exposure to treatment. (p. 47)" Treatment participants are then matched with non-participants based on the probability of being in the treatment group. The control group for the Summer Boost analysis (questions 4-5) was created by calculating propensity scores using logistic regression and then matching program participants to non-participants using nearest neighbor matching. This method was recommended by Stone and Tang (2013) for small, preestablished treatment groups common in educational research. The advantage of this method is that its straightforward design is more accessible to researchers and policymakers (Stone & Tang, 2013). Nearest neighbor matching identifies the non-treated student with the closest propensity score to a treated student. In this analysis, only one non-treated student was matched to each participant and it was possible for a non-treated student to be matched to more than one treated student. Stata version 14.2 was used for this analysis including the commands *psmatch2* and *teffects psmatch.*

Fourteen covariates were used to calculate propensity scores (the complete list is reported in Table 1). Optimal covariates for calculating propensity scores are related to the outcome variable but not the treatment (Garrido, Kelley, Paris, Roza, Meier, Morrison, & Aldridge, 2014). In this case, the outcome variable is the summer gain score as measured both by the traditional schedule and the minimized school influence schedule. Covariates for free/reduced lunch status, race/ethnicity, grade, and subject were included based on the established link with summer learning loss identified in previous research as described in the *Other Relevant Summer Learning Patterns for Study Design* section of Chapter II. Free/reduced lunch status serves as a proxy for socioeconomic status. Three variables for African-American, Hispanic, and Other are included for race/ethnicity. A student coded 0 for each of these variables is White. Interaction variables between lunch status and race/ethnicity were also included because of the interconnectedness of

these variables in affecting summer learning loss in the literature. Subject was captured in the analysis by inclusion of the initial Spring score and a math achievement score for each student. The math score was included because it provides some insight into the academic abilities of the students beyond only reading fluency and possibly the support available to students at home. Other covariates were included in the analysis because of their possible connection to the outcome variable and their ability to reduce bias in the propensity scores. The students from this study came from four different elementary schools so three covariates were included for three campuses. This means a child coded 0 for each school in the analysis attends the fourth school. These variables will ensure differences in schools are accounted for and may also provide insight into the neighborhoods because the schools are neighborhood elementary schools with small attendance areas. A covariate for whether a student is identified for special education is also included because of the large sample size and relationship to summer learning loss.

Table 1: Covariates for Propensity Score Matching

Covariate	Coding
Spring Score	Number of words per minute (continuous)
Grade	1 = 2nd grade, 0 = 1st grade
Free/Reduced Lunch Status	1 = qualified for free/reduced lunch 0 = did not qualify
African American	1 = Yes 0 = No
Hispanic	1 = Yes 0 = No
Other (includes Asian, Pacific Islander, American Indian, and multiple)	1 = Yes 0 = No
Math Concepts and Applications	Percentile Score 1-99
School 1	1 = Attends School 1 0 = Does not attend School 1
School 2	1 = Attends School 2 0 = Does not attend School 2
School 3	1 = Attends School 3 0 = Does not attend School 3
IEP (Special Education)	1 = Student has an IEP 0=Student does not have an IEP
Lunch Status X African American	1 = Qualifies for free/reduced lunch AND is African American

Lunch Status X Hispanic	1 = Qualified for free/reduced lunch AND is Hispanic
Lunch Status X Other	1 = Qualified for free/reduced lunch AND is Other ethnicity

Summary

This research project is designed to address one of the most significant methodological issues in past summer learning research: the lack of overlap between the school year and the testing calendar. Past estimates of summer learning loss are likely underestimated because of the significant amount of school time included in the estimates. This study seeks to minimize the amount of school time within summer learning estimates in order to get a more realistic view of summer learning loss and to get a better idea of the implications for summer program research.

This study also utilizes a relatively new instrument in summer learning research called R-CBM. R-CBM is a quick measure of reading fluency that is predictive of overall reading ability. Because of how easy it is to administer the R-CBM, it is possible to get assessment scores much closer to the last and first days of school. This tool also allows for the exploration of learning rates within the school year and within summer vacation instead of focusing on the difference between the two time periods.

Chapter IV- Results

Introduction

This quantitative study is designed to test how measuring summer learning loss at different periods affects the evaluation of summer interventions as well as the measurement of summer learning loss in general. Students were assessed approximately one month before school ended in the spring (spring #1), again near the last day of school in the spring (spring #2), as close to the first day of school as possible in the fall (fall #1), and again approximately one month after school started (fall #2). Data was collected for 556 students. Five students were excluded from the analysis because both demographic data was unavailable and because of missing test scores. 402 students had a full summer score profile (recorded scores for all 4 periods). 393 students had a complete score profile (including all 4 summer scores as well as Fall 2016 and Winter 2016 scores). Students without full profiles are included in the analysis as much as possible because transiency is an important possible explanation for the effects of socioeconomic status on learning rates. However, some statistical models require full score sets for students to be included. Summary statistics for the full sample of 551 students are included in Tables 2 and 3 below.

Table 2: Summary Statistics for Sample

Variable	Observations	Mean	Stand. Deviation
Summer Boost (1 = participant, 0 = non-participant)	551	0.145	0.353
Grade (0 = 1st, 1 = 2nd)	551	0.483	0.500
Lunch Status (1=free/reduced, 0=paid)	551	0.659	0.475
Gender (1=female, 0=male)	551	0.474	0.500
Race/Ethnicity			
White	551	0.532	0.499
African American	551	0.174	0.380
Hispanic	551	0.114	0.319
Other	551	0.180	0.384

Table 3: Tabulation of Important Statistics for Entire Sample

Variable	1	Percentage	0	Percentage
Summer Boost (1 = participant, 0 = non-participant)	80	14.52%	471	85.48%
Grade (0 = 1st, 1 = 2nd)	266	48.28%	285	51.72%
Lunch Status (1=free/reduced, 0=paid)				
Gender (1=female, 0=male)	261	47.37%	290	52.63%
Race/Ethnicity				
White (1=yes, 0=no)	293	53.18%	258	46.82%
African American (1=yes, 0=no)	96	17.42%	455	82.58%
Hispanic (1=yes, 0=no)	63	11.43%	488	88.57%
Other (1=yes, 0=no)	99	17.97%	452	82.03%

For analysis, gain scores were calculated utilizing similar procedures to past summer learning research (i.e. Entwisle & Alexander, 1992; Burkam, et al., 2004). the assessment scores were used to calculate differences in scores to measure gain (or decline) during the traditional schedule (spring #1 subtracted from fall #2), minimized school influence schedule (spring #2 subtracted from fall #1), and the school year (fall 2016 subtracted from spring #1). Because the time periods had significant differences in length, gain rates were also calculated by dividing these differences by the number of days included in the time period. Table 4 below show the average score during each time period as well as the average number of school days included. These numbers were used to build Figure 2 and 3 that demonstrate how removing as many school days from summer learning loss estimates dramatically changes the slope of the learning loss. These numbers were calculated with the entire sample of students but much of the analysis excludes Summer Boost participants to reduce bias unless the summer program is being analyzed. Figures 4 and 5 show how the relationship is affected by free-reduced lunch status.

Table 4: Average Scores and Number of School Days Included for Entire Sample by Time Period

Time Period	Average Score	Average Number of School Days Included
Fall 2016	**32.021**	**264.253 (9/4/16)**
Winter 2016	**48.682**	**165.945 (12/12/16)**
Spring #1	**66.351**	**32.914 (4/23/17)**
Spring #2	**74.731**	**2.246 (5/23/17)**
Fall #1	**58.396**	**2.028 (8/16/17)**
Fall #2	**59.671**	**22.144 (9/5/17)**

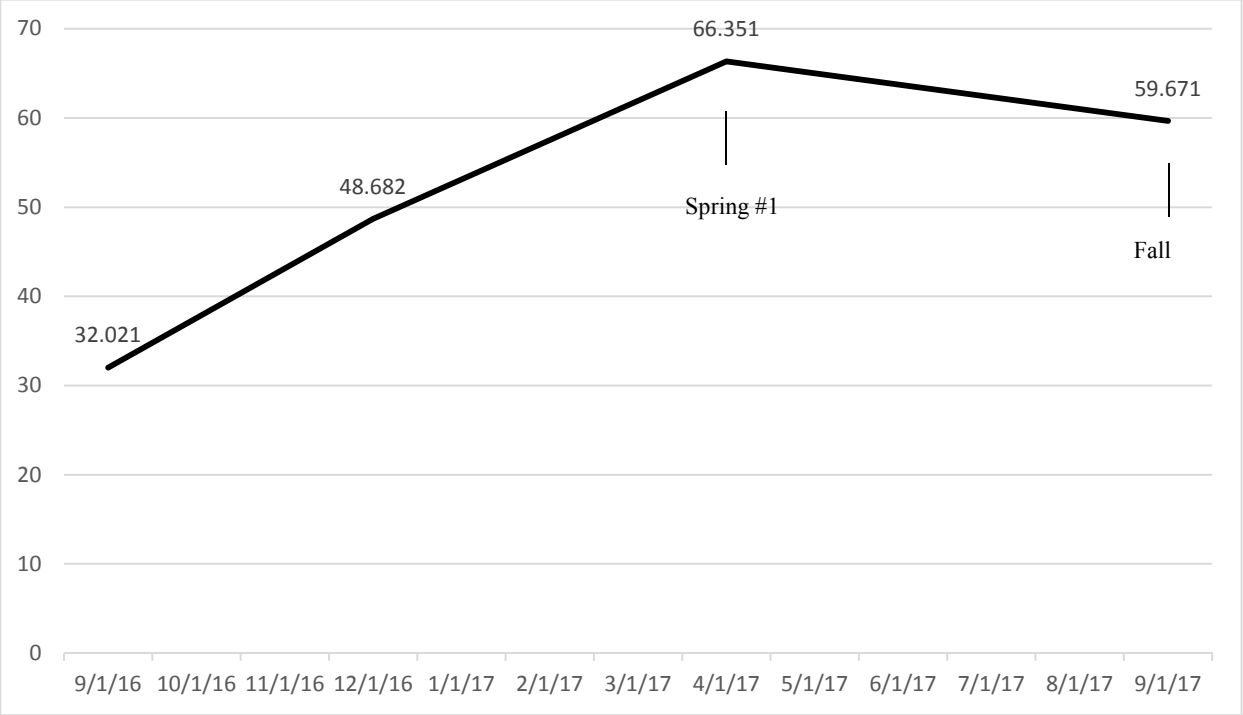

Figure 2: Growth Rates for All Students (Traditional Schedule)

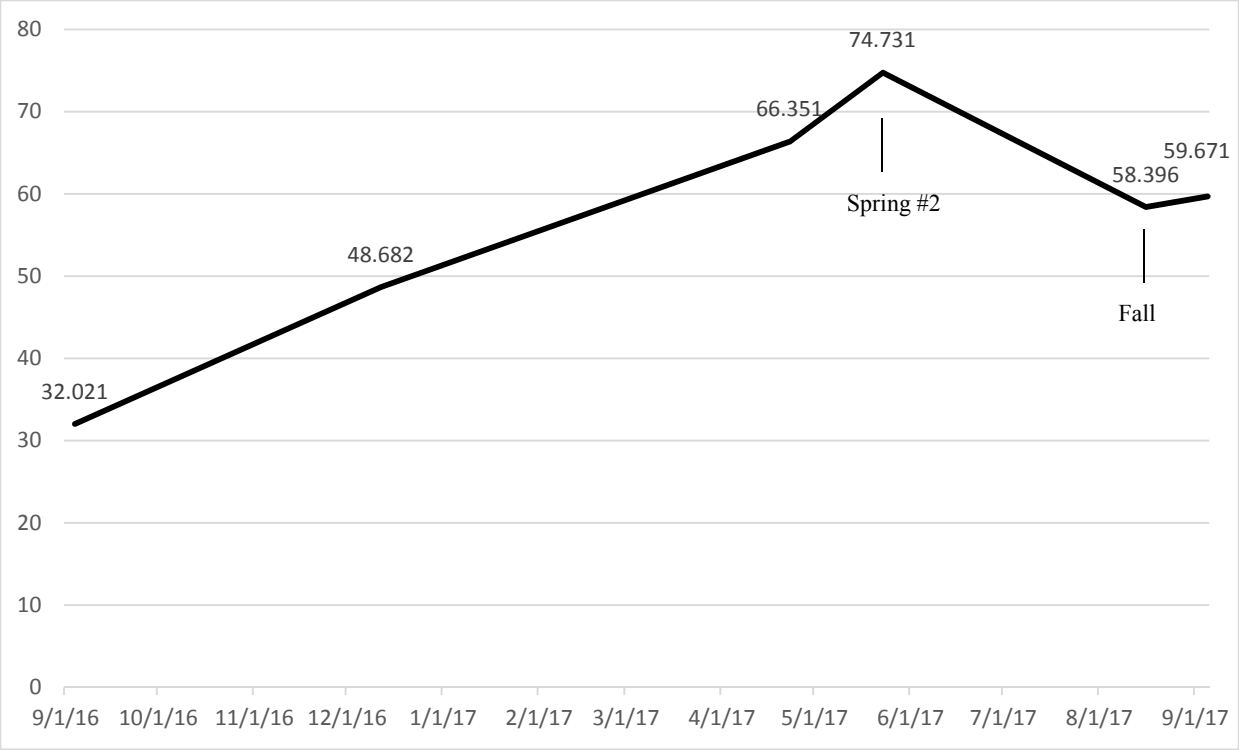

Figure 3: Growth Rates for All Students (Minimized School Influence)

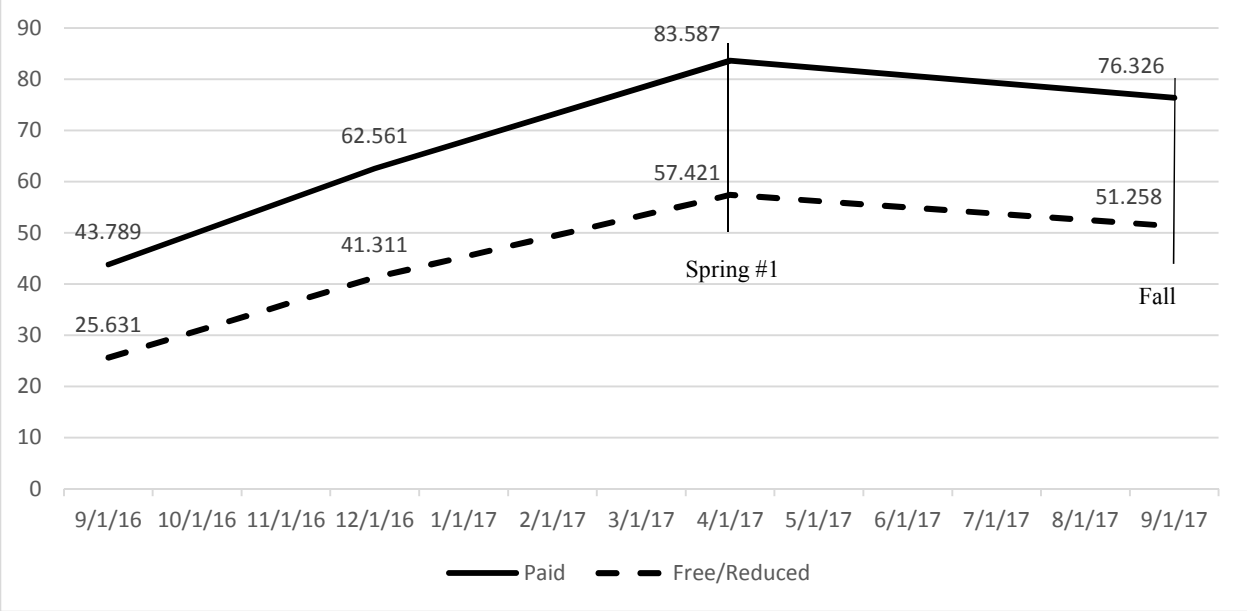

Figure 4: Growth Rates by Free/Reduced Lunch Status (Traditional Model)

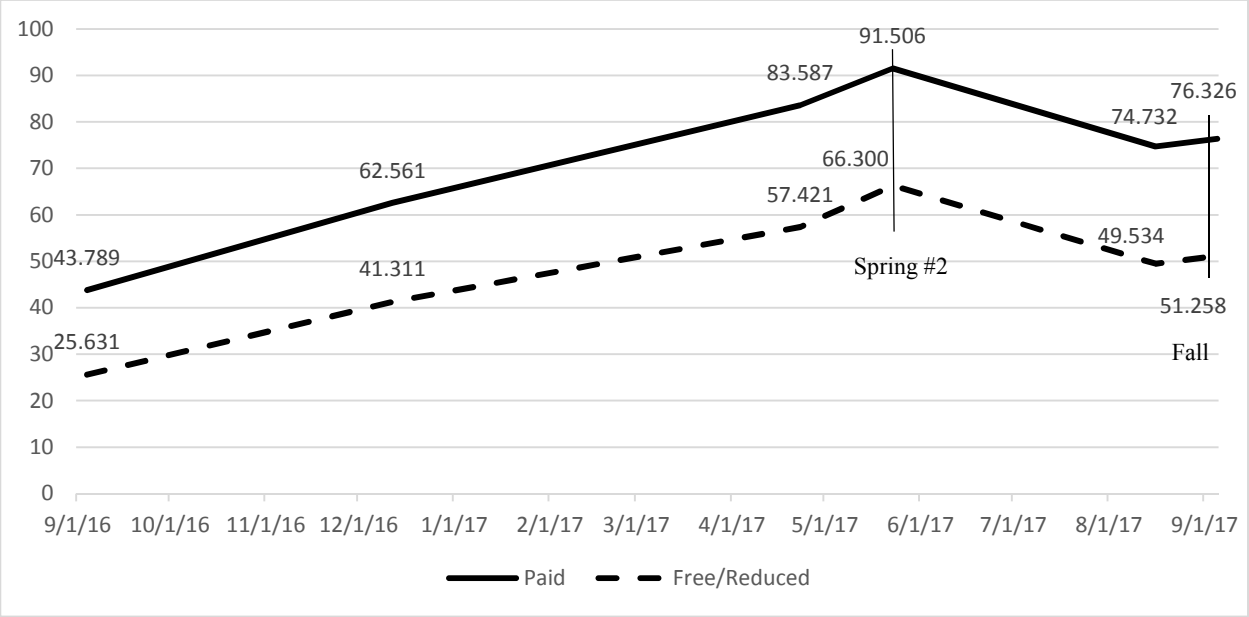

Figure 5: Growth Rates by Free/Reduced Lunch Status (Minimized School Influence)

Question #1

Does assessing students with the minimized school influence schedule (spring #2 to fall #1) significantly change the measurement of summer learning loss with respect to reading fluency compared to the traditional schedule (spring #1 to fall #2) employed in other studies?

Null Hypothesis $H_0$1: Assessing students with the minimized school influence schedule does not significantly change the measurement of summer learning loss with respect to reading fluency compared to the traditional schedule.

Paired *t* test results showed the difference of the mean of the traditional schedule gain scores (2nd fall score subtracted by 1st spring score) and the mean of the minimized school influence gain scores (1st fall score subtracted by 2nd spring score) was 11.806 word per minute. It is highly likely the difference in the means is statistically significant, $Mean_{trad} = -5.564$, $Mean_{msi} = -17.370$, $t\ (334) = 10.6998$, $p<.001$. Therefore, null hypothesis $H_0$1 is rejected. On

average, a student's score decreased by 5.564 words per minute during the traditional schedule compared to decreasing by 17.370 words per minute during the minimized school influence schedule. Summer Boost participants were excluded from this analysis to prevent the bias of results although similar results were obtained when they were included. The number of observations, means, standard errors, and standard deviations are reported in the Table 5 below.

Table 5: Comparing Summer Learning Loss Based on Traditional vs. Minimized School Influence Schedules

Variable	Observations	Mean	Standard Error	Standard Deviation
Traditional (Spring 1 to Fall 2)	335	-5.564	0.778	14.237
Minimized School Influence (Spring 2 to Fall 1)	335	-17.370	1.090	19.958
Difference	335	11.806	1.159	21.211

The rate of gain is also of interest for this research question because the length of each time period is different. The gain rates for each student were calculated by subtracting the earlier assessment score from the later assessment score and then dividing by the number of days between assessments. Paired *t* test results at the 95% confidence level show the difference in the mean of the gain rates for the traditional schedule and minimized school influence schedule show was 0.160 words per minute each day. It is highly likely the difference in the means of the gain rates between the two models is statistically significant, $Mean_{trad} = -0.042$, $Mean_{msi} = -0.201$, $t(334) = 12.872$, $p<.001$. Therefore, this analysis supports the conclusion that null hypothesis $H_0 1$ should be rejected. On average, a student's score decreased by 0.042 words per minute each day during the traditional schedule compared to decreasing by 0.201 words per minute each day during the minimized school influence schedule. The number of observations, means, standard errors, and standard deviations are reported in Table 6.

Table 6: Comparing Summer Gain Rates Based on Traditional vs. Minimized School Influence Schedules

Variable	Observations	Mean	Standard Error	Standard Deviation
Traditional (Spring 1 to Fall 2)	335	-0.042	0.006	0.105
Minimized School Influence (Spring 2 to Fall 1)	335	-0.201	0.013	0.231
Difference	335	0.160	0.012	0.227

Question #2

Does the gain rate for reading fluency achievement differ significantly for the first and last month of the school year compared to the rest of the school year?

-Null Hypothesis: $H_0$2a: The gain rate for reading fluency achievement does not significantly differ for the last month of the school year compared to the rest of the school year.

-Null Hypothesis $H_0$2b: The gain rate for reading fluency achievement does not significantly differ for the first month of the school year compared to the rest of the school year.

The rate of gain was used for this analysis because it accounted for the substantial differences in time period lengths being compared. The gain rates for each student were calculated by subtracting the earlier assessment score from the later assessment score, and then dividing by the number of days between assessments. Summer Boost participants were excluded from this analysis to prevent the bias of results although similar results were obtained when they were included. Paired *t* test results at the 95% confidence level showed the difference in the mean of the gain rate for the academic year (fall 2016 to spring 1) and the last month of school (spring 1 to spring 2) was -0.147 words per minute each day. It is highly likely the difference in the means is statistically significant, $Mean_{schlyr} = 0.135$, $Mean_{spring} = 0.282$, $t\ (381) = -4.999$, $p<.001$.

Therefore, null hypothesis $H_0$2a is rejected. On average, a student's score increased by 0.135 words per minute each day during the school year compared to increasing by 0.282 words per minute each day during the last month of school. The number of observations, means, standard errors, and standard deviations are reported in Table 7.

Table 7: Comparing Gain Rates for the School Year and Last Month of School

Variable	Observations	Mean	Standard Error	Standard Deviation
School Year (Winter 2016 to Spring 1)	382	0.135	0.006	0.119
Last Month (Spring 1 to Spring 2)	382	0.282	0.028	0.542
Difference	382	-0.147	0.029	0.575

A similar analysis was conducted to compare the first month of school to the rest of the school year. Paired *t* test results at the 95% confidence level showed the difference in the mean of the gain rate for the academic year (fall 2016 to spring #1) and the first month of school (fall #1 to fall #2) was -0.016 words per minute each day. It is not likely the difference in the means is close to 0 and not statistically significant, $Mean_{schlyr} = 0.138$, $Mean_{fall} = 0.154$, $t\ (350) = -0.415$, $p=0.678$. Therefore, null hypothesis $H_0$2b is not rejected. On average, a student's score increased by 0.138 words per minute each day during the school year compared to increasing by 0.154 words per minute each day during the first month of school. The number of observations, means, standard errors, and standard deviations are reported in the table below.

Table 8: Comparing Gain Rates for the School Year and First Month of School

Variable	Observations	Mean	Standard Error	Standard Deviation
School Year (Winter 2016 to Spring #1)	351	0.138	0.006	0.117
First Month (Fall #1 to Fall #2)	351	0.154	0.038	0.714
Difference	351	-0.016	0.039	0.723

Question #3

Does the gain rate for reading fluency achievement during the academic year and summer differ significantly based on SES?

-Null Hypothesis $H_0$3a: The growth rate for reading fluency achievement during the academic year does not significantly differ based on SES.

-Null hypothesis $H_0$3b: The growth rate for reading fluency achievement during the academic year does not significantly differ based on SES.

The rate of gain was used for this analysis because it accounted for the substantial differences in time period lengths being compared. The gain rates for each student were calculated by subtracting the earlier assessment score from the later assessment score, and then dividing by the number of days between assessments. Summer Boost participants were excluded from this analysis to prevent the bias of results although similar results were obtained when they were included. Two-sample *t* test results at the 95% confidence level showed the difference in the means of the gain rate for the school year (winter 16 to spring 1) for students who did and did not qualify for free/reduced lunch was 0.036 words per minute each day. It is highly likely the difference in the means is statistically significant, $Mean_{paid} = 0.177$, $Mean_{free} = 0.140$, $t\ (400) = 3.811$, $p<.001$. Therefore, null hypothesis $H_0$3a is rejected. On average, the score of a student qualifying for free/reduced lunch increased by 0.140 words per minute each day during the school year and the score of a student not qualifying for free/reduced lunch increased by 0.177 words per minute each day. The number of observations, means, standard errors, and standard deviations are reported in the table below.

Table 9: Comparing Gain Rates for the School Year by Free/Reduced Lunch Status

Variable	Observations	Mean	Standard Error	Standard Deviation
Paid	141	0.177	0.008	0.093
Free/Reduced	261	0.140	0.006	0.900
Combined	402	0.153	0.005	0.093
Difference		0.036	0.010	

A similar analysis was run to compare gain rates during the summer based on free/reduced lunch status. Two-sample *t* test results at the 95% confidence level showed the difference in the means of the gain rate for the summer (spring 1 to fall 2) for students who did and did not qualify for free/reduced lunch was 0.017 words per minute each day. It is likely the difference in the means is close to 0 and is not statistically significant, $Mean_{paid} = -0.027$, $Mean_{free} = -0.044$, $t(369) = 1.480$, $p=0.140$. Therefore, null hypothesis $H_0 3b$ is not rejected. On average, the score of a student qualifying for free/reduced lunch decreased by 0.044 words per minute each day during the summer and the score of a student not qualifying for free/reduced lunch decreased by 0.027 words per minute each day. The number of observations, means, standard errors, and standard deviations are reported in Table 10 below.

Table 10: Comparing Gain Rates for the Summer by Free/Reduced Lunch Status

Variable	Observations	Mean	Standard Error	Standard Deviation
Paid	128	-0.027	0.010	0.118
Free/Reduced	243	-0.044	0.006	0.100
Combined	371	-0.038	0.006	0.107
Difference		0.017	0.012	

To further examine the relationship between free/reduced lunch status and school-year and summer growth rates, a multivariate linear regression model was used that included race/ethnicity, gender, and interaction terms as control variables. The details of this analysis can be found in Appendix B. This analysis was completed to consider if free/reduced lunch status performs a principal role in different seasonal learning rates or if other demographic categories

strongly influence the relationship. If the results showed the effect of free/reduced lunch status was significantly reduced by the inclusion of other demographic and interaction variables, the results of the *t* test would be less impactful because it would exclude other important information. That being said, similar results were obtained compared to the two-sample *t* test model above. Oral reading fluency gain rate score differences were statistically significant based on free-reduced lunch status during the school-year but not statistically significant during the summer. The results of the *t* test bear scrutiny during the school-year because free/reduced lunch status proved to have a strong independent effect even after the inclusion of additional control and interaction variables. Of note in the summer analysis, the variable representing the interaction between free/reduced lunch status and gender was statistically significant. The oral reading fluency scores for female students qualifying for free/reduced lunch decreased by 0.044 words per minute each day compared to male students not qualifying for free/reduced lunch. The regression analysis upholds the *t* test results from above because the seasonal effects of free/reduced lunch status mirrored the results of the *t* tests.

Question #4

Is there a significant difference in reading fluency achievement for students who participate in the Summer Boost academic program compared to similar students who do not participate using test scores collected using the minimized school influence schedule?

> -Null Hypothesis $H_0 4$: There is no significant difference in reading fluency achievement for students who participate in the Summer Boost academic program compared to similar students who do not participate using test scores collected using the minimized school influence schedule.

The average treatment effect on the treated was calculated for this analysis comparing Summer Boost participants to a control group created by using nearest neighbor matching based on propensity scores. On average, the gain scores over the summer (spring #2 to fall #1) for Summer Boost participants were 2.735 words per minute greater than control group members, $Mean_{treated}$ = -11.309, $Mean_{control}$ = -14.044, Standard Error = 2.903, $z = 0.94$, $p = 0.346$, $n_{treated}$ = 68. This difference was not statistically significant; therefore, the null hypothesis was not rejected. However, the difference in the means is somewhat large so the lack of statistical significance may be more a reflection of small sample size. The covariates for the propensity score model are listed in the Table 11. The balance of the model was checked by comparing the means of the treated and control groups to see if there were any substantial differences. None of the differences in means were statistically significant so the propensity scoring method successfully identified a control group with similar characteristics to the treated group.

Table 11: Balance Check of Propensity Score Covariates (Spring #2 to Fall #1)

Covariate	Treated Mean	Control Mean	*t* score	Significance (*p*)
Spring Score	56.779	60.632	-0.52	0.602
Grade	0.441	0.574	-1.55	0.124
Lunch Status	0.632	0.706	-0.91	0.366
African American	0.250	0.279	-0.39	0.700
Hispanic	0.162	0.206	-0.66	0.510
Other Ethnicity	0.176	0.147	0.46	0.644
Math Concepts	35.721	31.441	0.88	0.382
School 1	0.191	0.162	0.45	0.656
School 2	0.132	0.118	0.26	0.797
School 3	0.235	0.250	-0.20	0.843
IEP	0.265	0.25	0.19	0.846
Lunch X African American	0.162	0.221	-0.87	0.387
Lunch X Hispanic	0.118	0.162	-0.74	0.462
Lunch X Other	0.118	0.118	-0.00	1.000

Question #5

Is there a significant difference in reading fluency achievement for students who participate in the Summer Boost academic program compared to similar students who do not participate using test scores collected using the traditional schedule?

-Null Hypothesis $H_0$5: There is no significant difference in reading fluency achievement for students who participate in the Summer Boost academic program compared to similar students who do not participate using test scores collected using the traditional schedule.

The average treatment effect on the treated was calculated for this analysis comparing Summer Boost participants to a control group created by using nearest neighbor matching based on propensity scores. On average, the gain scores over the summer (spring 1 to fall 2) for Summer Boost participants were 0.306 words per minute less than control group members, $Mean_{treated}$ = -3.467, $Mean_{control}$ = -3.160, Standard Error = 1.139, z = -0.27, p = 0.788, $n_{treated}$ = 75. This difference in the means was not statistically significant; therefore, the null hypothesis was not rejected. However, the signs did change in the average difference compared to the analysis in Question 4. The covariates for the propensity score model are listed in Table 12 below. The balance of the model was checked by comparing the means of the treated and control groups. None of the differences in means were statistically significant so the propensity scoring method successfully identified a control group with similar characteristics to the treated group.

Table 12: Balance Check of Propensity Score Covariates (Spring #1 to Fall #2)

Variable	Treated Mean	Control Mean	*t* score	Significance (*p*)
Spring Score	49.160	44.187	0.83	0.406
Grade	0.427	0.480	-0.65	0.515
Lunch Status	0.627	0.707	-1.04	0.302
African American	0.253	0.213	0.58	0.566
Hispanic	0.160	0.147	0.23	0.822
Other Ethnicity	0.173	0.213	-0.62	0.538
Math Concepts	35.747	30.867	1.05	0.295
School 1	0.187	0.187	-0.00	1.000
School 2	0.133	0.173	-0.68	0.500
School 3	0.227	0.253	-0.38	0.704
IEP	0.267	0.333	-0.89	0.376
Lunch X African American	0.173	0.160	0.22	0.828
Lunch X Hispanic	0.120	0.120	-0.00	1.000
Lunch X Other	0.120	0.160	-0.70	0.484

Conclusion

This dissertation study considered several questions regarding how differences in testing intervals affect summer learning measurement and intervention evaluation. The results for Question 1 show that the differences in summer learning loss averages as measured by the traditional schedule and minimized school influence schedule are significantly different. Question 2 showed achievement rates were different at the end of the school year compared to the rest of the year. This result has important implications for researchers using models that control for the date of assessment. Question 3 considered if the Carsonville CISD dataset replicated the seasonal learning pattern common in earlier research: low SES students learned at similar rates compared to high SES students during the school year, but then fall behind during the summer. The analysis for this question showed the learning rates for low SES were lower than high SES students during the school year. Question 4-5 considered how participants

in a summer program grew academically over the summer compared to a similar control group identified through propensity score matching. The analysis for neither question yielded statistically significant results but the differences in results by testing schedule show that including school year time in evaluation methods can change results in important ways.

Chapter V- Follow-Up Study

Introduction

Because of the reported findings in the quantitative analysis, this study features follow-up interviews with teachers in Carsonville CISD in order to help explain some of the patterns. The interviews are focused on how the last month and first month of the school year are different from the rest of the school year. The interviews also explored why teachers prefer to test at certain periods.

Sample Selection

Seven teachers were selected for interviews. The teachers were chosen to represent different patterns of R-CBM growth among their students in the 2016-17 school year. Teachers were selected from three groups: average growth rate during the school year but high growth rate in the last month of school, high growth rate during the school year but average growth rate in the last month of school, and high growth rate during both the school year and the last month of school. Table 13 shows the average growth for students by the primary teacher who instructs them in first and second grade in Carsonville CISD. The table also identifies the teachers selected for the interviews.

Theoretical Framework

This follow-up study works from the assumption that teachers may change how they teach and how they interact with students depending on the time period during the school year. Most relevant to the study was if students have experienced summer learning loss when they return to school, teachers must work with them to recoup any lost learning before they can begin to embark on new learning. Teachers may also spend a significant amount of time

Table 13: Identifying Teachers for Follow-up Interviews Using Growth Averages

Grade	School Year Growth (Fall 2016-Spring #1)	Last Month Growth (Spring #1-Spring #2)	Selected for Study; Identifier
1st	58.53	15.94	Yes; Teacher A
1st	29.85	12.07	No
1st	35.62	13.92	Yes; Teacher B
1st	52.96	7.67	No
1st	32.76	7.38	No
1st	49.67	5.44	No
1st	38.67	7.95	No
1st	43.47	7.2	No
1st	52.25	6.29	No
1st	60.17	3.82	Yes; Teacher C
1st	26.93	7.86	No
1st	21.89	7.72	No
1st	24.94	8.47	No
2nd	58.83	14.84	No
2nd	41.06	12.06	No
2nd	50	17.21	No
2nd	63.62	12.38	Yes; Teacher D
2nd	38.87	12.82	Yes; Teacher E
2nd	53.19	15.72	Yes; Teacher F
2nd	45.63	8.94	No
2nd	29.75	8.65	No
2nd	23.63	0.37	No
2nd	41.82	0.45	Yes; Teacher G
2nd	25.71	6.73	No
2nd	53.1	6.67	No

making students feel welcome, building relationships with them, teaching students how to function daily in school, and just generally building the culture of the classroom as documented by Harry and Rosemary Wong in *The First Days of School*. A number of things may influence how teachers teach close to the end of the school year, including the end of strenuous testing periods and the excitement as summer break approaches.

Research Design and Procedures

Interviews were scheduled for a total of forty-five minutes. The estimated length of each interview was fifteen to thirty minutes, but the extra time was allocated just in case they extended past the projected time limit. A specific protocol was developed to guide the interviews, but additional follow-up questions were asked when an interviewee raised an

important topic that warranted it. The protocol focused on questions about testing schedules and instructional differences at the beginning and end of the school year. The testing schedule questions were crafted to consider what preferences teachers have for the timing of assessments and why they possess such predilections. The instructional questions were crafted in order to explore the different ways that teachers change their instruction, why they change their instruction, and the impact of said changes. The interviews took place in the classrooms of the teachers at times when students were not present (i.e. before school, after school, and during planning time). The interviews were recorded using the Voice Recorder application available for Apple iOS devices. The application transcribed the interviews automatically, but the transcripts were carefully reviewed for grammar, formatting, and contextual accuracy. The researcher printed paper copies of the transcripts and then coded them for themes. The themes were beginning of year testing schedule, end of year testing schedule, beginning of year instructional changes, and end of year instructional changes. Additionally, possible quotes for inclusion in the results of the study were marked. The quotes were organized by theme and then by sub-themes. The theme and sub-theme structure was then used to inform the narrative.

Testing Schedules

In their meta-analysis, Cooper, et al. (2000) found that the Sustaining Effects Study was single-handedly responsible for the more conservative measurements of summer learning loss it reported. The Sustaining Effects Study included substantially more students than the other studies in the meta-analysis combined and was the only one to actually report summer learning gains. However, an important consideration is that the Sustaining Effects Study contained eight to nine weeks of school time in its testing interval. Students were typically assessed five weeks before the school year ended and three weeks after the school year began in the fall. Researchers

reported that teachers preferred to give the assessments at this interval because they believed learning gains were unlikely in a shorter testing window and students needed a chance to get used to being in school again (Cooper, et al., 1996). While teachers and school districts may have other legitimate reasons for assessing students when they do, the inclusion of substantial school time in the assessment interval between spring and fall has major implications for summer learning research.

In the current study, all seven teachers expressed a desire to wait a few weeks before administering any assessments of consequence at the beginning of the school year. Teachers gave many reasons for this preference including: letting students get used to being in school, having a desire to set the culture/routine of the classroom before getting into serious academic work, needing time to get to know the students before administering tests, and a strong belief the students will not show their true ability earlier.

Several of the teachers expressed a need to give students time to get back to being in school. As Teacher F put it, "because they're still trying to get in the swing of waking up, and they're not really in the 'school-mode' yet." Teachers expressed concern that students were coming from environments that were less intellectually stimulating and lacked structure. Teacher E was concerned about testing too soon because she believed students needed to get used to thinking at a rigorous level:

> "Because the students are still getting in the groove of things. I feel like, at the beginning, they're still getting used to being back at school and, you know, used to getting their brains working again. And a lot of the kids don't do anything over the summer so then, when they come back, they're not used to thinking so much."

Lack of structure at home during the off-school summer months was one of the main reasons teachers identified specifically as evidence as to why they believed students would not do well on the assessments if given too soon. Students were shifting from an environment of very little structure to an environment of heavy structure upon transitioning to school. Teacher G described how students need to learn the pattern of the school day in order to be successful:

> "Yes, I do think that the lack of structure is huge. Some of our kids are not, if you ask them: did you read this summer? 'No.' What did you do this summer? 'Ooh, I just played.' You know, playing is great, but playing along with today you have to read twenty minutes and that type of thing I think is huge, and so we, here, try to establish that structure. Today, we're going to read for twenty minutes. During this time, we're going to do reading here, we're going to do phonics here, and when they start falling into that pattern every day, that's what we're going to do, their behavior starts calming down because they're not, I think, afraid of what comes next. That's just me."

Several teachers identified the need to establish the routine and environment of their classrooms before administering any important tests. Some teachers even believed giving tests too soon would undermine their efforts to establish a positive classroom environment. Teacher D believed students thrived in less stressful environments:

> "When I did first grade, I felt like we tested letters and sounds right away, beginning of the year. And now that I've come to second-grade, I feel like they need to feel that this is more of a less-stress environment at the beginning. I want them to get, they've had the summer off, accustomed around the academic world before we test. So now, we don't do our Aimsweb till September, like the first week maybe, which I think is better than right when we get here because their brain is not thinking that way right after. So, yeah, we

> don’t do spelling tests the first couple weeks or any kind of evaluation, really. I just want to, kind of, get to know them and what they can do.”

Teacher C points out that students need to enjoy school to be successful and that testing too early can disrupt the enjoyment:

> “Well, because I'm focusing on really getting to know my students and just having fun. I think it's really important for the first three weeks of school. We just really focus on having a really good time in school and having fun and I devote all my time to my students and I'm just focused on that– focused on the routines and having fun and making sure the kids really like school and it's an enjoyable time.”

Like Teacher C, several teachers expressed a desire to get to know their students before administering any assessments. This belief is especially prevalent in the primary grades where one-on-one assessments are common; having information about the background of students and how they operate can help a proctor ensure students can be successful. For example, if instructors can tell a student is having a bad day, they can wait until the next day to give the assessment rather than getting a score that is not indicative of the student’s ability. Teacher A describes why it is important to get to know students before administering assessments:

> “So, the first couple weeks can't be about academics. It's got to be social. And then the testing can't be that first week because I don't know them yet. I don't know what sets them off. I don't know who has a rough home life yet. I don't know who is up at the crack of dawn; who's got to be at day care at six in the morning. I don't know who doesn’t have a bed time. So, once I kind of feel them out, I would know that, okay, this person I’m going to want to test in the morning because they have a really rough home life, and they're getting closer to going home, and it's not going to work well, so they're stressed

> about going home. I don't know who didn't have breakfast, and maybe I need to make sure they're on the 'free eater' list for breakfast, so they're not going to do well, so I should test them after lunch. I don't know any of that yet. So, with little kids that stuff is important, really important for them."

Teacher B actually expressed a desire to give the assessment both when students first arrive back at school and a few weeks after the start of school. Because the test is short and easy to administer, this teacher likely believes this is a realistic option:

> "I say both. I know we don't do Aimsweb right away, but it would be nice to see exactly what they come in with and then, in a month, how much did they retain and see, kind of, where kids are again, which I think we do informally all the time anyway. But also, there's another part of me that doesn't want to give them all that stuff at the beginning."

Similarly, most teachers expressed a desire to assess students several weeks before the end of the school year. However, this time it was less conclusive. Five teachers preferred to test several weeks before school ended, while two teachers preferred to test closer to the last day of school. Once again, teachers gave many reasons for assessing several weeks before the end of the school year, including: student behavior changes as summer break approaches, lack of academic focus for the students, the need to complete other tasks at the very end of the year, and, once again, a belief the scores would not be a true reflection of a student's abilities.

Several teachers reported that behavior changes in students in the last few weeks of school make it less conducive to learning. Students are excited about summer break and not paying attention as much as teachers desire. Teacher D expressed she prefers to assess students sooner because she has some control of the student's focus:

> "Even though I keep my expectations the same, the kids know, in their mind, that the time is getting short. And, I'd rather have their focus brain– about the beginning of May, very end of April, kind of where I still have a little bit more control than normal I think."

Similarly, Teacher E prefers to test several weeks before the end of school because students are more focused:

> "Because, they're checked out by the end of the year. If you did it a week before school is out, they're checked out already and they don't care. But if you do it a few weeks before, or a month before, I feel like that's, like, a great time, that they're still focused and not checked out."

Teacher A also prefers to test several weeks before the end of the school year because other priorities, such as report cards, become more important. In addition, Teacher A points out how over-testing, in general, can have a negative effect on assessment scores:

> "Not right at the end because I'm trying to get my own things done to get their last grade card out because there isn't any time; it has to go home the last day, so I would say probably maybe three weeks before the end of the year. They're not going to make that many gains in those last three weeks that would change that all that much, but that last week or two of school, I'm trying to get my own assessments finished, and that's just too much testing. We test them too much already, and they're six. Come on, they've only been in school two years. We assess them way more than is necessary because it doesn't show you anything more than what they can do on a given day, and the test doesn't reflect that Johnny's mom and dad broke up over the weekend, and now he's staying at grandma's house, and yeah, he tanked it on the test, but here's why. I could've told you that, but on a piece of paper, it doesn't show you what he can and can't do. Or he tested

> with somebody he doesn't know, and if he's shy, he's not going to talk to you. Whereas with me, he would've: 'da da da da da da dit.'"

Alarmingly, Teacher F shared the concern with over-testing and the negative consequences of the testing culture that has taken root in schools. Teacher E believes testing affects students in negative ways:

> "I think because they're stressed out and behaviors spike and then we're focused on basically what they're going to be tested on and maybe not like the whole picture of what they need to be learning about."

Moreover, Teacher E actually questioned the validity of testing students closer to the end of the school year and emphasized the importance of having strong relationships with students for accurate testing:

> "I really don't know if assessment the last two weeks or even the last month of school is very authentic at all. Just because of where they are. I even think of the validity when we did all the intervention when all of us used to go around and test. How valid is a test of a shy student that has me come in and doesn't know me and doesn't want to test the right way or behavior or whatever that day compared to what the classroom teacher could do? I always thought about those scores especially when I was the second-grade interventionist (math and reading) and the teacher would say, 'well this kid's not doing any of this in class' and I'll say, 'well look, they are with me so there's something there, either relationship or, you know, something that affects them, and I think just the whole 'end-of-the-year' thing probably plays a big part into it too.'"

Two teachers expressed a desire to test students closer to the end of the school year. Teacher G prefers to wait because it gives struggling students more time to learn the material they need to learn:

> "I think that depends on the kiddo. That's difficult because my high achieving kids could take it well before the school year ended and do just fine. My kids that are just making progress on a daily basis would benefit from waiting until the end of the year."

Interestingly, Teacher C prefers to wait until closer to the end of the school year because the benchmark testing acts as a signal to the students that the end of the school year has come. This marker makes it difficult to continue academic learning through the end of the school year.

> "Well, my preference is that we not assess them too soon. I think that we can give the assessments too early and then that almost tells the students that we're finished. So, I like it when we assess them maybe even the middle of May versus the very beginning of May. Sometimes, I think we start too early with our end-of-the-year assessments. I guess they've been told that they're giving a test, and it's the last one; it's the end of the year. I think that does have an impact. And it could be partially my fault because I'm gearing them up for taking those assessments and telling them that they do need to do a good job and to do their best. And, they've reached their goals so I guess people tend to relax after they've achieved their goal."

Both at the beginning and end of the school year, teachers generally prefer to assess students at times that will add substantial school time to summer learning estimates. At the beginning of the year, teachers prefer to wait in order to let students get used to being in school and in order to give themselves a chance to get to know their students. At the end of the school year, teachers prefer to wait because behavior changes in students can be detrimental to scores.

At both times, teachers expressed concern with the accuracy of scores if the tests are given too close to the last day or first day of school. The purpose of this study is not to question the reasoning of teachers and schools as to why they assess students when they do. Many of the reasons identified could potentially have legitimate evidence supporting them. Instead, the purpose is to point out that the normal operating procedure of a school district is not always ideally suited for educational research. Thus it is necessary for researchers to test on different schedules.

Beginning of Year Instructional Changes

In the quantitative analysis section, the differences in growth rates in the last month of school and first month of school compared to the rest of the school were an important finding of the study that warrants further consideration and explanation. The growth rates in the last month of school were substantially greater than the growth rates during the rest of the school. The differences in growth rates at the beginning of the school year were not statistically significant, but it is notable that the growth rates at the beginning of the year were lower than the rest of the year. This finding is important for summer learning research because it means getting an accurate measurement requires more than just controlling for the amount of school time in the estimates. Controlling for this element, of course, gets you a better estimate, but it can still be more accurate. In the follow-up study, teachers were asked how they treat the beginning and end of the school year differently and why that might account for the differing patterns in growth rates. This section will focus on the beginning of the year period and why growth rates dip below the growth rate of the rest of the school year.

All of the teachers interviewed talked about the importance of setting up the culture of the classroom in the first few weeks of school. This can take the form of explicitly teaching

students the procedures they will be following for the rest of the school year (i.e. how to line up, how to get the teacher's attention, etc.). Teachers work diligently to establish high expectations for students in both their academic achievement levels and social behavior. Having high expectations involves not only setting boundaries but also how a teacher goes about enforcing the boundaries. Teacher F points out the importance of enforcing the boundaries because inevitably "students will test the boundaries." Another way teachers set the culture of the classroom is to deliberately establish positive relationships with students and create experiences for students where they can create positive relationships amongst each other. In many cases, less time is focused on academic work in order to ensure an effective classroom culture is established. This choice might help to explain why the growth rate is lower at the beginning of the year. Nevertheless, this extra effort to establish an effective classroom culture can potentially produce returns later by building the independence of students.

Teachers reported several methods and activities used to establish the culture of the classroom. Many teachers reported they would explicitly review classroom procedures with students and then regularly practice with the students. Teachers utilized a number of 'get-to-know-you' activities in order to learn about their students' backgrounds as soon as possible. Many teachers utilized games more often during this period in order to make school fun for the students. Some teachers also tried to get parents involved in the classroom right away.

Six of the seven teachers emphasized the importance of establishing the structure of the classroom early because it will yield gains in student learning later in the year. Getting the structure in place involves teaching the rules and procedures to students and establishing the regular routine of the school day. Teacher C describes the process in this way: "And, we just practice, practice, practice. That's basically what we do. Every procedure that we need to know

in school." On that note, some of the procedures practiced include how to line up, how to get permission to get out of your seat, and how to put away your personal items when you enter the classroom. Students mainly know the way things run without explanation if structure is effectively put in place by the home room teacher. When asked about the importance of setting the structure of the classroom, teachers responded in similar ways. Teacher F responded:

> "So that I'm not spending the rest the school year teaching them the classroom procedures so that we can focus on the actual learning stuff and then building the relationship so that they can trust me and want to learn and do the best…because then I can spend the time working on reading and writing and math and all that stuff and not redirecting kids to get back on task."

Teacher G points out that getting the structure in place early allows the teacher the chance to relax and enjoy the school year:

> "I do, I do, and I think that if you do that the first three to four weeks then you can kind of take a breath, and you can kind of let go of some things also, and they can fall in without everything falling apart."

Another important strategy for establishing and maintaining a positive classroom culture was setting high expectations. Teacher A stated, "we do a lot of setting expectations, practicing talking about it, and then by the end of the year, those expectations are expected to be met." Expectations can come in many forms, including the level of academic work expected of a child or how they are expected to interact with other children. When asked why establishing high expectations were so important for student learning, Teacher A answered:

> "Because they don't know how to be successful without them. Without those expectations, they don't know what they're supposed to do, so they set their own

> boundaries, and their boundaries are largely not what I would use for them to do, and it causes disruption for other students, which slows their learning, so no, we don't do disruptions. We only have two rules—don't be hurtful and don't be disruptive. Those rules aren't going to change just cause it's the last week of school."

Teacher F highlights that setting and maintaining high expectations early will result in more independent learning behavior later:

> "In the first weeks, they kind of test the boundaries and are trying to kind of get a feel of what the school day's actually going to look like, and then, towards the middle of the year, they know what's expected of them, and they're doing pretty much what they're supposed to be doing."

Six out of seven teachers also identified building positive relationships with the teacher and among students as an important strategy for establishing a positive classroom culture. This took the form of teambuilding and 'get-to-know-you' activities. It also involved teaching students how to talk and interact with each other as friends. Several teachers pointed out the importance of the teacher building trust with the students in order to get the best academic work out of them. Teacher D explained why establishing trusting relationships helps students to realize the importance of school:

> "I think if, number one, they don't like me or trust me or you know respect me or think they're safe when they're here, they're not going to work for me; you know? School is not going to be important. It's going to be 'I don't want to get up today.' I just want them to be excited to come here every day, so I try to make something maybe at the beginning of those first two weeks, over-exciting things like some snack project or some art, so they

> can see that along with the structure that I have, a lot of structure, there is that's because I care but there is that fun, the fun piece too."

Teacher C emphasizes that strong relationship building early in the year is an effective way of winning the students over to the culture of the classroom:

> "Yes, that's the time I work on building relationships with my students and they understand the routines in school and I really want them to be excited about learning and want to learn and I think it's really important that I hook them in during that time."

Teacher G describes the importance of helping students build positive relationships with each other:

> "Well, for them to be nice to one another and to always know that they should have their classmates' back whether they like them or not because it's very difficult for them; some of them don't like each other; some of their parents don't like each other, and so, it's really hard for them to know or to understand because apparently, no one's teaching them that—that you have to be nice and respectful whether you like them or not; so, yes, that is the goal."

Six of the seven teachers also identified that the effort to establish a positive culture may require less academic focus or less introduction of new material in the first few weeks of school. A lot of time is utilized reviewing material from the previous grade level and teaching social skills. Teacher B points out that "teaching stepping stone skills make the rest of the year easier." Teacher C explains why reviewing material from the previous year and starting slowly with new material is so important:

> "Because it reviews all of the sounds, their letters, and sight words. So if we take our time and we don't rush and we don't try to overload them, they are just so much more

> successful. They're happier and it really sets a really good positive beginning for first grade as well."

Teacher A explains why teaching social skills should take precedence over the introduction of challenging academic work:

> "Less; the things that you know are not nearly as important as how you do it, so any job you go into as an adult, they're going to teach you how to do your job because they pretty much expect you to be able to get along with other people. If they don't know how to get along with others, it doesn't matter what you know because no one is going to work with you anyway, or you are not going to hold your job long enough for them to bother, so they have to have those social skills first, and then, you put the academics in place once the social skills are where they should be."

A critical goal of establishing a positive classroom environment is that students become more independent in their learning process and with how they interact with their peers. When students follow the rules and procedures of the classroom independently, it takes a lot of the burden of enforcement off the teachers, allowing them to focus on the academic experience of the students. Teachers establish strong classroom cultures by establishing structure, setting high expectations, and building positive relationships with students. Sometimes these activities take away from the introduction of new academic material which is a possible explanation for why growth rates are lower at the beginning of the year in reading fluency.

End of Year Instructional Changes

The end of year period is of particular importance to this follow-up chapter because the teachers selected for interviews were chosen, in part, because of average end of year growth for their students and because the finding that end of year growth is significantly greater than growth

during the rest of the school year runs counter to many of the explanations of teachers and schools for why they give important assessments several weeks before the last day of school. In the way many teachers talk about this period in terms of assessment timing, one would expect a sharp decline in growth rates, but in fact, the opposite was true. Interviewed teachers did report that the behavior of students changes for the worse as summer break approaches. However, teachers also adjust how they run their classrooms, including: maintaining high expectations, organizing more hands-on activities to keep students interested, moving from teacher-centered to student-centered activities, and synthesizing learning from throughout the school year. Teachers generally reported they had the freedom to teach the way they wanted during the school year, so it is important to understand why such strategies are employed at this time if they result in higher learning rates.

Importantly, all the teachers in the study described a significant behavior change in the students as the end of year approached. Students were less focused on their academic work and were more difficult to control than the rest of the school year. Teacher A identified the cause of the behavior change as excitement about activities during summer break:

> "Yes, yes, huge. One, by the time they're in first grade, they actually know what summer break is. Kindergarteners have never had a summer break. They don't know what that is yet. First graders have; they know exactly what's coming and that they'll be at daycare, on field trips or at grandma's or wherever they're going to be."

Teacher B finds the behavior changes as a building up of excitement as different activities occur throughout the year:

> "Yeah, it definitely gets more bizarre, definitely more movement going around, more excitement buzzing in the air. Obviously, I try and dial it down, you know, stating: 'We

> still have this to do, we can't get to this party or this, you know, planned activity if we can't get this done,' especially when, you know, like spring break's coming, with all the holidays we've had, Valentine's Day and St. Patrick's Day, now and then, kind of, field trip stuff coming up. It's just, it seems like there's so much building all year long that it's just, you know, they're kind of like a volcano ready to explode by the end of the year but, you know, again trying to make sure you have those expectations still in place that everyone's going to follow that way it's more manageable."

Teacher F reports the behavior changes in her classroom are caused more about apprehension and uncertainty about living environments during summer break:

> "The last month is definitely a struggle because some they won't have that structure at home, and so, they start kind of acting out, and it's getting warm outside, so they'd rather be outside playing, or maybe the night before they're out playing too late, so they didn't get enough sleep."

Teacher C attributes the behavior change to changes in the weather as well as the ending of assessments signaling the end of the year:

> "It's harder to teach them I would say and keep them focused and keep them with you. They just don't have as much of a desire to learn, I guess, as they do at the beginning of the year. It's just a harder time, I guess, because the weather is so nice. I know it's nice in August too, but everything is new at the beginning of the year. And, at the end of the year, it's just a different setting, and everybody is winding down. And, we have our assessments, and that's kind of at the end of the year turns into: a lot of assessments."

Connecting to what teachers reported for the beginning of year period, several teachers emphasized the importance of keeping expectations high even as student behavior changes.

Teacher A states, “The expectations are the same the first day as they are the last day. It’s the last day of school, it’s 10 minutes before dismissal; yeah, we still don’t throw things across the room. Yeah, there’s no slide in expectations in that last month of school.” Teacher D explains her motivation for keeping expectations high as “because I care about them.” The same teacher reported that she sets the tone in her classroom. If she demands high expectations, the students will deliver:

> “Well, number one, I think that at the end of the year, they're wild, and I feel like if I set the expectation, they know what they have to meet. If I’m more lackadaisical then they're going to become that way, you know? So, I feel that, yeah, the expectation is just as high; their behavior everything like those is the same as it is throughout the whole year.”

Teacher A reports she holds on to the expectations even more because she wants her students to be successful in the next grade level and beyond:

> “The last month of school I hold them to the expectations even tighter because you're getting ready to go to second grade, and the expectation is you need to be ready to be there so you have to show me that you're ready to be there, and after 130 days of school, I shouldn't have to be explaining not to throws things across the room, so there are fewer chances for misbehavior. If you're going to misbehave, there's less reminders, so we spend more time on academics cause there's less reminders. Cause, yeah, you knew that wasn’t going to happen. I just look at them and I just point my finger. I don't even have to say anything; they know why they’re going.”

Teacher A also describes the benefit of keeping expectations high:

> “No, except that I get more teaching done at the end of the year because I'm not spending so much time monitoring behaviors. They're monitoring their own behaviors; they’re

> monitoring each other's. Where they're telling so-and-so, 'hey that wasn't very nice' instead of having me come over and interrupt my lesson with my group to go over and monitor and stop this problem from happening and talk it through and get that sorted out and then come back. I'm not dealing with that as much, so since I hold those expectations so high, I have more time to teach, and I get less interruption and you can get more done when there's less interruptions over stuff that doesn't need to be interrupted for."

Similarly, Teacher C believes keeping expectations high prevents students from misbehaving and helps them stay focused:

> "No, they never change. I think it helps a lot because I think that, if you become very lax on your expectations, I think they start playing around and not doing what they're supposed to be doing."

Teacher B keeps the expectations high in order to help the students grow:

> "I think it helps because it gives them that challenge and that pressure, you know, setting high expectations for everybody all the time and continuing to make them grow."

Two teachers reported they relax their expectations at the end of the school year, but much of it is a matter of semantics. Many teachers reported changing how they teach in the last month of school which could be interpreted as a relaxation of expectations. Teacher E explained the students are pretty worn out by the end of the school year, so there is not much to be gained by nagging them:

> "I don't know. I feel like at that point, at the end of the year, the students are pretty checked out. So, I feel like me nagging at them for not pushing their chair in is something that, at that point, isn't a big deal."

One way that teachers changed their instruction is to plan more hands-on learning for students near the end of the school year. Teachers plan these activities in part to keep the students' attention and to keep them interested in learning through the last day of school. Teacher D uses hands-on activities to keep students motivated and to disguise learning through the last day of school:

> "Well, I always tell them, up until the last day, we're still working; we're still working; we're still working; we still have learning to do, but I try to make it so they don't notice the learning. Like last year, we did a camping theme the last week, so we made s'more's, so now, we're going to write the steps, so I try to kind of hide that in there but know that there has to be a little to keep class control, but those kinds of things. We had tents in here, but they were able to go and read in the tent, so they were still doing something academic, so I tried to hide that in there because I know that they're getting wild but they have to. Sitting at a desk, at that point, I feel it's kind of pointless."

When asked how planning more hands-on activities affects student learning, Teacher E responded:

> "It should affect it positively. Usually students like to do those activities. So, they're a little more engaged in the activities than tests, and worksheets, and stuff like that."

Teacher C describes how she uses hands-on activities in order to keep students focused:

> "Well, at the end of the school year, it's hard to keep everybody focused and involved, so I still like to have fun because I always like to have fun teaching, but I do more movement activities and use a lot more GoNoodle and things like that just because they are tired, and they're ready to be outside, and keeping them with you and focused is difficult."

Another way teachers adjust instruction is to plan more student-centered activities rather than teacher-centered activities. This notion connects back to a critical goal at the beginning of the year: building independence. Students are more capable of learning on their own or with peers, as the end of the year approaches, if their teachers have provided the structure and guided the students in the right direction. The students also mature greatly during the year, and this occurrence contributes to their independence. Consequently, students are allowed more free choice in how they learn and what they read. Teacher A describes the payoff of being very structured at the beginning of the year in order to develop independence later:

> "Well, they know the expectations. They know how to do it. We've built on it all year. It's not something that's just 'surprise, guess what we're going to do today,' and they know each other. We spent a lot of time getting to know each other as a team, getting to know each other as friends, seeing things from other people's point of view, talking to each other instead of coming and telling me, 'have you told them?' You can solve ninety percent of your problems without me. All you have to do is say, 'hey, did you know that that hurt my feelings,' and usually, they're like, 'oh sorry.' You don't need me for that. You got that now and most of the time, they're getting the hang of that by then."

Teacher B also describes the benefit of developing in the students the ability to solve their own problems independently:

> "When I ask them about blending, when I ask them about when you sound it out, it's like they've never practiced it a day in their life, so I'm hoping and maybe that is, you know, the difference that we've hit so hard at the beginning of the year that they're able to kind of independently help themselves grow because they have a lot more skill set."

Teacher C grants her students more independence because students have demonstrated the ability to learn without her:

> "I try to focus more on station time, the Daily Five, and group activities. I move more towards group activities, less towards teacher-centered activities or more of me directing things because they're more independent, and they can do more things without me. So, if anything, I've given them more independence, and they have more time to play games and go to stations versus direct instruction. My direct instruction time, I cut down on that, and I give them more freedom and more time to work in groups and do things with their friends."

Teacher B describes how she eventually expects her students to follow the procedures on their own after practicing with her:

> "Yes, because at the beginning of first grade, you're still kindergartners learning how to be a first grader. Midway through the year, you're full-fledged first graders, and sometimes, I start in January but usually February on, you're an almost second-grader. Here's what it looks like. We've already been through these procedures; we know what to do, so I would say as opposed to teacher, being like, directing procedures in the beginning of the year; at this time of the year, you guys need to show me how to do it. You've already practiced this. Show me what you've learned. So, more on your own things than, or I should say more student-led than teacher-directed at the beginning of the year."

Teacher E points out that students mature quite a bit through the school year:

> "Well, when I get them at the beginning of the year, they're like first graders. When I have them at the end of the year, they're, like, almost third graders. And, I feel like the

transition between second and third grade is a big jump, maturity-wise. Other than that, I don't know."

Teachers also utilized synthesis and review activities to review the learning during the school year. Synthesis activities, such as journaling and writing letters, help students to reflect and connect learning from the school year to different contexts. Review activities allow students to retain learning and be ready for the next grade level. Teacher C describes the importance of making connections in student learning:

> "Oh, yes. I think that's very important. I do that all the time. Anytime I can make a connection with something that they've learned before. I'm always making connections because that's how our brain works and finding those connections is so important, so they can learn it."

Four of the seven teachers actually accelerated academic learning in the last month of school in different ways. The overall goal was to make sure students are ready for the next grade level. Teacher A described pushing higher performing students in the last month of school:

> "No, well, I guess some. I start pushing the higher ones a little harder at the end of the year; the lower ones we're still plugging along right where we're plugging along just because they're not ready to go any faster, so they're probably getting, the higher ones are starting to get pushed harder just because they've mastered so many more skills, and they're ready to be pushed."

Teacher B described accelerating academics in order to ensure the students have been exposed to all the material they need to in order to be ready for the next school year:

> "Get it done; get it done; get it done. I feel like there's so much more that, at the end of the year, you're trying to get to. Not necessarily 'oh, I have to finish Wonders, I have to

> finish Envision,' just making sure that the standards that we're supposed to cover we've gotten; making sure that we've introduced things that they're going to need in second grade that way, you know, we can say, 'hey, we worked on it; this is how we focused on it in first grade.' That's why they are where they're at in second grade, and you know this, we talked about money and how, you know, we don't really teach it; it's not required anymore, standard-wise, but we still need to introduce it because second grade starts at coin collections. Well, if we don't spend even a little bit of time on it, you know, they've lost it, especially if they don't get it at home."

When asked if they have more freedom to teach the way they want to in the last month of school, teachers generally reported they typically had a lot of freedom throughout the school year. This is likely, in large part, because of the grade levels of the teachers and students. If the teachers were in a state tested grade level, they may not have been granted as much freedom. Teacher A reported she has a lot of freedom to experiment in part because she does not teach a tested grade level:

> "Yes, because since we're not doing state standardized testing, we move at their speed. Mr. Idacavage [principal] has been fantastic about saying, 'Ok, feel free to try something. If it doesn't work, it doesn't work. Figure out why it didn't work and change it.' He's not one who's going to hold our feet to the fire: 'oh my gosh, this kid doesn't know this skill, oh my gosh.' He would rather they have the social expectations mastered before the academic ones because the academics will come, but without the social expectations, it's not going to get you where you need to go in life, but he's very encouraging of us, letting us try new things, so by the end of the year, I know the kids better."

Teacher G also reports she does not feel restricted by the campus principal:

> "You know what, I don't feel restricted, and maybe, I don't know, maybe that comes from where I came from because I felt like we were pretty restricted when I was at David Brewer [Elementary School], but I don't feel restricted at any point in time. Several times, I've asked Craig [principal], 'would it be OK if?' But, he's never said no, so I don't feel restricted. I mean, I may feel like I have to justify it if somebody said, 'you guys were outside all day!' Well, this is what we were doing, but, you know, I don't feel restricted at all."

Teachers identified many important changes to instruction at the end of the school year in spite of the negative change in student behavior. This trend may help to account for the surprising finding that academic growth in reading fluency spiked at the end of the year.

Conclusion

The interviews revealed that the selected teachers had a strong preference for assessing students with the traditional testing schedule; thus, the result being more school time embedded within summer learning measurements. Furthermore, some teachers reported a belief that scores would not be a true reflection of the students' ability if the assessments were administered closer to the beginning or end of the school year. A common belief was that students needed time to get used to being back in the school environment when they returned from summer vacation. Teachers also believed behavior changes and lack of focus would adversely affect student performance right before the summer break.

At the beginning of the school year, teachers reported being more focused on establishing the classroom culture and relationship building rather than teaching new academic material. Therefore, this habit may explain, in part, why students did not demonstrate substantial learning gains compared to the rest of the school. However, establishing the structure of the classroom,

setting high expectations, and getting to know students may have paid dividends later in the form of a more easily managed classroom and increased learning rates. Teachers reported that spending extra time on non-academic activities at the beginning of the school year helped the entire school year be more successful.

Comparably, most teachers reported that the behavior of students changed significantly at the end of the year because they were excited or apprehensive about the coming summer vacation. Typically, students were less focused, less willing to follow directions, and more difficult to control. Hence, this information might suggest a significant drop in achievement during this time period, but the opposite was true. Teachers actually made substantive changes to instruction partly in response to the behavior changes. Important changes included more hands-on activities, academic acceleration for some students, and activities to synthesize learning. Correspondingly, many teachers worked hard to build independence in students during the school year resulting in more student-centered instead of teacher-centered activities in the last month. High expectations were also carried over from the rest of the school year by many successful teachers. The completion of stressful testing may also explain why some teachers change how they teach but less so for the interviewed teachers because their students do not take state assessments. These changes to instruction, regardless of how the students felt or behaved, could partially account for the phenomenon of learning acceleration.

Chapter VI- Conclusion

Introduction

This study considered the implications regarding how giving assessments to students on a schedule that does not overlap with the school district academic calendar indirectly affects estimates of summer learning loss. Moreover, findings regarding the effectiveness of summer programs designed to mitigate learning loss were also investigated. The review of literature revealed how the summer learning pattern suggests schools serving low-SES families are performing better than is popularly believed. This realization meant factors other than a student's in-school environment, such as family and neighborhoods, could play a role in creating eh achievement gap based on social class. Nevertheless, the evidence suggests schools could play a role in closing the achievement gap especially with the targeted deployment of resources specifically assigned to assist at-risk students. Summer academic programs are one promising strategy to accomplish this feat.

Most seasonal learning research potentially underestimates summer learning loss because students are not tested on the first and last days of school. In fact, many weeks of school time are often included in summer learning loss estimates. This prevalent issue was addressed by using curriculum-based measurement, a type of assessment relatively new to summer learning research. R-CBM, the specific assessment used for the study, only requires students to participate for one minute making it more practical to assess students closer to the first and last days of schools. This study compressed the amount of school time in summer learning estimates down to only a handful of days instead of several weeks. Having said that, under the traditional schedule of assessment, tests are given with several weeks of instruction included in the school year measurement, while under the minimized school influence schedule, tests are scheduled to

reduce the amount of school time included in summer learning estimates. The differences in the estimates, based on the two models, had a profound impact on summer learning estimates and summer program evaluations in the next section.

Underestimation of Summer Learning Measurements

The first research question considered how assessing students at different time periods affected the measurement of summer learning loss. Gain scores were analyzed based on assessments collected during the traditional schedule, the time period often favored by teachers that includes several weeks of school time in summer learning estimates, and a minimized school influence schedule, when assessments were given nearer to the first and last days of school. On average, a student's reading fluency score decreased by 5.564 words per minute during the summer break as measured by the traditional schedule compared to decreasing by 17.370 words per minute as measured by the minimized school influence schedule. This differences in the means proved to be statistically significant, t (334) = 10.6998, $p<.001$. These results coincide with the reasoning of Cooper, et al. (1996) that notes including a substantial amount of school time during assessment testing can dramatically underestimate summer learning loss measurements. In this case, the traditional schedule underestimated summer learning loss by 219% compared to the minimized school influence schedule. It is important to note this result is for all students, regardless of socioeconomic status, so there may be important differences in gain scores by groups that may be masked by this result.

The follow-up interviews revealed teachers have a strong preference for testing at periods consistent with the traditional schedule. School policies, at least in Carsonville CISD, also seem to favor this interval since the tests were originally scheduled during these time periods. Teachers reported they did not believe scores would be a true reflection of students' abilities

closer to summer break because of negative behaviors and a need to become familiar with the school environment. For this reason, the testing schedule in practice is unlikely to change for school districts, so researchers should carefully consider the underestimation of summer learning measurements when interpreting results from data collected with the traditional schedule.

Controlling for the Testing Interval Issue

The second research question considered if the learning rates during the first and last months of school were different from the remainder of the school year. This question is significant because of the efforts in previous research to control for the extra school time included in summer learning estimates. Controlling for the extra school time assumes a static growth rate for the school year. While the differences in growth rates during the school year may not be massive, it is worth considering if they are different. On average, a student's score increased by 0.135 words per minute each day during the school year and by 0.282 words per minute during the last month of school as measured by the traditional schedule. The difference in the means was statistically significant, t (381) = -4.999, p<.001. On average, a student's score increased by 0.138 words per minute each day during the school year compared to increasing by 0.154 words per minute each day during the first month of school. This difference in means was not statistically significant, t (350) = -0.415, p=0.678. Surprisingly, the last month gain rate was significantly different than the rest of the school year. These findings were not expected and so became the focus of the follow-up study described in the next few paragraphs. Interestingly, the results were somewhat at odds with the conclusions of Nese et al. (2012) that growth was greater during the first semester than the second semester. However, it is important to note the data used for Nese et. al's (2012) study followed the traditional schedule and did not account for the time periods just after school started and just before it ended in the spring. An important consideration

for this analysis is that the spring period included on average 31.914 days, while the fall period included only 22.144 days. The difference in the number of days may have contributed to the contrasting results of the significance tests.

Teachers identified in the follow-up interviews that, at the beginning of the year, they preferred to focus on cultivating a positive culture in their classrooms rather than introducing new academic material. Teachers spend a lot of time teaching procedures that include: how the day will flow, how students are expected to carry out regular activities, how students are expected to treat each other, and how students are expected to work together. Teachers also focus on teaching social skills and developing strong relationships with their students. Often, academic work was secondary, while classroom procedures and relationship-building took priority. Given this context, it would make complete sense that the students would not grow significantly in the first few weeks of school. Teachers devote this particular period to setting the classroom culture because, if they teach students what is expected of them and make them feel welcome, generally it has been known to yield gains in effort and behavior later in the year. Specifically, teachers noted that efforts to build independence helped at the end of the school year when students were able to learn on their own and follow the classroom procedures with less redirection. If the procedures are appropriately set from the beginning, students would be more attentive to each lesson taught throughout the year making the lessons appreciably more effective.

At the end of the school year, teachers reported a tremendous amount of frustration with changes in student behavior so it came as a surprise that the learning rates for this last month of school were so high in the quantitative analysis. However, after further analysis, many teachers tended to adjust their instruction but maintained their expectations during this period, potentially accounting for the increased growth rate. Teachers utilized different strategies in large part to

help keep their students' attention and keep them on track. On the surface, the behavior changes appeared to hurt academic achievement. Upon closer inspection, if teachers effectively channeled the energy of the students and adjusted instructional practices, it contributed to a deeper level of learning. A teacher who lowers his or her expectations may produce a detrimental slide in academic growth if he or she does not hold students to high standards. Additionally, if a teacher allows the behavior of the students to take control of the classroom, there is greater probability of academic progress faltering. Conversely, if a teacher is able to continue to enforce high standards, even as the behavior of the students change, the teacher is able to transform said energy into learning. Interestingly, the flow of the typical school day and instruction changed at then end of the academic year in such a way as to promote additional learning without the immediate recognition of the teachers or other adults.

Although this is more of a consideration for higher grade levels, high-stakes tests are typically completed at least a few weeks before school ends and might also help explain why instruction seems to change at the end of the school year. Many teachers may conventionally "teach to the test" which can artificially inflate the level of surface learning a student encounters right before an important assessment. However, this trend can inhibit instructional techniques that are designed to produce deeper, more sustained learning like those reported in this study. With the pressures of accountability lessened, the instruction during the last part of the school year may change in important ways. Some teachers choose to introduce innovative lessons that require students to synthesize and apply learning from the cumulative teaching up to that point in the academic year. Teachers might try to use these types of lessons during the school year but not be able to implement them fully because of testing accountability and the pressure to stick to a schedule of introducing testing-related concepts.

The results of this study have important implications for the efforts by researchers to control for testing interval differences when data is collected with the traditional schedule. This study provided evidence that students can experience different learning rates throughout the school year, particularly the last month. The late 1990's ECLS-K data included testing dates and utilized these figures as controls in their analysis (i.e. Burkam et al., 2004). However, using the dates in this way assumes a consistent growth rate through the school year (and summer). This study provides evidence indicative that the growth rates may be inconsistent. Efforts to control for the testing interval at least acknowledge this is an issue in summer learning research. Studies including these controls do a better job of accounting for the additional school time than studies that do not include them. As Cooper, et al. (1996) point out: not including the dates assumes the tests were given on the last day and first day of school. But researchers must consider in their interpretations that even these controls do not accurately account for the school time included in typical summer learning estimates.

Compensatory Effect of Schools

The third question considered the compensatory effect of schools for low-SES students. If time in school allows low-SES children to keep up academically with high-SES children or even grow at a greater rate, there is some evidence for a compensatory effect for schools. During the school year, on average, the score of a student qualifying for free/reduced lunch increased by 0.140 words per minute each day, and the score of a student not qualifying for free/reduced lunch increased by 0.177 words per minute each day. The difference in these means was statistically significant, $t(400) = 3.811$, $p<.001$. During the summer, on average, the score of a student qualifying for free/reduced lunch decreased by 0.044 words per minute each day, and the score of a student not qualifying for free/reduced lunch decreased by 0.027 words per minute each day.

The difference in these means was not statistically significant, t (369) = 1.480, p=0.140. The results of the multi-variate analysis supported the significance of the *t* test results. These results, in many ways, are not consistent with the results of the seminal works on summer learning. Alarmingly, the gain rate for low-SES students was significantly lower than high-SES students, indicating the achievement gap in reading fluency is growing during the school year. Interestingly, summer learning loss measurements were not significantly different along SES lines, which once again, is not consistent with previous research. These results provide evidence to support Quinn and Le's (2018) conclusions with more recent summer learning data that achievement gaps remained static or even widened during the school year. This shift possibly occurred because many of the equalizing effects of public schools might have been reversed in recent years, in the form of less equitable funding and re-segregation (Quinn & Le, 2018). If further research replicates this finding, this study magnifies a troubling trend for the public education system.

Effectiveness of Summer Boost

This dissertation study did not show that Summer Boost, a summer academic program for at-risk students, had a significant impact on reducing summer learning loss. However, this conclusion was likely the result of a small sample size (68 students were included in the treatment group). A propensity score matching model, utilizing logistic regression to calculate probability scores and nearest neighbor matching to construct the control group, was used to reduce selection bias. On average, the gain scores over the summer with the minimized school influence schedule for Summer Boost participants were 2.735 words per minute greater than control group members, $Mean_{treated}$ = -11.309, $Mean_{control}$ = -14.044, Standard Error = 2.903, z = 0.94, p = 0.346, $n_{treated}$ = 68. On average, the gain scores over the summer with the traditional

schedule for Summer Boost participants were 0.306 words per minute less than control group members, $Mean_{treated}$ = -3.467, $Mean_{control}$ = -3.160, Standard Error = 1.139, z = -0.27, p = 0.788, $n_{treated}$ = 75. The design of Summer Boost was guided by previous research in several areas including scheduling, academic focus and planning, total hours of programming, teacher recruitment practices, and transportation options. Cooper et al. (2000) and McEachin, et al. (2016) found that well-planned summer programs have a small-to-moderate impact on academic outcomes. The minimized school influence analysis for this study is consistent with this finding. The findings of this dissertation study highlight the difficulty past researchers have encountered with isolating specific design aspects of effective programs and the perspective that just having a summer program alone does not guarantee results.

There are also two important model design considerations. First, the number of Summer Boost participants was relatively small with 68 students in the treatment group. Two complete grade levels were included in the analysis to increase the number of participants. Even so, district selection policies ultimately determined the number of participants for these grade levels. With a larger sample size, the differences between Summer Boost participants and the control group measured by the minimized school influence schedule may have been statistically significant. Second, the model for calculating the propensity scores could have been strengthened with more covariates. This analysis utilized information readily available to the school district. However, more detailed information was not available. For example, socioeconomic status was represented by free/reduced lunch status which only measures family income. Socioeconomic status is a multifaceted concept and other important aspects of it, such as parental education and employment levels, were not included. Propensity score analysis is dependent on the assumption of ignorable treatment group assignment. The covariates should account for as many

characteristics independent of treatment as possible because the analysis assumes students with similar propensity scores have similar distributions of characteristics, observed or not (Rosenbaum & Rubin, 1983). While the covariates included were wide-ranging, the inclusion of additional covariates related to the outcome might have strengthened the model.

The difference between the analysis based on the minimized school influence schedule and the traditional schedule is also worth consideration because of the implications for program evaluation design. While neither result for questions four through five was statistically significant, the differences in the average treatment effect on the treated were important. Summer Boost participants had an average score almost three words per minute greater than control group members measured by the minimized school influence model. In contrast, using the traditional schedule, Summer Boost participants scored slightly less than control group members. Paired *t*-test results at the 95% confidence level demonstrate the gain scores for Summer Boost participants were on average 8.044 words per minute greater using the minimized school influence schedule compared to the traditional schedule, $Mean_{trad} = -3.265$, $Mean_{msi} = -11.309$, $t(67) = 3.398$, $p < .01$. The difference in school time included in the different models has important implications for measuring the pure effect of a summer intervention. The extra school time can mask the effectiveness of an intervention when specifically analyzing if participants experience less summer learning loss than non-participants. However, in practical terms, the traditional schedule measurement is still important for school officials because they are concerned with the overall academic growth of the child, including learning during the academic year and summer. The goal for a program seeking a long-lasting impact should be to affect student learning even after the school year has begun.

Recommendations for Future Research

The primary recommendation for future research is the replication of the methodology for this dissertation study because it utilized a unique model in summer learning research. Because students were assessed nearer to summer vacation, this study examined patterns in summer learning that could only be controlled for in previous research, not actually observed. The important findings also warrant further research to see if the results are replicated. The finding that low-SES students experienced significantly smaller gain rates during the school year than high-SES students deserves particular scrutiny because it runs counter to previous summer learning research results and may be indicating a pattern that is just beginning to emerge from the accountability era. The findings about the underestimation of summer learning loss and different learning rates at the end of the school year also need to be replicated because these results might have important implications for interpreting future summer learning research. This study utilized a type of assessment called R-CBM, which is not a common assessment in summer learning research. Consequently, it is important that results are replicated both using R-CBM and forms of assessment forms of assessment that have been used in past studies to continue to examine the relationship between R-CBM and more comprehensive assessments regarding summer learning research.

References

Alexander, K. L., & Condliffe, B. F. (2016). Summer setback in Baltimore: A review and update. In K. Alexander, S. Pitcock, & M. Boulay (Eds.), *The summer slide: What we know and can do about summer learning loss* (pp. 23-34). New York: Teachers College Press.

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2007a). Lasting consequences of the summer gap. *American Sociological Review*, 72(2), 167-180.

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2007b). Summer learning and its implications: Insights from the beginning school study. *New Directions for Youth Development*, 114, 11-32.

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2001). Schools, achievement, and inequality: A seasonal perspective. *Educational Evaluation and Policy Analysis*, 23(2), 171-191.

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2014). *The long shadow: Family background, disadvantaged urban youth, and the transition to adulthood.* New York: Russell Sage Foundation.

Allinder, R. M., & Fuchs, L. S. (1994). Alternative ways of analyzing effects of a short school break on students with and without disabilities. *School Psychology Quarterly*, 9(2), 145-160.

Atteberry, A. & McEachin, A. (2016). School's out: Summer learning across grade levels and and school contexts in the United States today. In K. Alexander, S. Pitcock, & M. Boulay (Eds.), *The summer slide: What we know and can do about summer learning loss* (pp. 35-54). New York: Teachers College Press.

Baker, S. K., & Good, R. (1995). Curriculum-based measurement of English reading with bilingual Hispanic students: A validation study with second-grade students. *School*

Psychology Review, 24(4), 561-578.

Borman, G. D., Benson, J., & Overman, L. T. (2005). Families, schools, and summer learning. *The Elementary School Journal*, 106(2), 131-150.

Bowles, S., & Gintis, H. (2002). Schooling in capitalist America revisited. *Sociology of Education*, 75(1), 1-18.

Bryk, A. S., & Raudenbush, S. W. (1988). Toward a more appropriate conceptualization of research on school effects: A three-level hierarchical linear model. *American Journal of Education*, 97(1), 65-108.

Burkam, D. T., Ready, D. D., Lee, V. E., & LoGerfo, L. F. (2004). Social-class differences in summer learning between kindergarten and first grade: Model specifications and estimation. *Sociology of Education*, 77(1), 1-31.

Chin, T., & Phillips, M. (2004). Social reproduction and child-rearing practices: Social class, children's agency, and the summer activity gap. *Sociology of Education*, 77(3), 185-210.

Condron, D. J. (2009). Social class, school and non-school environments, and Black/White inequalities in children's learning. *American Sociological Review*, 74(5), 683-708.

Condron, D. J., & Roscigno, V. J. (2003). Disparities within: Unequal spending and achievement in an urban school district. *Sociology of Education*, 76(1), 18-36.

Cooper, H., Charlton, K., Valentine, J. C., Muhlenbruck, L., & Borman G. D. (2000). Making the most of summer school: A meta-analytic and narrative review. *Monographs of the Society for Research in Child Development*, 65(1), 1-127.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, 66(3), 227-268.

Darby, D., & Rury, J. L. (2018). *The color of mind: Why the origins of the achievement gap matter for justice.* Chicago: The University of Chicago Press.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, 52, 219-232.

Deno, S. L. (1992). The nature and development of curriculum-based measurement. *Preventing School Failure: Alternative Education for Children and Youth*, 36(2), 5-10.

Deno, S. L. (2015). Problem-solving assessment. In R. Brown-Chidsey (Ed.), *Problem-Solving Based Assessment for Education Intervention* (2nd ed.) (pp.10-35). New York, NY: Guilford Press.

Deno, S. L., & Fuchs, L. S. (1987). Developing curriculum-based measurement systems for data-based special education problem solving. *Focus on Exceptional Children*, 19, 1-16.

Downey, D. B. (2016). Schools, families, and inequality: Strong empirical patterns in search of strong theory. In K. Alexander, S. Pitcock, & M. Boulay (Eds.), *The summer slide: What we know and can do about summer learning loss* (pp. 11-22). New York: Teachers College Press.

Downey, D. B., von Hippel, P. T., & Broh, B. A. (2004). Are schools the great equalizer? Cognitive inequality during the summer months and the school year. *American Sociological Review,* 69(5), 613-635.

Downey, D. B., von Hippel, P. T., & Hughes, M. (2008). Are "failing" schools really failing? Using seasonal comparison to evaluate school effectiveness. *Sociology of Education*, 81(3), 242-270.

Duffett, A., & Farkas, S., & Loveless, T. (2008). *High-achieving students in the era of No Child Left Behind.* Washington DC: Thomas B. Fordham Institute.

Entwisle, D. R., & Alexander, K. L. (1992). Summer setback: Race, poverty, school composition, and mathematics achievement in the first two years of school. *American Sociological Review*, 57(1), 72-84.

Farkas, G., & Beron, K. (2004). The detailed age trajectory of oral vocabulary knowledge: Differences by class and race. *Social Science Research*, 33(3), 464-497.

Fuchs, L. S., & Deno, S. L. (1991). Effects of curriculum within curriculum-based measurement. *Exceptional Children,* 58(3), 232-243.

Fuchs, L. S., & Fuchs, D. (1991). Curriculum-based measurements: Current applications and future directions. *Preventing School Failure*, 35(3), 6-11.

Fuchs, L., Fuchs, D., Hamlett, C., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review*, 22, 27-48.

Garrido, M. M., Kelley, A. S., Paris, J., Roza, K., Meier, D. E., Morrison, R. S., & Aldridge, M. D. (2014). Methods for constructing and assessing propensity scores. *Health Services Research*, 49(5), 1701-1720.

Gershenson, S. (2013). Do summer time-use gaps vary by socioeconomic status? *American Educational Research Journal*, 50(6), 1219-1248.

Godwin, R. K., & Kemerer, F. R. (2002). *School choice tradeoffs: Liberty, equality, and diversity.* Austin, TX: University of Texas Press.

Gold, K. M. (2002). *School's in: The history of summer education in American publc schools.* New York: P. Lang.

Hamiltion, C., & Shinn, M. R. (2003). Characteristics of word callers: An investigation of the

accuracy of teachers' judgments of reading comprehension and oral reading skills. *School Psychology Review*, 32(2), 228-240.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore: Paul H. Brookes.

Heyns, B. L. (1978). *Summer learning and the effects of schooling.* New York: Academic Press, Inc.

Kansas State Department of Education (2019). *Kansas Report Card 2017-18*. Retrieved February 6, 2019, from http://ksreportcard.ksde.org/home.aspx?org_no=D0453&rptType=2

Kim, J. S., & Quinn, D. M. (2013). The effects of summer reading on low-income children's literacy achievement from kindergarten to grade 8: A meta-analysis of classroom and home interventions. *Review of Educational Research,* 83(3), 386-431.

Lareau, A. (2011). *Unequal childhoods: Class, race, and family life* (2nd ed.). Berkeley: University of California Press.

Lee, V. E., & Burkam, D. T. (2002). *Inequality at the Starting Gate: Social background differences in achievement as children begin school.* Washington, D. C.: Economic Policy Institute.

McCombs, J. S., Pane, J. F., Augustine, C. H., Schwartz, H. L., Martorell, P., & Zakaras, L. (2014). *Ready for fall? Near-term effects of voluntary summer learning programs on low-income students' learning opportunities and outcomes.* Santa Monica, CA: RAND Corporation.

McEachin, A., Augustine, C., & McCombs, J. (2016). Best practices in summer programming. In K. Alexander, S. Pitcock, & M. Boulay (Eds.), *The summer slide: What we know and can do about summer learning loss* (pp. 193-207). New York: Teachers College Press.

National Center for Education Statistics (2019). *Common Core of Data.* Retrieved February 6,

2019, from https://nces.ed.gov/ccd/

Nese, J. T., Biancarosa, G., Anderson, D., Lai, C. F., & Tindal, G. (2012). Within-year oral reading fluency with CBM: A comparison of models. *Reading and Writing*, 25, 887-915.

Pallas, A. M. (2016). The sunny home advantage: Class differences in resources to support summer learning. In K. Alexander, S. Pitcock, & M. Boulay (Eds.), *The summer slide: What we know and can do about summer learning loss* (pp. 111-130). New York: Teachers College Press.

Phillips, M., & Chin, T. (2004). How families, children, and teachers contribute to summer learning and loss. In G. D. Borman (Eds.), *Summer learning: Research, policies, and programs* (pp. 255-278). Mahwah, NJ: Lawrence Erlbaum Associates.

Pitcock, S. (2016). School's out: Summer learning and the opportunity gap. In K. Alexander, S. Pitcock, & M. Boulay (Eds.), *The summer slide: What we know and can do about summer learning loss* (pp. 70-89). New York: Teachers College Press.

Quinn, D. M. (2015). Black-White summer learning gaps: Interpreting the variability of estimates across representations. *Educational Evaluation and Policy Analysis*, 37(1), 50-69.

Quinn, D. M., & Le, Q. T. (2018). Are we trending to more or less between-group achievement inequality over the school year and summer? Comparing across ecls-k cohorts. *AERA Open,* 4(4), 1-19.

Reardon, S. F., Valentino, R. A., & Shores, K. A. (2012). Patterns of literacy among U.S. students. *The Future of Children*, 22(2), 17-37.

Reschly, A. L. Busch, T. W. Betts, J., Deno, S. L., & Long, J. (2009). Curriculum-based

measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology,* 47(6), 427-469.

Rosenblatt, M. L. (2002). The effects of summer vacation on children's reading performance: An examination of retention and recoupment using curriculum-based measurement. (Unpublished master's thesis). Syracuse University, Syracuse, NY.

Sandberg Patton, K. L., & Reschly, A. L. (2013). Using curriculum-based measurement to examine summer learning loss. *Psychology in the Schools,* 50(7), 738-753.

Stein, M. L. & Fonseca, E. (2016). A review and analysis of evaluations of summer programs. In K. Alexander, S. Pitcock, & M. Boulay (Eds.), *The summer slide: What we know and can do about summer learning loss* (pp. 271-283). New York: Teachers College Press.

Stone, C. A. & Tang, Y. (2013). Comparing propensity score methods in balancing covariates and recovering impact in small sample educational program evaluations. *Practical Assessment, Research & Evaluation*, 18(13), 1-12.

Wayman, M. M., Wallace, T., Wiley, H. I., Ticha, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education*, 41(2), 85-120.

Wilson, W. J. (2009). *More than just race: Being Black and poor in the inner city (issues of our Time).* New York, NY: W. W. Norton.

Wiseman, A. W., & Baker, D. P. (2004). The American summer learning gap from an international perspective. In G. D. Borman (Eds.), *Summer learning: Research, policies, and programs* (pp. 255-278). Mahwah, NJ: Lawrence Erlbaum Associates.

Yeo, S. (2010). Predicting performance on state achievement tests using curriculum-based measurement in reading: A multilevel meta-analysis. *Remedial and Special Education*,

31(6), 412-422.

Appendix A- Descriptive Charts for Carsonville CISD

Table 14: Pre-Kindergarten Through Twelfth Grade Enrollment by Grade for Carsonville CISD (2016-17)

Grade	Number of Students
Pre-Kindergarten	151
Kindergarten	285
Grade 1	288
Grade 2	264
Grade 3	275
Grade 4	268
Grade 5	251
Grade 6	252
Grade 7	250
Grade 8	205
Grade 9	345
Grade 10	335
Grade 11	363
Grade 12	338

Source: National Center for Education Statistics, 2019

Table 15: Carsonville CISD Enrollment by Race/Ethnicity (2016-17)

Category	Number of Students
American Indian/Alaska Native Students	10
Asian or Asian/Pacific Islander Students	44
Hispanic Students	434
Black Students	641
White Students	2,248
Hawaiian Nat./Pacific Isl. Students	11
Two or More Races Students	482

Source: National Center for Education Statistics, 2019

Table 16: Carsonville CISD Enrollment by Grade and Race/Ethnicity (2016-17)

Grade Level	American Indian/ Alaska Native	Asian or Asian/ Pacific Islander	Hispanic	Black	White	Hawaiian Nat./ Pacific Isl.	Two or More Races
Prekindergarten	1	0	21	13	84	0	32
Kindergarten	2	2	29	52	166	0	34
Grade 1	1	2	34	50	152	0	48
Grade 2	1	1	27	47	149	0	39
Grade 3	0	1	35	59	136	1	43
Grade 4	2	4	19	46	153	1	43
Grade 5	0	3	34	35	138	1	40
Grade 6	1	3	29	51	137	1	30
Grade 7	0	2	29	47	149	0	23
Grade 8	1	1	18	31	129	1	24
Grade 9	0	6	36	61	209	0	33
Grade 10	0	6	36	44	225	1	23
Grade 11	0	8	49	54	208	2	42
Grade 12	1	5	38	51	213	2	28

Source: National Center for Education Statistics, 2019

Table 17: Carsonville CISD Enrollment by Grade Level and Gender (2016-17)

Grade Level	Male	Female
Prekindergarten	89	62
Kindergarten	148	137
Grade 1	162	126
Grade 2	137	127
Grade 3	162	113
Grade 4	147	121
Grade 5	137	114
Grade 6	123	129
Grade 7	132	118
Grade 8	100	105
Grade 9	181	164
Grade 10	178	157
Grade 11	194	169
Grade 12	183	155

Source: National Center for Education Statistics, 2019

Table 18: Performance in Mathematics for State Assessment- District and State Averages

Group- Year	Level 1	Level 2	Level 3	Level 4	Not Tested
District- 2017	41.3%	36.66%	15.9%	4.58%	1.54%
District- 2018	46.43%	34.51%	15.12%	3.92%	0
State- 2017	27.77%	37.5%	24.07%	9.21%	1.42%
State- 2018	29.23%	37.89%	23.71%	9.14%	0

Source: Kansas State Department of Education, 2019

Table 19: Performance in English/Language Arts for State Assessment- District and State Averages

Group- Year	Level 1	Level 2	Level 3	Level 4	Not Tested
District- 2017	44.46%	27.6%	21.15%	4.62%	2.14%
District- 2018	43.84%	29.45%	21.23%	5.46%	0
State- 2017	27.37%	33.73%	28.12%	9.08%	1.68%
State- 2018	29.16%	33.96%	28.2%	8.66%	0

Source: Kansas State Department of Education, 2019

Table 20: Percentage of Male and Female Students Enrolled in Carsonville CISD

School Year	Male	Female
2014	53.5%	46.5%
2015	53.5%	46.5%
2016	53.4%	46.6%
2017	53.6%	46.4%
2018	53.4%	46.6%

Source: Kansas Department of Education, 2019

Table 21: Percentage of Students Enrolled by Race/Ethnicity in Carsonville CISD

School Year	White	African-American	Hispanic	Other
2014	59.7%	18.7%	9.5%	12.1%
2015	59.4%	19%	9.7%	12%
2016	57.4%	18.2%	10.3%	14.1%
2017	58.1%	16.6%	11.2%	14.1%
2018	58%	15.8%	11.6%	14.6%

Source: Kansas Department of Education, 2019

Table 22: Percentage of Students Enrolled by Economic Disadvantage in Carsonville CISD

School Year	Econ. Disadvantaged	Non-Econ. Disadvantaged
2014	61.3%	38.7%
2015	63.45%	36.55%
2016	63.21%	36.79%
2017	58.4%	41.6%
2018	58.07%	41.93%

Source: Kansas Department of Education, 2019

Table 23: Percentage of Students Enrolled by English Proficiency in Carsonville CISD

School Year	ELL	Non-ELL
2014	2.43%	97.57%
2015	2.06%	97.94%
2016	2.19%	97.81%
2017	2.46%	97.54%
2018	1.93%	98.07%

Source: Kansas Department of Education, 2019

Table 24: Percentage of Students With/Without Disabilities Enrolled in Carsonville CISD

School Year	Students with Disabilities	Students without Disabilities
2014	18.28%	81.72%
2015	17.98%	82.02%
2016	17.26%	82.74%
2017	16.62%	83.38%
2018	17.17%	82.83%

Source: Kansas Department of Education, 2019

Appendix B- Question #3 Linear Regressions

This additional analysis was completed to further explore the relationship between free/reduced lunch status and seasonal growth rates (school-year and summer) and to see how additional demographic variables including race/ethnicity and gender affect that relationship. Multi-variate linear regression was used to analyze the relationships between the variables. For the school-year period, Model 1 includes growth rate as the dependent variable and free/reduced lunch status as the independent variable. Model 2 adds control variables for race/ethnicity and gender and Model 3 adds interaction variables. Table 25 shows the results for this analysis.

Table 25: Analysis of School-Year Growth Rate (Fall 2016 to Spring #1)

Variable	Model 1			Model 2			Model 3		
	Coef.	**SE**	***t***	**Coef.**	**SE**	***t***	**Coef.**	**SE**	***t***
Constant	0.172	0.007	24.96***	0.165	0.008	20.07***	0.162	0.011	15.05***
Lunch Status	-0.034	0.008	-4.02***	-0.035	0.009	-4.00***	-0.031	0.014	-2.26*
African American				-0.016	0.011	-1.38	0.031	0.025	1.23
Hispanic				0.004	0.013	0.27	0.002	0.025	1.23
Other				0.019	0.011	1.66	0.029	0.022	1.33
Gender				0.013	0.008	1.58	0.008	0.015	0.56
Lunch Status X African American							-0.039	0.027	-1.45
Lunch Status X Hispanic							-0.002	0.028	-0.08
Lunch Status X Other							-0.023	0.025	-0.92
Lunch Status X Gender							0.013	0.018	0.74
Gender X African American							-0.032	0.023	-1.39
Gender X Hispanic							0.008	0.028	0.28
Gender X Other							0.009	0.023	0.39
Adjusted R-Squared	*0.039*			*0.039*			*0.038*		

t test significance*:* *** p<.001 ** p<.01 * p<.05

The free/reduced lunch status variable was statistically significant for all three models. The race/ethnicity and gender variables did not prove to be statistically significant. For Model 1, the oral reading fluency score for students qualifying for free/reduced lunch on average decreased by 0.034 words per minute each day compared to students not qualifying for

free/reduced lunch. For Model 2, the oral reading fluency score for students qualifying for free/reduced lunch on average decreased by 0.035 words per minute each day compared to students not qualifying for free/reduced lunch after controlling for race/ethnicity and gender. For Model 3, the oral reading fluency score for students qualifying for free/reduced lunch on average decreased by 0.031 words per minute each day compared to students not qualifying for free/reduced lunch after controlling for race/ethnicity, gender, the interaction between lunch status and race/ethnicity, the interaction between lunch status and gender, and the interaction between gender and race/ethnicity.

The analysis for the summer period was set up similarly to the analysis for the school-year period, with control variables added to Model 2 and interaction variables added to Model 3. Table 26 shows the results of this analysis. Like the results with the *t* test in Chapter IV, free/reduced lunch status did not prove to be statistically significant during this period. However, the interaction between free/reduced lunch status and gender was statistically significant in Model 3. For Model 1, the oral reading fluency score for students qualifying for free/reduced lunch on average decreased by 0.016 words per minute each day compared to students not qualifying for free/reduced lunch. For Model 2, the oral reading fluency score for students qualifying for free/reduced lunch on average decreased by 0.014 words per minute each day compared to students not qualifying for free/reduced lunch after controlling for race/ethnicity and gender. For Model 3, the oral reading fluency score for students qualifying for free/reduced lunch on average decreased by 0.005 words per minute each day compared to students not qualifying for free/reduced lunch after controlling for race/ethnicity, gender, the interaction between lunch status and race/ethnicity, the interaction between lunch status and gender, and the interaction between gender and race/ethnicity. All the other control and interaction variables in

Model 3 were not statistically significant except the interaction between free/reduced lunch status and gender. The oral reading fluency score for a female student qualifying for free/reduced lunch status decreased by 0.044 words per minute each day compared to male students not qualifying.

Table 26: Analysis of Summer Growth Rate (Spring #1 to Fall #2)

Variable	Model 1			Model 2			Model 3		
	Coef.	**SE**	***T***	**Coef.**	**SE**	***t***	**Coef.**	**SE**	***t***
Constant	-0.025	0.008	-3.08**	-0.025	0.010	-2.47*	-0.025	0.013	-1.90
Lunch Status	-0.016	0.010	-1.60	-0.014	0.011	-1.33	-0.005	0.017	-0.27
African American				-0.021	0.014	-1.45	-0.052	0.031	-1.64
Hispanic				0.008	0.016	0.49	-0.004	0.028	-0.13
Other				-0.007	0.013	-0.54	-0.048	0.027	-1.78
Gender				0.003	0.010	0.26	0.019	0.018	1.02
Lunch Status X African American							0.034	0.034	1.00
Lunch Status X Hispanic							0.006	0.033	0.18
Lunch Status X Other							0.032	0.029	1.11
Lunch Status X Gender							-0.044	0.021	-2.08*
Gender X African American							0.016	0.029	0.56
Gender X Hispanic							0.021	0.032	0.66
Gender X Other							0.040	0.027	1.51
Adjusted R-Squared	*0.004*			*0.001*			*0.004*		

t test significance*:* *** p<.001 ** p<.01 * p<.05

Appendix C: Letter of Consent for Teacher Interviews

February 13, 2018

Dear Educator:

Summer vacation can have an interesting effect on the learning of students. Some students may thrive during the summer and come back to school better prepared for academic success. Other students may regress and require significant remediation when they return to school. I am completing my dissertation research and I have found some interesting patterns in learning rates as students go into summer vacation and return to school in the fall. I am interested in hearing your perspective as a classroom teacher about how you may treat the first and last months of the school year differently than the rest of the year.

Would you be willing to participate in an interview lasting approximately thirty minutes to one hour? There are no right or wrong answers to my questions and your insights would be most appreciated. Participation in the interview is voluntary and you are free to withdraw at any time. Your decision on whether or not you participate will in no way affect your relationship with your school or the University of Kansas. I anticipate there are no risks associated with your participation in the interview. Responses you share will be treated confidentially. Your identifiable information will not be shared unless (a) it is required by law or university policy, or (b) you give written permission. With your permission, I will audio record and transcribe the interview. If you request, I will turn off the audio recorder at any time. All information collected will be stored securely and the audio recording will be destroyed within one year of the interview.

I believe this research will be beneficial to schools that serve at risk students. It will help to shine a spotlight on the positive effect of public schools as well as provide guidance on changes schools can make in their programming to better serve at risk students. I plan to share the results with the school district as well as including it in my dissertation.

Thank you very much for your consideration. If you have any questions, feel free to contact me at (913) 240-5190 or b455w342@ku.edu. You may also contact my faculty advisor, Dr. John Rury, at (785) 864-9697 or jrury@ku.edu. If you have any additional questions about your rights as a research participant, you may call (785) 864-7429 or (785) 864-7385, write the Human Research Protection Program (HRPP), University of Kansas, 2385 Irving Hill Road, Lawrence, Kansas 66045-7568, or email irb@ku.edu. Please sign below if you agree to participate in the research study.

Sincerely,

Bryan Walker

_____
Participant's Printed Name

_____ _____
Participant's Signature Date

Appendix D: Teacher Interview Protocol

Beginning of School Year

How do you treat the first few weeks of school differently than the rest of the school year in your classroom?

Do you spend more or less time on academics at the beginning of the school year?

Do you do any special activities in order to set the culture of your classroom at the beginning of the year? Please provide me with some examples.

Do you prefer to wait a few weeks before doing any assessments with students that will have any consequence? Why?

End of School Year

How do you treat the last month of school differently than the rest of the school year?

When do you prefer to assess students in order to most effectively measure student learning most accurately? Several weeks before the end of school or closer to the end of school? Why?

Do your expectations change as the end of the school year approaches? Do you think this helps or hurts student learning? Why?

How does the behavior of the children change in the last month of school? Do you think this helps or hurts student learning? Why?

Do you change how you teach as the end of the school year approaches? How? Why? How does it affect student learning? Do you feel like you have more freedom? Do you do any activities to synthesize learning?

(Share teacher-specific data with interviewee) Do you have any other insights as to why this pattern occurred in your classroom?

Appendix E: IRB Approvals

THE UNIVERSITY OF KANSAS
Research

APPROVAL OF PROTOCOL

May 18, 2017

Bryan Walker
b455w342@ku.edu

Dear Bryan Walker:

On 5/18/2017, the IRB reviewed the following submission:

Type of Review:	Initial Study
Title of Study:	Improving the Evaluation of Summer Interventions: How Testing Intervals Affect Summer Learning Loss Measurements
Investigator:	Bryan Walker
IRB ID:	STUDY00140939
Funding:	None
Grant ID:	None
Documents Reviewed:	• letter_superintendent.pdf, • Walker_SummerLearning_HumanSubjects.pdf

The IRB approved the study on 5/18/2017.

1. Notify HRPP about any new investigators not named in original application. Note that new investigators must take the online tutorial at https://rgs.drupal.ku.edu/human_subjects_compliance_training.
2. Any injury to a subject because of the research procedure must be reported immediately.
3. When signed consent documents are required, the primary investigator must retain the signed consent documents for at least three years past completion of the research activity.

Continuing review is not required for this project, however you are required to report any significant changes to the protocol prior to altering the project.

Please note university data security and handling requirements for your project:
https://documents.ku.edu/policies/IT/DataClassificationandHandlingProceduresGuide.htm

You must use the final, watermarked version of the consent form, available under the "Documents" tab in eCompliance.

Sincerely,

Jocelyn Isley, MS, CIP
Interim IRB Administrator, KU Lawrence Campus

Human Research Protection Program
Youngberg Hall | 2385 Irving Hill Rd | Lawrence, KS 66045 | (785) 864-7429 | research.ku.edu/hrpp

KU THE UNIVERSITY OF KANSAS Research

APPROVAL OF PROTOCOL

February 23, 2018

Bryan Walker
b455w342@ku.edu

Dear Bryan Walker:

On 2/23/2018, the IRB reviewed the following submission:

Type of Review:	Modification
Title of Study:	Improving the Evaluation of Summer Interventions: How Testing Intervals Affect Summer Learning Loss Measurements
Investigator:	Bryan Walker
IRB ID:	STUDY00140939
Funding:	None
Grant ID:	None
Documents Reviewed:	• Consent Letter for Follow-up Interviews, • Consent Letter for Follow-up Interviews, • Protocol for Follow-up Interviews, • Protocol for Follow-up Interviews

The IRB approved the study on 2/23/2018.

1. Notify HRPP about any new investigators not named in original application. Note that new investigators must take the online tutorial at https://rgs.drupal.ku.edu/human_subjects_compliance_training.
2. Any injury to a subject because of the research procedure must be reported immediately.
3. When signed consent documents are required, the primary investigator must retain the signed consent documents for at least three years past completion of the research activity.

Continuing review is not required for this project, however you are required to report any significant changes to the protocol prior to altering the project.

Please note university data security and handling requirements for your project:
https://documents.ku.edu/policies/IT/DataClassificationandHandlingProceduresGuide.htm

You must use the final, watermarked version of the consent form, available under the "Documents" tab in eCompliance.

Sincerely,

Jocelyn Isley, MS, CIP
IRB Administrator, KU Lawrence Campus

Human Research Protection Program
Youngberg Hall | 2385 Irving Hill Rd | Lawrence, KS 66045 | (785) 864-7429 | research.ku.edu/hrpp