# ARE YEAR-ROUND SCHOOLS A VIABLE OPTION FOR IMPROVING STUDENT ACHIEVEMENT, COMBATING SUMMER LEARNING LOSS IN DISADVANTAGED YOUTH, CONTROLLING EXPENSES, AND REDUCING TEACHER BURNOUT?

Adrienne Smith

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of Education.

Chapel Hill
2011

Approved by:

Chair: Dr. Gregory J. Cizek

Advisor: Dr. Gary T. Henry

Reader: Dr. Kathleen M. Brown

Reader: Dr. Harris Cooper

Reader: Dr. William B. Ware

© 2011
Adrienne Smith
ALL RIGHTS RESERVED

**ABSTRACT**

ADRIENNE SMITH: Are Year-Round Schools a Viable Option for Improving Student Achievement, Combating Summer Learning Loss in Disadvantaged Youth, Controlling Expenses, and Reducing Teacher Burnout?
(Under the direction of Gary T. Henry)

Interest in year-round schooling is motivated by international comparisons of time spent in-school and efforts by policymakers to identify viable policy avenues for improving achievement and reducing costs. Prior research on effectiveness of modified year-round schools finds modest support for a modified year-round calendar, but much of the research is weak. Both memory and time-on-task literatures provide a framework for understanding how patterns of schooling and non-schooling intervals could impact student learning. While there is some evidence that the learning losses from summer breaks are greatest for students of low socio-economic status, there are few inquiries into the effects of year round schooling on these students, or other important student subgroups such as English language learners and students in special education. Using an extensive micro-level longitudinal database I compare the achievement of students under a traditional or modified year-round calendars. Capitalizing on a natural experiment in Wake County, NC wherein schools were switched from a traditional to a year-round calendar, I apply a student fixed effects method to isolate the effect of calendar arrangement on student achievement and student absenteeism. To complement the student fixed effects analysis and to increase the study's external validity, I use a growth curve analysis to compare outcomes for students attending a modified year-round calendar to students attending similar schools operating under a traditional calendar. In addition, I examine whether the modified year-round calendar is advantageous for increasing retention and reducing costs.

There were five major findings in this dissertation. First, the modified year-round calendar leads to improved student achievement for students of low socio-economic status and second, the modified year-round calendar is also beneficial for students with special needs. Third, the modified year-round calendar is detrimental to student performance for students who are English language learners. Fourth, the link between the modified year-round calendar and lower rates of student absenteeism is supported in the student fixed effects methodology. Fifth, higher teacher retention rates are correlated with a change from a traditional calendar to a modified year-round calendar. Future research efforts are suggested including an investigation of potential mediators for the modified year-round calendar effect.

## ACKNOWLEDGEMENTS

Upon completion of this work I must credit the individuals who have been instrumental in its development. Special thanks to Dr. Henry for his expertise, Dr. Cizek for his guidance in this process, and to all of my committee members, Dr. Brown, Dr. Cooper, and Dr. Ware, for their valuable input.

I am also indebted to my fabulous parents, who always shine the light; Jeannie, my music motivator; "Dream then do" Tim; my superbly classy friends; my husband Eric, whose love sustains me; and everyone who helped to remind me that I am so much more than the author of a really long paper.

**TABLE OF CONTENTS**

## LIST OF TABLES

Table

## LIST OF FIGURES

Figure

# CHAPTER 1

## INTRODUCTION

Blue ribbon commissions in the U.S. have acknowledged that the organization and amount of time spent on instruction can influence learning (Berliner, 1990; National Commission on Time and Learning, 1994). Spurred by international comparisons, concern about the traditional US calendar has grown since the mid 1990s (Berliner, 1990; National Education Commission on Time and Learning, 1994). School districts have experimented with several types of calendar reform including adding days to the school calendar, extending the school day, and modifying the distribution of the 180 days in the standard school calendar (Beaton et al., 1996). Extending the amount of time students spend in school is likely to add additional costs to compensate teachers for additional hours of work and other costs of school operations. In times when schools are pressured by taxpayer groups and elected officials to operate with limited budgets, calendar reforms such as modifying the school calendar emerge as a potential option. Calendar modifications offer three potential benefits: raising student achievement, improving teacher morale, and containing costs. Relationships between learning time and student performance are addressed in memory and time-on-task research. Both literature strains support the idea that calendar reform may influence student achievement. Within calendar reform options, modifying the distribution of the 180 schooling days is an economical choice because instead of requiring additional time in school, the modified calendar rearranges the 180 schooling days to extend beyond the August to May traditional schooling time span. Teacher retention is another way to gauge whether calendar reform could gain public support to become a viable option for reforming schools.

The number of modified calendar year-round schools has been steadily increasing since the 1980s. According to the latest figures just over two million children in America are enrolled in modified year-round schools (National Association for Year-Round Education, 2007). There were 3,000 modified calendar year-round schools (including schools from kindergarten to twelfth grade) during the 2006-2007 school year. Students enrolled in a modified calendar year-round school is up 11% from the 2000-2001 school year and up 39% from ten years ago. Modified calendars in particular are adopted more readily by elementary schools, as opposed to high schools, because of the scheduling conflicts that arise with after-school activities, especially athletics (St. Gerard, 2007). States with the highest number of modified year-round schools are California, Hawaii, Arizona, Nevada, and Texas; which are also states experiencing high rates of population growth (Cooper, Valentine, Charlton, & Melson, 2003).

Modified year-round schooling calendars are commonly arranged by nine weeks of schooling followed by a three week break. The modified year-round calendar eliminates the long summer break. If the time-learning relationship is not only affected by the amount of schooling time, but also the length of time between schooling periods, the modified calendar may have an advantage over the traditional calendar. Indeed, research has quantified the summer learning loss phenomenon, which is even more striking for economically disadvantaged students (Alexander, Entwistle, & Olsen, 2001; Heyns, 1978). Eliminating the long summer break may also be beneficial for other sensitive subgroups unexamined by previous research, such as English Language Learners and students in special education.

Several literature strands may inform calendar directives in multiple ways:

1. An extensive memory literature tradition, with roots in Ebbinghaus's experiments in the late nineteenth century, uncovers nuances between learning time and memory (Dempster, 1989). Conceptualizing time in a concrete manner, researchers in this tradition focus on the spacing of learning activities and subsequent memory tests to see if certain patterns of study and varying intervals of time passage increase retention.

2. From Carroll's work springs another line of research focusing on unpacking time into separate components including student motivation and innate ability, commonly labeled "time on task" (Carroll, 1989).

Memory studies and the time-on-task literature give insight into the effective use and placement of schooling and non-schooling time periods. In other words, both theoretical frameworks speak to calendar reform.

Yet, improving student achievement is not the only factor in the decision to initiate calendar reform. Calendar reform can be economical. Interestingly, financial concerns motivated the adoption of a particular kind of modified calendar, referred to as a multi-track year-round calendar, in growing regions. To handle an influx of students and the lack of time and money for building new schools, policymakers have turned their attention toward multi-track modified year-round schools as a feasible cost-savings option. Multi-track schools contain several groups of students who follow different schedules. Multi-track schools have a group of students on break at any given time. Each track has its own schedule. For example, a school on a multi-track modified calendar has 4 classrooms designated for third graders. There are 5 classes of third graders that rotate through the four classrooms. Every week one class is on a break (referred to as "tracked-out") while the other four classes are in session. Once a track resumes instruction, they occupy the classroom where previous students who are now "tracked out" once resided. Thus the school needs fewer classrooms and fewer resources (books, desks, chairs, storage cabinets, etc.) to serve more students. Few studies have quantified the effects of multi-track modified calendar schools relative to the traditional calendar (Daneshvary & Clauretie, 2001).

Calendar reform, however easy to manipulate in terms of a policy intervention, experiences the greatest challenge from cultural systems which are slow and oftentimes resistant to change. The standard calendar with 9 months of schooling followed by a three month break was a compromise between urban and rural interests in the early 1900s and is a central structure upon which families and society organizes their time, including school personnel. The satisfaction of school personnel is a

critical component for garnering support for any school reform. Distinct patterns of teacher turnover can reveal a great deal about personnel satisfaction. Higher levels of teacher turnover can indicate less satisfaction with a reform, in this case, modified school calendars. On the other hand, more frequent breaks in the modified year-round calendar could prevent teacher burnout and lead to a higher rate of teacher retention. To date, no evidence exists to support or reject the idea that teacher turnover may be lower in modified year-round calendar schools compared to traditional calendar schools.

The main focus of the proposed study is to investigate whether the modified year-round calendar improves student achievement outcomes, including the outcomes of groups of disadvantaged students, above and beyond achievement outcomes associated with a traditional calendar. Additional investigations of the school calendar's effect on school finance and teacher retention will also render evidence as to the potential viability of the modified year-round calendar reform. I hypothesize that, on average, students in modified year-round calendars will outperform students on a traditional calendar, and that the differences will be greater for students of low socioeconomic status, English Language Learners, and students in special education. Also, I predict that schools operating under a modified year-round calendar will spend less per pupil than schools under a traditional calendar. Finally, I hypothesize that the modified year-round calendar will reduce teacher burnout, resulting in lower teacher turnover rates compared to schools with a traditional calendar.

To study the effects of modified year-round schools will require three separateapproaches. First a comparison of modified year-round and similar traditional schools over a three year period will yield information on the growth curves of students from contrasting calendar arrangements. The lack of methodological and statistical rigor in the literature on modified year-round schooling is only beginning to be addressed by newer studies rooted in Ruben's causal framework; capitalizing on multi-level modeling strategies that enable researchers to test between school effects while controlling for individual student and teacher differences. Examining four-year growth curves will give information about the achievement trajectories for modified year-round and traditional students

and speaks to the different trajectories that may be observed in the various subgroups (i.e. students with differing levels of socioeconomic status).

Yet, the growth curve approach still suffers from bias, namely selection bias. If students who attend a modified year-round school are systematically different from students who are enrolled in a traditional calendar school in ways that are not measured by control variables such as prior test scores, race/ethnicity, or poverty status then these unobserved differences would confound the calendar effect. For example, perhaps parental motivation regarding their children's success in school influences the decision to reside in a high growth suburban community with high performing schools, which increases the chance their children will enroll in a year-round school. Differences in student achievement may be due to parental factors (when they are uncontrolled for in a study) instead of school calendar arrangement. Conveniently, a natural experiment in North Carolina allows for a second exploration to estimate the effect of the modified year-round calendar with little chance of selection bias. In the 2007-08 school year, Wake County Public Schools switched 23 elementary and middle schools from a traditional calendar to a modified year-round calendar just before the school year began, leaving parents with little opportunity to move their children to a school with a traditional calendar. Using a fixed effects design, students serve as their own controls, and statistical tests contrast the effect of traditional and modified school calendars on individual students. However, to consider the impact estimates unbiased, we have to assume that year-to-year growth is constant over from grade to grade.

Teacher morale, measured by differences in turnover, will be addressed using schools in Wake County with a difference-in-differences design. The difference-in-differences design will allow a comparison of teacher turnover rates prior to switch to a modified year-round calendar, adjusting for the differences due to the passage of time.

Expenditure differences will be examined similarly, where the variables of interest consist of several school-level expenditure categories related to resources (both material and human resources) that may be affected by calendar reform: regular instruction, special instruction, student services,

instructional support, school maintenance and utilities, school leadership, and capital outlay. Note that multi-track modified schools will be the focus of this analysis. Using difference-in-differences, the effect of calendar arrangement on operational expenditures can be measured. Differences in per pupil spending in various expenditure categories related to calendar type will inform policymakers as to whether calendar reform is an efficient or lower cost alternative.

The proposed study utilizes several advanced research designs and statistical techniques; however, limitations remain. First, it is impossible to know whether all relevant variables have been included in the model. Apart from an experiment that randomly assigns calendar arrangements to schools, sound methodologies that utilize the latest techniques for minimizing bias are needed to estimate the effects of school calendar on student performance. Second, to take of advantage of an exogenous switch in the school calendar, the fixed-effects and difference-in-differences analyses are limited to a sample of schools from a specific county in North Carolina. The study provides high levels of interval validity but since Wake County, North Carolina may not be representative of the population of schools that could be converted to a modified year-round; the external validity of the effect estimates may be reduced. The remainder of the dissertation includes a review of the literature on the theoretical framework suggesting benefits for a modified school calendar (Chapter 2), the policy rationale behind calendar reform (Chapter 3), details the methodology behind the current study (Chapter 4), presents model results (Chapter 5), and discusses findings with suggestions for future research (Chapter 6).

# CHAPTER 2

# THEORETICAL LITERATURE REVIEW

Calendar reform restructures schooling and non-schooling time. Depending on the reform, calendar changes may increase the effectiveness of schooling time because of qualitative differences in schooling that occur based on the specific spacing of learning sessions. Both memory and time-on-task theories provide a context for understanding how patterns of schooling and non-schooling intervals impact student learning. Next is a description of the literature in mass versus distributed theory and theories based on the time-on-task conceptual framework.

## Memory and Spacing

The central difference between the modified year-round calendar and the traditional calendar is the distribution of in-school and out-of-school time. The traditional calendar lumps together in-school time while the modified year-round calendar creates chunks of in-school time (on average, nine weeks) separated by out-of-school time (on average, three weeks). The effect of the dispersion of in-school learning time in each calendar arrangement may be examined through the mass versus distributive practice theoretical frame. The traditional calendar can be compared to the mass condition, where learning time is massed together in one period, while the modified year-round calendar is likened to the spaced condition, where learning time is divided into multiple periods. Next, is a literature review of the spacing effect including a summary of various theories attempting to account for the phenomena and landmark studies within the field.

### Early Learning Theory

The spacing effect is grounded in early learning theory. Performing memory experiments on himself in the late nineteenth century, Ebbinghaus surmised that spacing learning trials improved

retention compared to learning trials packed closely together (aka "the spacing effect"). Ebbinghaus's groundwork spurred a number of studies on the spacing effect in the early twentieth century (Dempster, 1989). Although methodological procedures were far less advanced, many of the studies confirmed the presence of the spacing effect.

Within the context of early learning theory, the two main facets of learning are: changes in resistance to forgetting and changes in lag time between a question and an answer (Estes, 1955). According to Estes, both facets, combating forgetting and reflex response, are in line with Skinner's work on the use of reinforcement schedules that increase or decrease the likelihood of eliciting a specific behavior (Skinner, 1938). Labeled "habit strength" and "response strength," Estes quantifies both facets through a series of probability curves. Using this approach he finds that the number and spacing of conditioning episodes is related to habit strength, meaning that forgetting can be manipulated by changing the number of learning occasions and the amount of time between learning occasions. Increasing the time interval between conditioning episodes results in greater retention. Response strength (how quickly a participant answers a question) appears greater with very short intervals between conditioning episodes at first, but over time spaced conditioning episodes tend to produce faster recall (Estes, 1955). Grounded in behaviorism, Estes reasons that improved retention is a direct function of the number of similar elements between conditioning episodes and recall trials. This approach is later fleshed out by component level theorists (Glenberg, 1979).

**Defining the Spacing Effect**

A flurry of research occurred in the 1960s and 1970s, when different theories such as the components level theory, were put forth to help further explain the consistency of the spacing effect. Interest in the relationship between studying intervals and retention resulted in an explosion of experiments. For example, students studied spelling through six computer exercises that were either spaced one day apart (the spaced condition) or all occurred on the same day (the massed condition) (Fishman, Keller, & Atkinson, 1968). Memory tests of this sort revealed differences in retention for

participants in spaced versus massed conditions, favoring spaced conditions with remarkable regularity (Hintzman, 1974; Melton, 1970).

Dempster's (1989) review of spacing effect research summarizes several important features of the spacing effect that have been highlighted repeatedly in the literature. First, the overall size of the spacing effect is pronounced, especially compared to other learning variables which have comparably weaker effects. Second, the size of the retention interval (the amount of time between the last learning episode and the test) is also related to the size of the spacing effect with longer retention intervals favoring more spaced learning trials. Contemporary research has focused on this aspect of the spacing effect and is detailed later in the review. Third, it appears that the spacing effect has some boundaries, specifically in terms of material and context. Massed reading passage trials were just as effective as spaced reading passage trials (Austin, 1921). Recalling nonsense syllables may be less influenced by the spacing effect phenomenon than the acquisition of more meaningful information (Underwood, 1961). However, even though some material may be less amenable to the spacing effect, the spacing effect has been evidenced in virtually all traditional learning tasks (see Dempster, 1989 for full review).

There may also be a developmental component to the spacing effect. Toppin and DiGeorge (1984) postulated that the neural mechanisms associated with the spacing effect mature in time. Using evidence from similar experiments performed with preschoolers and first graders, the spacing effect was present for first graders, but not for preschoolers. However, other researchers have reported effects in young children (Cornell, 1980; Rea & Modigliani, 1987). By and large, the spacing effect is extremely robust over different subject matter, interval spacing, and throughout a long research history.

**Encoding Variability Theory**

Explanations for the spacing effect can be classified into two categories; encoding variability theories and deficient-processing theories (Hintzman, 1974). Underlying the encoding variability theory is the assumption that information is encoded and that as the number of encodings increases so

does the number of pathways for retrieval. The explanation for recall differences between participants in massed versus distributed conditions is due to partial versus full processing. In massed conditions, repeated trials spaced closely together require less processing because simple retrieval of the previous encoding is so accessible. The spaced condition requires deeper processing upon each trial because the prior encoding happened longer ago. Using deeper processing for successive trials is what encoding variability researchers believe accounts for the higher long-term retention of participants in a spaced versus massed learning condition. Addressing the need for further explanation of the role of context within the encoding variability theory, Glenberg (1979) provides a more explicit account in his component levels theory.

**Component Levels Theory**

In an effort to explain the spacing effect, Glenberg (1979) describes the component-levels theory which outlines three types of components in relation to the retrieval process. In component-levels theory various stimuli are encoded by the individual either voluntarily or involuntarily resulting in the formation of a complex episodic trace. Multiple components make up the episodic trace; contextual components, structural components, and descriptive components. Contextual components include all stimuli related to the individual's environment, including the physical setting, time of day, and the individual's mental and emotional state. Contextual components are encoded automatically. When information is encoded in different contexts, the amount of contextual components increases. With more contextual components to aid in retrieval of the episodic trace, an individual is more likely to recall information. Unlike contextual components, structural components are encoded based on an individual's voluntary processes to manage new information. The ways in which an individual organizes or chunks information together is the structural components part of the episodic trace. Lastly, descriptive components relate new information with prior experiences. By linking the semantic memory to the input of new information, the descriptive components are formed on a deeper level of processing. Glenberg (1979) reasons that spaced learning conditions allow for greater exposure to a rich set of contextual components and more opportunities to utilize different structural

and descriptive components to create a more elaborate memory trace resulting in higher rates of retention.

Component-levels theory goes one step further by attributing the encoding feasibility to different features of the trial process that create an "episodic trace." Surface structures, seemingly superficial features of a learning trial, can aid in the creation of an episodic trace. Component-levels theorists argue that variations in learning trials create more episodic traces which lead to easier retrieval in later tests. Dellarosa and Bourne (1985) set out to test this facet of component-levels theory by changing the surface structure of learning sessions. The researchers manipulate surface structure by allowing free recall versus verbatim responses and substituting different speakers during learning trials. Indeed, the results supported the component-levels theory because in both manipulations students in the massed condition performed just as well as the students in the spacing condition.

Recall that Glenberg's components level theory offered three elements that lead to the creation of a rich episodic trace; contextual, structural, and descriptive. Dellarosa and Bourne (1985) point out that their manipulation of speaker voice and use of free recall instead of requiring simple recognition may have altered one or many of these three elements. Precisely which elements are altered by the study manipulations and understanding which element is the most consequential is still in debate. Dellarosa's research team suggests that perhaps differential weights should be assigned to the three elements, the weakest to contextual elements. Additional experiments in the 1980s did not support encoding variability theory or related theories, and it is no longer a preferred interpretation for the spacing effect (Dempster, 1987; Postman & Knects, 1983).

**Deficient Processing Theory**

The second broad classification of theories explaining the spacing effect is deficient-processing theory. The deficient processing theory stipulates that massed presentations require less mental strain than distributed presentations. The theory relates the amount of processing to retention. Since massed presentations result in less processing, less information is retained. Grounded in

deficient processing theory, four explanations for differential processing have come forward: (a) consolidation, (b) habituation, (c) rehearsal, and (d) attention (Dempster, 1989). The first two contend with involuntary processing; that is, processing that occurs without effort from the individual. The latter two represent techniques a person actively uses in order to increase retention.

Proponents of the consolidation mechanism say that information must be converted from just-learned memory to long term memory storage (Dempster, 1989; Landauer, 1969). Since the conversion requires time, distributed presentations allow time for consolidation, and therefore result in greater long-term retention. While the consolidation theory has been set forth as an explanation for the spacing effect, little research has been conducted to test the theory explicitly.

According to habitation theory, after the first presentation of new material the brain shuts down. If a second presentation occurs soon after the first it cannot receive full processing treatment. Additional presentations are further limited in processing. Interestingly, the habituation explanation has also been neglected by researchers.

Switching to voluntary behaviors that flesh out the deficient processing theory, the rehearsal hypothesis suggests that spaced presentations allow more time for rehearsal. The increased rehearsal is what cements learning. Research on the rehearsal hypothesis has been limited, however the finding of the spacing effect in young children, who are not old enough to engage in regular rehearsal of new information, does not support the rehearsal mechanism (Cornell, 1980).

Lastly, for the attention explanation, providing time in-between presentations enables individuals to devote more attention to subsequent presentations. Massed presentations result in less attention given to all presentations beyond the first presentation because the subject chooses to pay less attention to all subsequent presentations (Dempster, 1989). Attention theory has received support from a large body of research on attitudes of participants. Using a survey on attitudes during a memory activity, respondents noted that interest and attention were two areas that differed based on mass versus distributed presentation (Dempster, 1986). (Additionally, students in the spaced condition had higher rates of retention.) Pupil dilation, one way in which researchers measure

attention, is related to time between presentations (Magliero, 1983). Following up on the voluntary nature of participants' attention levels, researchers isolated several reasons why an individual may choose to pay less attention under massed conditions. First, a false sense of confidence after the initial presentation may prompt individuals to believe they have mastered material at a higher rate than has actually been retained. Zechmeister and Shaughnessy (1980) found that college students were more confident about the amount of material they believed they had remembered in massed condition versus the spaced condition. In actuality, individuals in the spaced condition outperformed those in the massed condition on memory tests. Second, individuals may find massed repetitions boring and redundant, resulting in the choice to pay less attention. Elmes, Dye, and Herdelin (1983) had participants rate the "pleasantness" of spaced versus massed word presentations and found that the spaced words were rated as more pleasant.

However, the persistence of the spaced effect lends itself more to an involuntary process. Memory researchers tweaked the attention hypothesis in proposing a less voluntary mechanism, sometimes known as the reconstruction or accessibility hypothesis (Dempster, 1989). In a massed condition, the first encoding is easily accessible during the second presentation. The ease of remembering the first encoding requires less attention from the individual. Spaced conditions require more effort during the second presentation because the encoding from the first presentation is not as readily available, resulting in more effort in processing during the second presentation. Studies that restructure repeated presentations in order to seem less similar resulted in similar retention levels for individuals in massed and spaced presentation conditions (Dellarosa & Bourne, 1985).

**Assumptions and Expansions of the Spacing Effect Literature**

Bahrick (1979) brought to light many of the underlying assumptions and points of weakness within recall research. He classified memory research into two camps, successive relearning in a laboratory situation and the investigation of complex knowledge maintenance using realistic non-laboratory situations. Bahrick asserted that almost all memory research seems preoccupied with the former, although statistical advances are making the latter a more viable alternative. Bahrick pointed

out that restricting memory research to the laboratory has limited the type of information used for experimentation to material that can be easily mastered within relatively few, short trials and oftentimes restricts the retention interval as well. By adopting this mode of operation, memory researchers have effectively assumed that no fundamental difference exists between the acquisition and retention of simple information compared with more complex cognitive learning. Bahrick explained that differences in mass versus distributive practice are related to differences in the memory process, more specifically, to two distinct phases; the acquisition phase and the maintenance phase. He cautioned researchers to pay particular attention to which phase of memory is being manipulated.

Most interesting is Bahrick's (1979) examination of time differences between study trials and tested trials. Assigning participants to learning trials spaced a day, 7 days, or 30 days apart Bahrick found that after six learning trials the groups were not significantly distinguishable in terms of the amount of material remembered. However, he conducted a second experiment identical to the first, except that the final test was given 1 month after the last learning trial. In this experiment, Bahrick found that the group whose study periods were spaced 30 days apart recalled more information than participants whose study periods were lumped closer together. Bahrick had uncovered a relationship between sustained knowledge and the spacing of learning sessions. He explained that subjects change the encoding method they employ when they encounter items forgotten since the previous learning trial. Since the learning trials spaced further apart lead to more forgetting, the need to relearn during each study session, and thus trying new encoding strategies, is essential for improving retention. Using what Bahrick called mediators, participants in the long intersession intervals experiment and test various mediators that enhance encoding. Short intervals between test sessions do not allow participants to discover less durable mediators. Without testing various mediators, when there is a long delay between the last study session and the test, these participants underperform. The results were consistent across subject matter.

**Recent Research**

Following from the reconstruction or accessibility theory and borrowing ideas from Bahrick's (1979) body of work, a line of research developed on spaced testing. Honing in on the relationship between the intersession interval (time between learning presentations) and the retention interval (time between the last presentation and test), Roher and Pashler (2007) modeled the optimal interval lengths for the highest rates of retention. According to these scholars, increasing the retention interval leads to lower rates of retention regardless of the intersession interval. Retention is improved when the length of time between final study session and final test is similar to the length of time between study sessions. The model suggests that longer intersession intervals are best when long term retention is the goal.

In order to tap into the variability of intersession gaps and test delays (aka retention interval), Cepeda and colleagues (2009) tested different spacing and testing intervals to see which combination yielded the greatest retention rate. Using various study material such as Swahili-English word pairs and unfamiliar visually presented objects, the experimenters found that increases in test delay led to increases in the optimal intersession gap. For example, in one experiment, a ten day test delay led to optimal results with a 1 day intersession gap. In a second experiment a six month test delay produced high retention rates when paired with 28 day gaps between learning sessions. The research team recommends the use of repeated cumulative tests to practitioners aiming for long term retention.

Kornell (2009) examined the spacing effect through the lens of the learner. Using flashcards, learners were divided into two groups; massed and spaced. Participants in the massed condition studied a stack of flashcards four times throughout one day. The spaced condition studied the same cards once per day for four days total. Participants in the spaced condition outperformed participants in the massed condition on retention tests. One might believe that differences in the amount of study time between the two conditions could lead to differential retention rates; however the two groups of participants did not significantly differ in the amount of time used to study with the flashcards. Interestingly, participants incorrectly judged the effectiveness of the two study conditions.

Participants, whether under the massed or spaced condition, believed that persons in the massed condition had a greater command of the material than participants in the spaced condition. Based on prior research, Kornell speculated that participants mistakenly equated the relative ease of retrieval in the short-term with long-term learning (Benjamin, Bjork, & Schwartz, 1998; Kelley & Lindsay, 1993).

Karpicke and Roediger (2010) also manipulated the intersession gap, but instead of using constant intervals between study sessions, they expanded the learning session intervals over time. The researchers found no difference on memory tests based upon whether participants experienced equal or expanding intervals between learning sessions. Next, is an examination of the time on task literature, which also highlights a relationship between learning time and achievement.

## Time on Task

Shifts in the school calendar reorganize time spent in school, but time is not the only pre-requisite for learning. Established in the literature as "time on task," a line of research outlines the importance of time engaged in meaningful learning activities as a central tenet for increasing student achievement. A modified year-round calendar may impact time on task.

Carroll's (1963) classic paper presented a model of school learning, a product of years of research on the processes of learning a foreign language. According to the model, five variables accounted for student performance: aptitude, opportunity to learn, perseverance, quality of instruction, and ability to understand instruction. The first three variables were said to be affected by time. Aptitude was defined as the quantity of time required for a particular student to master a task. Opportunity to learn was simply being given the time necessary to learn. Perseverance was thought of as the student's motivation to learn. The basic outline for the time-on-task model influenced other bright scholars in the field of education research as well as spurring a number of school policy changes as well (Karweit, 1984).

Many researchers have used Carroll's model to create expanded models of the time-on-task construct and to test different components of the model. Synthesized into five categories, Anderson

(1981) organized the field of time-on-task literature into the following categories: ecological psychology, learning for mastery, survival skills, studies of beginning teachers, and a large scale evaluation.

**The Ecological Psychology Approach**

The ecological psychology approach originated from Kansas in the 1960s and outlined the concept of behavior patterns. According to ecological psychologists, certain classroom practices (e.g. reading circles) elicit specific behavior patterns regardless of the particular natures of the individuals themselves. Gump (1967) found that structures of the classroom such as whole-class versus small group instruction yielded differences in students' time-on-task. Specifically, students were more engaged during small group instruction and externally-paced activities (Gump, 1967). Kounin, another ecological psychologist, looked at how specific teacher behaviors influenced student time-on-task. Interestingly, Kounin (1970) found that the relationship between teacher behaviors and student time-on-task was moderated by the classroom setting. For example, variety in task was highly correlated to student time-on-task in seatwork settings, but not in whole-class instruction settings. Kounin, through collaborations with Gump and Doyle, fleshed out a new concept to explain their research findings coined "continuity of signal systems" (Kounin & Doyle, 1975; Kounin & Gump, 1974). Grounded in the idea of the importance of classroom settings, characterized by the information presented, the presentation style, the materials used, and the nature of the specific behavior setting contribute to the signal system which ultimately elicits certain behaviors from students. When continuous and aligned signals are communicated, students are more apt to experience high rate of time-on-task.

In the 1980s, research by Haertel, Walberg, and Weinstein (1983) also took an ecological approach by adding components to Carroll's model that were associated with the social environment. Referencing social factors such as the social nature of classrooms, the home life of the learner, the power of peers to shape an individual's actions, and the effect of the media, researchers sought to

expand Carroll's model. Synthesizing prior research on teaching, Walberg (1986) used past studies to draw support for the additional components he and colleagues had added to the model.

**Learning for Mastery**

One such scholar whom was profoundly influenced by Carroll is Benjamin Bloom, who used the time-on-task framework to ground his concept of mastery learning (Carroll, 1989). With mastery learning, Bloom asserted that every individual is capable of grasping specific material and the goal of educators is to either increase the amount of time spent on learning the material or to use teaching strategies that lessen the amount of time needed to master the task (Bloom, 1968). Preferring to focus on the "alterable" variables of Carroll's model, Bloom's work called educators to action by creating a direct link between classroom practices and student achievement (Anderson, 1981).

Bloom's learning for mastery approach includes several fundamental components including specific objectives communicated clearly to students, small chunks of information, short well-developed tests for the purpose of assessing whether an individual has understood a specific objective (also known as formative assessment), remediation when necessary and monitoring of remediation using assessments until all students have performed well (Bloom, 1968).

Several researchers have used Bloom's lens to examine time-on-task. For example, one series of studies gauges the influence of teacher practices on students' time-on-task dependent upon the nature of the students. Students with high initial ability experienced higher rates of time-on-task during lecture orientated classes than students with lower initial ability (Anderson & Scot, 1978).

Aldridge was also influenced by Carroll's model, expanding it into a mathematical formula and testing the formula in classroom settings (Aldridge, 1983; Johnson & Aldridge, 1985). While the effort was an attempt to quantify the process of mastery learning, it has yielded promising, but not definitive support for the model. Shulman and Carey (1984) included affect into the classic model. These researchers postulate that ignoring the passions and emotions of the pupil oversimplifies the learning process. Still others quibble over the use of the term "allocated time." Fleshing out the

theory, researchers reason that teachers must also monitor and discipline students, and divide allocated time into allocated exposure time and useable time (Wiley & Harnischfeger, 1974).

**Survival Skills**

In the survival skills line of thought, time-on-task is operationalized as four skills: attending, effort, volunteering, and obedience (Cobb, 1972). A set of experimental studies revealed a positive relationship between teachers' reinforcement of the four survival skills and students' time on task. Further experiments established an additional link between variations in students' time on task and student achievement. Higher performance was consistently related to higher rates of time-on-task (Cobb & Hops, 1973; Hops & Cobb, 1974; Walker & Hops, 1976).

**Beginning Teacher Evaluation Study**

In a three-part effort to better understand the effect time has on student learning, the Beginning Teacher Evaluation study (BTES), divided time into three types: (a) allocated time, (b) engaged time, and (c) academic learning time. Allocated time, the most literal meaning of time, was the time assigned for a particular task. Engaged time is synonymous with time-on-task. Lastly, academic learning time is the time students spend successfully completing tasks. In terms of engaged time, researchers found that interacting with teachers, teacher monitoring, and teacher-led group work corresponded with high time-on-task. Other classroom methods such as multiple activities occurring within the classroom and long stretches of independent work corresponded with low time-on-task. The findings suggested that communication of directions and expectations could be the driving force behind engaging students (Fisher et al., 1978; MacDonald & Elias, 1976).

**Follow-Through**

Proponents of the follow through approach use student achievement rather than time-on-task as the dependent variable. Connecting the amount of time engaged in certain classroom activities and a student's achievement, researchers isolated certain activities that may lead to higher student performance (Stallings, 1975; Stallings, Corey, Fairweather, & Needels, 1977). In summarizing the most successful classroom practices, some of the same themes echoed by other time-on-task research

are evidenced in this line of research. Continuous instruction, clear goals, teacher monitoring, immediate feedback, use of praise, and sufficient allocated time were all associated with higher student achievement in the Follow Through observational study (Stallings, 1980).

The various approaches outlined differ only slightly in their conceptualization of time-on-task, most retaining the fundamental components outlined by Carroll. Yet even small differences in defining the concept has led to very different conclusions (Karweit & Slavin, 1982). Trying to operationalize time-on-task in order to measure and study variation in time-on-task is another daunting task for this research vein. Researchers have measured time-on-task in a variety of ways, from counting the number of rehearsals to questionnaires (Bloom, 1984; Naus, Ornstein, & Aivano, 1977). Yet, the literature on time-on-task is vast and established in educational research as a well supported phenomenon (Hawley & Rosenholtz, 1984).

**The Economics Approach**

Another model expressing the relationship between time and student achievement is grounded on Carroll's work (Levin, 1986). The model of instructional time views student achievement as the function of four components; capacity, effort, time, and learning resources (Levin, 1986). Capacity refers to intellectual capacity to process information as well as the current physical and mental state of an individual. Effort is related to task perseverance and time is defined in terms of the duration of a learning activity. Learning resources concern the quality and the quantity of items that assist in the acquisition of knowledge. Interestingly, Levin acknowledges the role of fatigue in defining the effort component. While largely tied to student motivation, a student's ability to attend to a task over a large quantity of time is relevant, and ultimately impacts student performance. Another insight shared in Levin's work concerns the role of the student in manipulating time. Based in some of economics literature, Levin sees students as decision-makers, dividing time and energy on activities that would maximize personal well-being (Becker, 1965).

## Connection to Calendar Reform

The purpose of the dissertation is the estimate the unbiased effects of modified year-round schools on student achievement, student absenteeism, teacher retention, and school expenditures. Therefore it is important to understand how the modified year-round calendar could be related to the theoretical foundations of prior educational research. Next is an exploration of how educational theory may explain the possible link between the modified year-round calendar and schooling outcomes.

Calendar reform has theoretical foundations within the cognitive memory processing and time-on-task literatures. Time is the central component in research on both the spacing effect and the effect of time-on-task. Calendar reform requires the reallocation of schooling time. In particular, the use of a modified year-round calendar manipulates the pattern of schooling and non-schooling time periods. Both memory and time-on-task literatures suggest that manipulations of time should impact student learning as described in greater detail below.

Memory research presents a number of mechanisms that can explain how the modified year-round calendar may influence student performance. First, the act of creating spaced learning sessions (versus a single massed learning session) may facilitate greater academic success. Spacing effect research confirmed that individuals tend to perform better on tests when learning trials were spread out over time, as opposed to learning trials conducted back-to-back. Although theorists point to different specific explanations for why spaced trials enhance learning, a common theme shared among researchers is the value of the time *in-between* learning trials. The distinguishing characteristic of the modified year-round calendar is the frequent breaks between school sessions which create a spaced learning condition. According to memory theorists, breaks between learning trials provide an opportunity for deeper processing. Deeper processing is primarily due to an increased effort in recalling previously learned information which strengthens cognitive connections for long-term retention. For example, giving several opportunities spaced over time for recalling information allows individuals to test the robustness of their encoding strategies and ultimately select

an encoding strategy that favors long-term memory. For students in the modified year-round calendar, information must be reviewed several times throughout the year when students return from each break. Each break gives students time to process newly learned information and each return from break allows students to test encoding strategies to assess their durability. By the time students confront end of grade examinations they have experienced multiple opportunities to recall learned information and are thus better prepared than traditional calendar students whose recollection techniques may be less sophisticated.

Second, the benefits of the modified year-round calendar may be underestimated by the end of grade (EOG) tests. As recent research suggests, the strength of the spacing effect is related to the time between the last learning trial and the retention test. The value of the spacing effect is underestimated by an EOG test given soon after the final learning presentation. If the goal of schooling is to increase retention beyond the end of grade examination, students in the year-round calendar have a distinct advantage since they are more likely to retain information long after the final information presentation. Comparing student performance after a single year may minimize the advantage accrued by students in the modified year-round school calendar if the effects are cumulative over time. Ultimately the effect may be more evident after a few years of exposure to year round schools. Cumulative returns, due to the spacing effect, imply a steeper slope in children's learning trajectories when educated under a modified year-round calendar as opposed to a traditional calendar.

Time-on-task research is also a lens through which to examine the possible benefits of a modified year-round calendar. For ecological theorists, frequent breaks may impact multiple variables in a student's environment, such as the mood of teachers or the behavior of peers. Teachers, a critical part of a student's environment, select classroom practices based on a host of factors, including conditions in the classroom such as student attitudes, expectations of student behavior, anticipation of disruptions, or even a teacher's energy level (Harris, 1979; Rivkin, Hanushek, & Kain, 2005). If frequent breaks improve or enhance classroom conditions, the instructional quality for an

individual student is subsequently impacted. Frequent breaks may also stave off teacher burnout. Modified year-round calendars give more breaks during the academic year than traditional calendars. If teachers feel less stressed after each break, they may make different instructional choices which benefit students. Additionally, teachers may feel more rejuvenated resulting in lower rates of teacher turnover. Or, if students feel refreshed after breaks and come back to the classroom eager to learn, a collection of motivated students can directly impact the learning environment for each individual student in the classroom (Wentzel, 1998). Ultimately the time-on task literature supports a possible link between the modified year-round calendar and higher student achievement and between the modified year-round calendar and greater teacher retention.

Several explanatory models rooted in the time-on-task literature contain an effort component. Drawing on the associations among student effort, engagement, and ultimately performance, individuals on a year-round calendar may be uniquely positioned for greater success. For example, students under a modified year-round calendar receive multiple opportunities throughout the academic year to rest. More rest means students are less likely to experience high levels of fatigue; fatigue that decreases student effort, which leads to lower engagement, and ultimately results in poorer academic achievement. Students on a modified year-round calendar have also experienced a more recent break close to the final test compared to traditional calendar students. The recent break may help to alleviate stress due to the ramp up of activity that usually occurs during the final weeks of preparation for end of grade examinations. Again the links between student effort, student engagement, and student performance are at play.

Both the spacing effect and time-n-task literatures support a line of reasoning that favors the modified year-round calendar over the traditional calendar. For example, adopting the modified year-round calendar leads to spaced learning intervals which may enhance the cognitive processing of students resulting in greater retention of information. Or, according to time-on-task theorists, student effort is one key in determining the value of schooling time. Sustained periods of schooling as seen in the traditional calendar where student are in school for 9 months with very few short breaks, may

lead to student fatigue and decrease student effort. Frequent non-schooling breaks might reduce student fatigue and raise student awareness during in-school learning sessions leading to more positive academic outcomes. Building on the prior research, there is reasonable support for the notion that the modified year-round calendar may be a successful tactic in maximizing student achievement by manipulating the timing of schooling sessions without increasing the overall amount of schooling time.

# CHAPTER 3

# POLICY RATIONALE FOR CALENDAR REFORM

Calendar reform, specifically the adoption of a modified year-round calendar arrangement, may save taxpayers money and increase student achievement. Whereas the traditional calendar consists of 9 months of schooling followed by a three month summer break, the modified year-round calendar distributes schooling time throughout the year, substituting the long summer break with more frequent breaks dispersed throughout the year. Students, especially disadvantaged students, suffer from summer learning loss during the traditional three-month summer break. Extending learning time by adding summer school interventions does not consistently counter summer learning losses. Instead, eliminating the summer break by rearranging the school calendar could reduce learning loss during out-of-school periods. Additionally, the modified year-round calendar allows for multiple cohorts of students to rotate in and out of classrooms, leading to higher capacity schools using fewer resources. Shorter more frequent breaks may affect teacher retention, if teachers prefer more frequent shorter breaks to one long summer break. The literature review outlines three main areas of research: summer learning loss, summer school interventions, and the impact of modified year-round calendar reform including cost-saving multi-tracking and teacher retention.

## Summer Learning Loss

The most compelling argument from educational researchers for year-round schooling is explained through a phenomenon called summer learning loss. All students are susceptible to summer learning loss if their current educational experiences cannot be equaled during the summer break. Certain populations are especially vulnerable because of the need for consistency in maintaining learning. For example, children with behavioral disorders thrive on routines. Without

developing similar routines during the break these children are more likely to relapse into old maladaptive behavior patterns. Children who are learning English as a second language often live in homes where English is not spoken. The absence of listening to English and having opportunities to converse in English during a long break can hinder the language development process. Poorer students may not experience enrichment activities such as visiting museums or public libraries that wealthier families are able to provide for their children over the summer break. Teachers must take additional instructional time at the beginning of the school year to review information that students have forgotten over the course of the summer. Modified year-round schooling may be thought of as a way to combat the summer learning loss experienced by students, especially populations who traditionally have lower levels of achievement. Policy strategies that target struggling populations may result in reductions of the achievement gap. The next part of this section will focus on the literature surrounding summer learning loss and its implications for student achievement.

Early studies found summer loss effects for speed, but not accuracy in math computation (Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996). In the early 1900s, a study revealed less summer learning loss for children engaged in work over the summer months. Here were the first clues that the activities a child engages in during the summer may impact achievement levels. Researchers in the 1920s began differentiating between different types of students, observing a greater loss for students with lower IQ scores (Brueckner & Distad, 1924; Patterson & Rensselaer, 1925). Summer loss lost momentum in research circles and relatively few studies were produced in the 1930s and 1940s and virtually no studies in the 1950s. The 1960s saw a rise of studies examining summer learning loss. A meta-analysis of 26 studies implemented prior to 1975 reveals some patterns (Cooper et al., 1996). In the 26 studies, 48 out of 80 comparisons of summer learning indicate a loss during the summer months. The overall results were not statistically significant, however there were significant results once broken down by subject matter. Mathematical computation showed losses in math skill ($p < .001$) while math reasoning results were mixed.

Reading also showed mixed results, but in 11 comparisons of spelling, there was a negative change in spelling skills over the summer.

Studies between 1975 and 1996 gave more support to the notion that, at least in some areas, summer learning loss does occur. Using students in both public and private schools, Pelavin and David (1977) examined summer learning rates in students identified as at-risk for cognitive impairment or low educational achievement. In five separate evaluations with highly consistent results, Pelavin and David found that in-school learning could not be sustained over the long summer break. In response to Pelavin and David (1977), the National Institute of Education (1978) issued a report to Congress. The report took issue with Pelavin and David's conclusions that students in compensatory education (CE) programs, who may show large gains over the school year, cannot sustain achievement over the summer. Separate calculations by the NIE showed that changes in reading and math scores over the summer did favor non-CE programs, but that changes in achievement levels were not as extreme as those reported by Pelavin and David.

A breakthrough came in 1987 when Heyns observed a natural experiment on the effects of schooling existed because of the traditional calendar. During the school year students were influenced by schooling practices and the outside environment. However, during the summer, students were influenced by all of the same environmental factors except schooling. Methodologically, summer could serve as a control for all of the influences of family background and cognitive ability. Therefore differences in learning rates during a summer month compared to a month in school could reveal the true effect of school. Using middle-schoolers in Atlanta, Heyns explored schooling effects and found that learning rates over the summer were different from learning rates during the school year. On average students learn less over the summer. Upon further exploration, she found that minority students and students of a lower socioeconomic status experience even slower learning rates over the summer.

Tying summer learning loss to issues of equity grew in the 1980s. Summer learning loss was revisited in terms of subject matter, socioeconomic status, and race. In general, math skills had a

greater deterioration over the summer than reading skills, and poorer and black students tended to fare worse in summer loss. The topic, however, remained controversial as a large study known as the Sustained Effects Study gave way to different modes of analysis and interpretation.

The sentinel piece in the literature on summer learning loss was published in 2001 by Alexander et al. Taking a seasonal perspective, Alexander and colleagues looked at learning growth rates during the school year as well as during the summer and aggregated their results by socioeconomic status (SES). Using the latest methodology for clustered data, hierarchical linear modeling (HLM) these researchers examined growth rates using a within-person design. They note that differences in preschool achievement could be explained by out of school factors. The SES variable is a strong predictor of student achievement at the beginning of first grade (Entwisle, Alexander, & Olsen, 1997). Believing that SES does not just affect achievement prior to schooling, but also throughout the school years, Entwisle, Alexander, and Olsen sought to examine how time in school (versus time out of school) moderates the SES effect. Using longitudinal data that spanned 5 years and contained fall and spring data, the researchers were able to examine seasonal differences in individual achievement. Participants were selected from the Baltimore area and were tested beginning in first grade (1982-1983) through fifth grade (1986-1987). The results of this important study are as follows: Lower SES youth were, on average, 0.7 standard deviations behind upper SES peers; children show higher gains during school months, than summer months; iIn the summer, verbal gains exceed quantitative gains suggesting that mathematic learning is more strongly associated with the in-school experience. With a more detailed examination, gains during the school year were found to be similar across different student levels of SES. Interestingly, the summer results showed a different pattern. In verbal tasks, students with lower levels of SES show small improvements some summers and small losses in other summers. For this same population with quantitative tasks, summer loss is more dramatic, especially over the first two summers. Scores for fall achievement were either roughly equivalent to the past spring levels or lower. The story for higher SES students presents a much different picture. Upper SES children saw gains in both verbal and quantitative

learning over the summer months. Upper SES students started off the new school year ahead of their spring achievement. Plotting the HLM analysis showed that African American youths' growth rate is slower than Whites and that boys' lagged behind girls, in addition to confirming that lower SES students showed slower growth rates than higher SES students. When adding level 2 predictors to see if the SES, race, and gender effect operates similarly in the summer and the school year, the researchers found that only the SES effect disappeared. Conceptualized as the "faucet theory", the school provides learning opportunities during the school year, but during the summer the water is "turned off" and students are not regularly receiving educational resources. Middle and upper class parents react by supplementing learning opportunities during the summer, while students with lower SES are unable to receive similar opportunities. Thus, the achievement gap is stable throughout the school year, but grows over the summer when students with high SES continue to grow academically (albeit at a slower rate than the school year) while students with low SES do not grow or even regress.

**Summer School Interventions**

One possible solution to summer learning loss is to supplement the summer break with summer programs, especially summer programs aimed at disadvantaged youth. However, the research on the effectiveness of summer programs is mixed. Five meta-analysis reviews on summer programs conclude that summer schooling has failed to eliminate the disparities between higher and lower SES achievement growth rates during the summer (Ascher, 1988; Cooper, Charlton, Valentine, & Muhlenbrock, 2000; Heyns, 1987; Karweit, 1993; Pipho, 1999). Effective summer programs benefit higher SES students more, which would exacerbate the inequities between upper and lower SES populations (Cooper et al., 2000). However, summer school attendance patterns may explain why higher SES students see a greater return. Attendance in a summer school program, such as Baltimore's Teach Baltimore summer school program, is associated with higher achievement during the summer months (Borman, Benson, & Overman, 2005). Not all of the studies on summer schooling have been methodologically sound. Randomized field experiments are the gold standard for scientific research. Schacter and Jo (2005) showed that a summer reading program did show

noteworthy differences between treatment and control groups. Yet other randomized experiments have not shown significant differences in achievement scores for children who have participated in summer programs and those who do not, even when high fidelity to the treatment is reported (Kim, 2007). In response to Alexander and colleagues's study, Borman and Dowling (2006) evaluated the longitudinal achievement effects of multiyear summer school. They found that treatment and control groups were not significantly different in achievement, until they examined attendance rate. Dividing up the participants in the summer programs into compliers and non-compliers and reanalyzing the data based on this distinction showed that compliers achieved 40-50% of one grade level higher than the control group counterparts. Roderick, Jacob, and Bryk (2004) investigated the effectiveness of the Summer Bridge Program, targeted towards students who fail to pass minimum competency requirements, in Chicago. Although they were able to isolate program effects on reading and math achievement, there was considerable variance across schools, surprising researchers since the Summer Bridge Program is a highly centralized and closely supervised program. Transportation issues are also important in order to ensure lower SES students can receive summer schooling.

Another part of the problem in analyzing summer programs seems to be the isolation of the specific summer experiences that are improving the achievement of higher SES students (Burkam, Ready, Lee, & LoGerfo, 2004) and how to replicate these experiences in a summer program. Trying to pick apart the skills to focus on during the summer, as well make the activities enjoyable for students whose peers are engaging in other summer activities, make designing an effective summer program difficult. The research indicates that children are learning at faster rates during the school year and therefore making sure they receive the quality schooling experience during the summer, without sacrificing some breaks for leisure and dramatically increasing school spending, year round schooling with a modified calendar is the most viable and potentially effective solution to reducing the SES achievement gap. Cooper et al. (2000) completed a meta-analysis on summer programs and created a set of summary conclusions. The first five, which were heavily supported by the literature, found that summer programs that (a) help students with learning challenges, (b) focus on accelerating

learning, (c) recruit middle class as opposed to disadvantaged students, (d) contain a small number of participants, and (e) deliver small-group instruction are most successful in impacting student achievement. Yet, the particular components that yield consistent summer school effects across schools and student subgroups remain elusive. The next section will review the literature on the effectiveness of year-round schooling to date.

## Modified Year-Round Schools

Instead of adding learning time with interventions such as summer school, calendar reformers argue that redistributing schooling time is another alternative to eliminating summer learning loss. Since the growth of modified calendar schools in the 1980s and 1990s, researchers have examined the potential benefit of calendar reform on student achievement, as well as other outcomes, such as teacher satisfaction, that impact the school context. Meta-analyses of modified calendars yield mixed results, but overall there is a small effect in favor of modified calendars (Cooper et al., 2003; Kneese, 1996, 2000). Kneese (1996) report that the type of track—single versus multi-track—influences student achievement. The effect size of modified year-round schooling ranges from .12 to .15, favoring single-year tracks. Other syntheses of year-round schools in the 1980s and 90s showed weak support for modified calendars, concluding that year-round schools certainly did no harm (Zykowski, Mitchell, Hough, & Gavin, 1991). In an update of Kneese's (2000) review, modified calendars again received modest support with 67.8% of comparisons between modified and traditional calendars favoring year-round schooling. In an effort to improve upon the selection criteria and statistical analysis of past analyses, Cooper and colleagues (2003) conducted a new meta-analysis to assess the effects of year-round schooling. In addition, the researchers gathered information and drew conclusions about parental, student, teachers, administrators, and the public's attitude toward year-round schooling. One particular important criterion for their meta-analysis was to include only year-round schools that redistribute the 180 calendar days and to discount data on schools with extended time in school. This way they could draw a distinction between the dispersal of days on achievement without the confounding factor of time. Of the studies found, 39 were

consistent with their criteria and could estimate an effect size. These studies were conducted between 1973 and 2000. Cooper and colleagues note that none of the studies comparing traditional and modified calendars were part of a randomized trial and less than half used statistical controls to make sure the initial populations were equivalent on important dimensions (such as achievement). This is an important point because inferences made from randomized experiments would be much more beneficial to the study of whether year-round schooling truly is a better alternative. Cautiously, the results of the meta-analysis revealed very small positive effects of year-round schooling (d = .01-.09). Students who had been on a modified calendar for more than one year showed positive effects (d = +.07), but when they had been on a modified calendar for more than two years, the students achievement went down (d = -.11). In regards to potential moderating effects of the calendar—achievement relationship, the students socioeconomic status did have an effect. Poorer students on the modified calendar school year performed on average .2 standard deviations higher than students on a traditional calendar. Thus, Cooper and colleagues surmise that year-round schooling may be considered as a policy solution for achieving equity with respect to socioeconomic status. Students in elementary school also show more positive effects of year-round schooling (d = +.09) than high school students (d = .00). With respect to attitudes, people were more positive about year-round schooling after experiencing a modified calendar than they were prior to its implementation. Yet these estimates should be interpreted cautiously, because the research literature on year round schooling leaves much to be desired. Error-prone techniques and a lack of research contribute to distrust of modified calendar effect estimates. There are a number of methodological issues plaguing the year-round calendar literature.

The article by McMillen (2001) entitled, "A Statewide Evaluation of Academic Achievement in Year-Round Schools," adds to the literature on the effectiveness of year-round schools. McMillen uses a sample of North Carolina students and examines whether changes in achievement scores are attributable, in part, to attending a year-round versus traditional calendar school. Controlling for individual characteristics, he finds student achievement, on average, is slightly higher for a student

following a year-round calendar than a student following a traditional schooling calendar. Besides the inherent strengths, such as the use of multilevel modeling to study a school level variable and the large sample size which maximizes the power to detect a statistical difference, there are several methodological flaws.

McMillen controlled for student-level variables of previous score, gender, minority/majority, and parental education. However, he does not control for student grade level. Prior studies suggest that the year-round calendar is more effective in the earlier grades (Alexander, Entwisle, & Olsen, 2007) for many possible reasons such as lower student attention spans which benefit from more frequent breaks, and smaller variance in achievement scores because achievement disparities grow over time due to differential summer learning loss. Also, the author does not control for other individual factors which may contribute to achievement differences such as whether the student is enrolled in special education, is an English Language Learner, or the student's socioeconomic status. If a variable that is unspecified in the model is responsible for the treatment effect then the model suffers from omitted variable bias. A more thorough list of individual covariates would make omitted variable bias less likely. Interestingly, no other school composition controls were included at level 2 of the model, besides the treatment variable. Peer effects and teacher effects were ignored. While McMillen's study takes advantage of a extensive detailed multilevel database, the study suffers from several methodological flaws and does not take full advantage of the available data.

Another facet of year-round modified calendar research is the ability to increase the number of students a school can serve by multi-tracking. Modified year-round calendars can be categorized as single-track of multi-track. Single track schools contain students who are all on an identical schedule, while multi-track schools contain several groups of students who follow different schedules. Rotating students in and out of classrooms means fewer resources, such as textbooks and tables, are needed. These structural changes are accompanied by changes in personnel. Specialist teachers, such as music teachers, may work 12 instead of 9 months out of the year, or additional staff is hired. Teachers may choose to substitute teach during their 3 week breaks. Overall, however,

year-round schools are thought of as a fiscally sound alternative to dealing with growing school-age populations and limited resources. One of the most comprehensive studies on the financial savings of year-round schooling was conducted by Daneshvary and Clauretie (2001). Using an economic model, Daneshvary and Clauretie compared costs between traditional calendar elementary schools and elementary schools operating on a year-round calendar. They found that the total cost per student drops by 7.5% for year-round schools, which is equivalent to about $400 per pupil (controlling for test performance and daily attendance). Operational costs dropped 12.3% per pupil for year-round schools. Year-round schools were most advantageous in the real estate savings, with a reduction of 31% per student. Many multi-track schools are located in high-poverty areas (Orellana & Thorne, 1998). While single-track calendar change is motivated by community concerns, parents' voices regarding multi-track modified calendar change are often not solicited in calendar shift decisions, adding to the controversy of calendar change.

Perhaps the biggest organizational issue for administrators of multi-track year-round schools is student assignment. A study by Burns and Mason (1995) claims that multi-track systems in year-round schools limit the principal's flexibility in student assignment, changing the compositional nature of classrooms, with homogenous grouping of students in classes resulting in heterogeneity among classrooms. A troubling study by Mitchell and Mitchell (2005) revealed the startling implications of multi-tracking: de facto segregation. In their comprehensive study of California year-round schools, Mitchell and Mitchell found that specific demographic characteristics were associated with students in the four different tracks. They note that assignment to tracks for year-round schooling is usually conducted by attendance boundaries, associations with certain programs, and choice for both teachers and students. Most students continue in the same track from one year to the next. Thus issues of equity in terms of the classroom composition as well as school resources (which includes teacher assignment) are pertinent to the study of track assignment. Mitchell and Mitchell make the following conclusions based on their California data. The tracks with more out-of-school time during the summer (thus resembling the traditional calendar arrangement) are the most popular.

Students in the traditional-like track tend to be white, wealthier, and English-speaking. This most popular track also has a higher mean achievement score than the least popular track, by the equivalent of 1½ years of schooling. Children in the least popular track start off with lower achievement and progress more slowly. This track has a high concentration of English Language Learners, is 2½ times more likely to contain poorer students compared to student composition in the most popular track, and twice as likely to be non-white. Mitchell and Mitchell find that program differentiation creates these demographic differences between the tracks. However, teacher sorting may explain the differences in learning rates. They find that teachers of students of the track with significantly poorer students have less years of experience and are 4 times more likely to have teachers with alternative credentials (such as lateral entry) than teachers in the traditional-like track. Mitchell and Mitchell's (2005) study challenges the notion that year-round schools can create more equity between students of various economic backgrounds. Clearly, year-round schools must respond to the scheduling challenges of a multi-track system without contributing to segregation in America's schools.

The change from a traditional to a modified calendar can be stressful for educators and administrators. Yet, teachers who have experienced year-round schooling have positive reports (Cooper et al., 2003). Flexible scheduling is one benefit for teachers. For example, having short breaks has allowed schools to charge for academic intersession interventions. These intersessions are a way for teachers to make extra money and for teachers on maternity leave to stay in the loop (Haser & Nasser, 2003). Additionally, principals report less teacher absenteeism with a modified schedule, which they purport results from less mental-health days. Intersessions can involve remediation as well as enrichment. Teachers enjoy designing units on topics they would like to teach about, but do not have ample time in the regular schedule. One school, Timber Lake, which operates on a year-round schedule, says it offers teachers supplements for meeting with other teachers at a designated time during the three-week break for planning and reflection. The main drawbacks for teachers are scheduling difficulties when children are on different calendars and the shorter amount of preparation time between school years (St. Gerard, 2007).

## Weaknesses in the Current Year-Round Research

There is not much empirical research on calendar change. Part of the reason may lie in the social forces that constrain calendar change and result in few year-round schools. Or the reason may lie in application researchers' reservation to study and suggest such an emotionally-charged contentious intervention such as adopting a novel school calendar. Whatever the reason, Kneese, Cooper, and colleagues agree that few studies of the relationship between calendar change and student achievement exist. Next, is a detailed look at the problems in design, sample, outcome measures, and analysis in the modified year-round calendar effectiveness literature.

### Problems in Design

In the small number of studies that address calendar change, few are longitudinal. Longitudinal studies would establish whether the magnitude of the effect varies over time. Most studies do not report the number of years a modified calendar has been in place (Kneese, 1996). If the current literature samples modified calendars during their initial year or years of implementation, it is very difficult to determine whether the modified calendar's advantage is due to its novelty (aka the Hawthorne effect) or whether it truly promotes greater student learning. A recent study by Alexander, Entwisle, and Olsen (2007) where the ninth-grade achievement gap is partially explained by the widening gap over elementary summer breaks, suggests that the modified calendar could have cumulative effects, but without specific modified-calendar longitudinal studies, this remains speculation.

### Problems in Sample Size

The lack of stronger effect sizes for year-round schooling could be due to small sample sizes. For example, within the year-round calendar literature, the highly cited Gandara and Fish (1994) study included only three schools operating on a year-round schedule. A lack of statistical power could certainly dampen the prospect of identifying an effect of school calendar.

**Problems in Outcome Measures**

Rigorous testing of the year-round modified calendar effect requires reliable outcome measures. Due to regulations from No Child Left Behind legislation (2001), well-developed tests that correspond with school curriculum are largely available in the subjects of reading and math and are only required of students in third grade and above. Reliance on end of grade student performance outcomes limits generalizablity in terms of subject matter and student age.

Many studies of year-round modified calendars rely on self report questionnaires, cognitive tests that are not directly tied to the school curriculum, and observational reports that have less discernible criteria for measuring effectiveness (Gandara & Fish, 1994, Orellana & Thorne, 1998). Also developing is a new line of research that seeks to identify what activities are taking place during schooling breaks that contribute to differential learning rates. Unfortunately some of the most intuitive hypotheses such as increased parental interest and increased attention to activities such as reading or attending cultural events (such as visiting a museum) have not explained much of the variation between students in differing strata of SES (Borman et al., 2005; Burkham et al., 2004). However, most of the information on parents and students' activities over break is gleaned from self reports, which can be biased. While education researchers may be interested in the mechanisms behind differential learning rates, policy makers are less concerned with underlying mechanism and more concerned with results. Apart from a randomized trial, sound methodological studies that utilize the latest techniques for minimizing bias is the best manner for examining the school calendar—student performance relationship across subgroups.

**Problems in Analysis**

Differences based on year-round schooling implementation could also be clouding the effect estimates. For example, studies are not always disclosing the exact schedule, such as nine weeks on and three weeks off. The way the calendar change was implemented may also influence the modified calendar-student achievement relationship. For example, parent and teacher satisfaction influences attitudes toward year round schooling and may ultimately affect whether calendar change affects

student achievement (Cooper et al., 2003). If parents were involved in the calendar change decisions, the effectiveness of the modified calendar may improve. Orellana and Thorne (1998) found that when parents were informed about intersessions and incorporated into the change process it led to greater satisfaction with the year-round calendar. Teachers may like a modified calendar to prevent burnout, but administrators or specialists may not like it as much because they are often changing from a 10-month to a 12-month calendar. Community support in terms of providing options in child care or flexibility for high-schoolers in job sharing seems important for calendar change success. Considering possible moderators, such as parental input or teacher satisfaction would enhance understanding of how modified calendars benefit students.

Another factor that could be minimizing the effect size of the modified calendar is achievement differences between the students on different tracks at multi-track calendar schools. The assignment of students to tracks could result in different schooling inputs for different tracks. For example, placing all the English Language Learners on the same track may lead to differential average achievement by track (Mitchell & Mitchell, 2005). If the schooling experience changes as a result of track placement, this could confound the year round treatment effect.

Another issue that must be confronted in studies comparing traditional and modified calendars is the presence of summer school or intersessions. Controlling for the additional instruction is necessary in order to obtain a clean estimate. Controlling for additional instruction will assess whether it is the time dispersion or the extra instruction (in which case quality summer school could substitute) that benefits students. Additionally, intersession and summer school quality varies (Gandara & Fish, 1994). Therefore the quality of the intersessions should be assessed by answering questions such as: Who is targeted? What subjects are addressed? What is the class size? Is the focus on remediation or enrichment?

In sum, studies of modified calendars are not of the highest quality and the literature is only recently taking advantage of the latest statistical methods. Kneese's (1996) meta-analysis of studies from 1982-1996 found that out of 80-plus studies, only 15 met the minimal criteria of involving year-

round schools with at least one year of implementation that include a control or comparison group and has student achievement as a dependent variable. However, the lack of methodological and statistical strength is beginning to be addressed by newer studies that are taking advantage of the multi-level modeling framework and utilizing statistical packages where researchers can look at between school effects while controlling for individual differences.

## In Summary

Interest in year-round schooling is partially based on prior research findings and efforts by policymakers to find viable policy avenues for improving achievement and reducing costs. Prior research on effectiveness of year-round schools finds modest support for a modified year-round calendar, but much of the research is weak. While there is some evidence that the learning losses from summer breaks are greatest for students of low socio-economic status, there are few inquiries into the effects of year-round modified calendar schooling on these students, or other critical subgroups such as English Language Learners and students in special education. Equally important in terms of policy ramifications, is to probe whether assumptions about cost efficiency in year-round schools are accurate. Additionally, the effect of a policy change on other important parties, such as teachers, is essential in garnering public buy-in. Probing whether teacher retention is influenced by calendar change is one way to understand how the schooling calendar impacts teacher satisfaction. To justify a recommendation to switch schools to a year-round modified calendar requires careful attention to multiple facets of the school context. Whether students benefit, whether achievement gaps shrink, what teachers prefer, and the financial impact are all essential components to evaluate in order to advise policymakers who control school calendar arrangements.

## Next Steps

The current study seeks to contribute to answering those questions. Using advanced methodology and examining specific subgroups, the research design outlines a methodology and analysis plan for obtaining information to address the important questions at hand. Next is a discussion of the sample, specific research questions, methodology, and statistical analysis plan.

# CHAPTER 4

# METHODOLOGY

Modifying the academic calendar has the potential of influencing a number of important factors in education; namely (a) student achievement, (b) student absenteeism, (c) teacher retention, and (d) school expenditures. The objective of the current study is to provide an unbiased causal estimate of the effect of calendar type (traditional versus the modified year-round calendar) on those schooling outcomes.

## Design

The desire to identify causal relationships forms the foundation of most research designs. Following Rubin's work on the theory of causal inference (1974), the ideal situation on which to base an estimate of a treatment effect is to observe the same unit under the treatment and control condition at the same moment in time. In this case, that would mean observing the same individual under a modified year-round calendar and a traditional calendar during the same time period. The more formal definition of the causal model is expressed in the equation: $\tau_i \mid_{Z=1} = E(Y_{i1} \mid Z_i = 1) - E(Y_{i0} \mid Z_i = 0)$ where $\tau_i$ represents the treatment effect, *i* represents the individual, $Y_{i1}$ is the score for an individual who received the treatment ($Z_i = 1$) and $Y_{i0}$ is the value for the same individual who does not receive treatment. Commonly referred to as the "fundamental problem of causal inference" (Holland, 1986), the ideal situation upon which to base effect estimates is unattainable since an individual cannot experience both treatment and control conditions at the same time. Therefore, a series of alternative research designs, with pros and cons, have been developed to estimate treatment effects.

The design that most closely approximates the ideal conditions outlined by Rubin is a randomized experiment. In a randomized experiment individuals in a population are randomly assigned to either the control or treatment condition. The randomization process means that individuals with observed and unobserved characteristics which may influence the effect of the treatment on the outcome of interest are assigned to a condition independent of these characteristics. While the same individual is not assigned to both conditions, assignment to the treatment or control group is without regard to individual factors, resulting in two groups that contain individuals with the same or approximately the same expected value of potential outcomes. Estimating the effect of calendar type with a randomized experiment would require the random assignment of students into modified and year-round calendar schools. Outlined by Cochran and Rubin (1973), there are a number of arguments against the use of randomized experiments, including random assignment may not be conducted for political or ethical reasons; randomized experiments are not conducted because of expense; difficulty enforcing study protocols, an observational study may be "more representative" of the overall population than the restrictive environment in which randomized experiments are often performed, and other designs that may be implemented quicker and with less cost may be helpful to decide if an experiment should be conducted or empirical hypothesis generation (Cochran & Rubin, 1973). Due to the difficulties inherent in conducting a randomized experiment of calendar type on student achievement, observational designs, also known as quasi-experimental or group comparison designs were considered.

In quasi-experiments, the assignment of individuals to treatment or control groups is not done randomly but many other features of a randomized experiment are retained. In some cases, the means by which individuals are assigned to treatment or control conditions is known. In other cases it is not. The assignment process is one factor that influences possible design and analysis strategies. The most difficult part of quasi-experiment design is identifying a strong comparison group. The most rigorous designs address the biggest problem that stems from the non-random sorting of students into treatment and control conditions called selection bias (Henry, 2010). Selection bias may bias group

comparisons because the various mechanisms through which students enter treatment or control conditions could explain differences in outcomes. For example, if parents who are highly concerned about their children's achievement enroll their children in modified year-round calendars, then differences between students of different academic calendars may not be due to the effect of the school calendar itself, but instead due to having parents who are highly motivated to act as advocates for their child's education. Choosing a design that minimizes selection bias is of utmost importance when providing strong evidence of a causal chain. Because we would like the individuals in the treatment group to have a similar background to individuals in the control group, a number of designs have been proposed to reduce the undesirable differences between treated and comparison groups.

First, the quasi-experimental design that most closely approximates randomized experiments is regression-discontinuity. In regression discontinuity designs the exact manner in which assignment to a treatment or control group is known and the assignment is done based on a score on a quantitative assignment variable. For example, in a study of the North Carolina Disadvantaged Student Supplemental Fund (DSSF) a composite score ranked districts on a scale of advantage. A cutoff point was established and districts below the cut-off were considered "disadvantaged" and received additional state funding. Researchers used the cutoff to identify a treatment and control group (Henry, Thompson, Fortner, Rickman, & Zulli-Lowe, 2008). Districts near the cutoff are assumed be to be similar in all respects other than the additional funding in this example. Outcomes for students in districts below the cutoff (also known as the "treatment" group) were compared to students in districts just above the cutoff (aka "control" or "comparison" group). In contrast, no single assignment variable was used to assign students to schools with a modified school calendar. A host of factors influenced decisions to create or switch schools to a modified year-round calendar, ruling out the use of a regression-discontinuity design in the current study.

As stated earlier, random assignment is the preferred method for creating treatment and control groups because all factors, observed and unobserved, that may influence the outcome of interest is randomly distributed across groups. The difficulty of using observational designs to make

causal inferences is the potential for pre-existing differences in the treatment and control groups to bias the effect estimate. Commonly referred to as "selection bias," selection bias is addressed by a variety of design strategies. While randomized experiments and regression-discontinuity were not feasible for the current examination of calendar type and schooling outcomes, other quasi-experimental designs including matching designs, natural experiments, and the use of longitudinal data with a rich set of covariates are all possible design choices and will be discussed next in detail.

Natural experiments, a form of interrupted-time series, are another design option where conditions in the environment led to experimental-like conditions. More specifically, a naturally occurring exogenous factor creates two groups where researchers have data on individuals before the treatment and after the treatment. Often used in analysis of policy changes, natural experiments allow for plausible explanations that changes in outcomes can be related to changes in exposure to treatment or control conditions. Natural experiments are still vulnerable to differences between pre-treatment and post-treatment groups. The potential for endogeneity requires close examination in time-series designs. Endogeneity is the idea that an independent variable in the model could be correlated with an unobserved variable captured in the error term. Natural experiments may be strengthened by subtracting the historical changes occurring within the comparison group from the changes seen in the treatment group. More commonly referred to as "difference-in-differences," estimates are based on outcomes due to movement from a control to a treatment condition net secular changes that occur in similar units over time (Wooldridge, 2009). The major difficulty with natural experiments is providing strong evidence that control and treatment groups are similar and accounting for time-specific factors such that differences in outcome can only be attributed to the causal variable (Morgan & Winship, 2007).

Matching designs identify a comparison group based upon similarities with the treatment group based upon known, or "observed," characteristics. Several matching procedures exist, including propensity score matching, where the comparison group is selected in a two-fold process. First, the probability for each unit to be in the treatment group is estimated using pre-intervention or

non-time varying covariates. Units in the comparison pool are selected to be a part of the comparison group based upon the similarity of probability scores for comparison units to the probability scores of units in the treatment group. While matching procedures are an improvement over using the all possible comparison units in the sample pool to be part of the comparison group, the matching is limited to observed characteristics. Recall that in randomized experiments observed and unobserved characteristics can be assumed to be equally balanced across treatment and control groups. Matching procedures are only equalizing groups based upon the observed characteristics used in the matching process. Therefore the reduction in bias due selection is more of a concern in matching designs than randomized experiments. Additionally, matching procedures require the potential for close matches.

Third, the use of longitudinal data with a rich set of covariates is another way to deal with the problem of selection bias. In longitudinal designs that employ fixed effects, a person (or school) can serve as his or her own control. In essence each person experiences both the treatment and control condition (although not at the same time). Within a student fixed effect design, researchers can control for any variation that occurs between individuals due to variables which remain constant over time (non-time varying characteristics). The major drawback to this type of design is that it limits the sample to individuals who have experienced both the conditions (treatment and control) which could limit the external validity of study conclusions, especially if the effects are heterogeneous. Rich covariates also help in controlling for factors besides the treatment that could impact the outcome. Rich covariates can be used at different levels of analysis to control for outside influences of individual, teacher, or school characteristics. One particular helpful covariate is the prior score on the dependent variable. Referred to as value-added modeling, controlling for a pre-score subsumes all the factors (observed and unobserved) that contributed to the pre-level of the dependent variable. In an educational context, controlling for the prior year's achievement, assuming for a moment that it is perfectly measured, accounts for all the factors that have contributed to the pre-level of achievement, including neighborhood effects, family inputs, and genetic endowment. Adding covariates beyond the pre-score helps to account for variation which occurs during the time period between the pre-score

and the post-score that could influence the post-score. Rich covariates are only helpful in so much as they capture variation related to treatment assignment and the outcome and simultaneously reduce the correlation between the treatment and the error term.

The goal of this study is to generate an unbiased estimate of the modified year-round school calendar on schooling outcomes. Accordingly I have to incorporate the problem of selection bias into my research design. Interestingly a natural experiment in Wake County provides an opportunity to analyze a naturally occurring change in calendar type. To prepare for rapid growth expected to impact area schools, the school board in Wake County decided to switch 23 schools operating on a traditional calendar to a multi-track modified year-round schedule, in the hope that multi-track schools could serve more students and delay the need for constructing new school buildings in the immediate future. The decision to switch schools happened in 2006-07 with the intent to begin the new modified year-round calendar in the following year 2007-08. The lack of parental involvement in decision-making led to court battles that lasted into May of 2007. The courts decided that parents must consent to the modified year-round calendar and schools asked parents to fill out consent forms at the beginning of the new school year, which, because it was beginning a modified year-round calendar, began in June 2007. If parents did not consent they could opt for their child to attend a traditional calendar school. Several reasons, including the delay in distributing consent forms, the short turnaround time for parental decision-making, and students' desire to remain at their present school led to a very small number of students deciding to leave these 23 schools and an overwhelming majority to stay put (see Appendix A for a full history and timeline of the events in Wake County).

The naturally-occurring situation where the same students at the same school received instruction under both the traditional and modified year-round calendar is an ideal set-up for a student fixed effects analysis. Two factors make the calendar change in Wake County an ideal natural experiment. First, the student population remained nearly the same when the new calendar was instituted. Robustness checks on student turnover revealed that student mobility for the year of the

calendar change (12% of students moved to another school) was in line with mobility trends during this time period. Fixed effects can only be estimated on students who experienced both the control (traditional calendar) and treatment (modified year-round calendar) conditions. If many of the students migrated to a different school after the calendar change then our estimate of the calendar effect would not include an important part of the sample and limit generalizeability. Additionally, if selection to stay at the school created a different population after the calendar change, then schooling outcomes could not be attributed to the calendar change, but something about students or families who select into a modified year-round calendar. However, in this situation a large majority of the students remained at the school after the calendar switch which increases my ability to minimize the potential of bias due to selection into the modified year-round calendar. Second, while the calendar type changed the schools themselves did not. If the schools changed then the study would be vulnerable to omitted variable bias. However, in this situation the schools drew students from the same neighborhoods, maintained the same leadership, the same staff, and the same curriculum suggesting that confounding factors are not responsible for differences in student achievement before and after the calendar change. Difference in the schools besides the calendar type could explain differences in schooling outcomes. The fact that the schools remained constant across control and treatment conditions means that the estimate of calendar change is not biased by omitting school level variables (that were constant across time). Adding a rich set of student and school level control variables also helped to limit susceptibility to omitted variable bias.

Although the situation in Wake County provides a nice natural experiment to analyze, it is restricted to a small sample of 23 schools. In order to generalize to a larger population, a larger and more representative sample would be best. Longitudinal data from the entire state of North Carolina on all modified year-round calendars was available. Growth rates of 3-8 grade students in North Carolina under a modified year-round calendar could be compared to growth rates of students on a traditional calendar using growth curve analysis. However, the major point of concern is the differences between students and families who choose to attend a modified year-round calendar

school and student and families who chose to attend a traditional calendar school. There is a danger that differences in student and family preferences is what contributes to differences in student outcomes that occur under different calendar types. If reasons besides calendar type are responsible for differences in student outcomes, the primary drawback of using pre-existing groups in my growth curve analysis, then the estimate on the modified year-round calendar is incorrect due to selection bias. However, one additional advantage of the growth curve analysis is that I can obtain the average student's achievement growth over a longer period of time. Since the natural experiment and growth curve analysis have somewhat different weaknesses and strengths, it seems appropriate to perform both types of analyses and compare results.

Research questions addressing school-level outcomes such as expenditures and teacher retention can be analyzed using the fixed effects methodology at the school rather than student level with the Wake County sample discussed previously. However, a second element, the difference-in-differences approach, can be added to strengthen the school level design. If changes in spending or teacher retention occurred over time due to factors other than switching the calendar type, then those changes could bias the modified year-round calendar estimate. To account for time trends, we can select a group of similar schools which operated consistently on a traditional calendar and isolate any changes in expenditures or teacher retention that occurred naturally over the same time period. Taking the difference in school-level outcomes pre and post the calendar conversion and then subtracting out a second difference in removing variation due secular trends, provides a cleaner estimate of the effect of the modified year-round calendar. The difference-in-differences approach requires balancing the treatment schools with similar comparison schools. One concern is how the schools were selected to switch from a traditional to a year-round calendar. If the switch was not exogenous, then identifying an equivalent comparison group would not be possible. However, Wake County maintains that switched schools were selected based on projections for population growth in the coming decade. Since there was not an identifiable characteristic that made switcher schools different from other Wake County schools, the selection of a comparison set of schools is plausible.

Selection bias is not the only concern in establishing a causal claim. Making a case for a causal linkage also requires attention to (a) the sequence of events, (b) the regularity of the association, (c) whether a plausible mechanism exists, and (d) the ruling out of rival explanations, which includes selection bias (Shadish, Cook, & Campbell, 2002). Since it can be verified that students first attend a school and then take the end of grade tests near the very end of the academic year (approximately 3 weeks before the last school day of the academic calendar regardless of calendar type), I consider all methodologies as satisfying the sequence of events criteria. In terms of the regularity of the association, all schools maintain a calendar type throughout the academic year. Schools may change calendar types between years. The plausible mechanisms prompting higher student performance under certain calendar types were expressed in detail in the literature review and policy rationale chapters. Lastly, in order to rule out rival explanations, two methods addressing the same outcomes (the growth curves and student fixed effects) were deliberately selected to address the weakness in the other. In conjunction, the results of both methodologies provide a strong case for ruling out alternative explanations for the link between calendar type and student achievement. Below is a list of the research questions stated explicitly followed by a description of the methodological procedures used to address each question.

## Research Questions

Several research questions guide this study: (a) "Do students perform better on achievement tests under a modified year-round calendar compared to a traditional calendar?"; (b) "Do students exhibit lower rates of absenteeism under a modified year-round calendar compared to students under a traditional calendar?"; (c) "Do specific student subgroups (low SES, English Language Learners, students in special education) perform better on achievement tests under a year-round modified calendar compared to a traditional calendar?"; (d) "Is there greater teacher retention in year-round modified calendar schools compared to traditional calendar schools?"; and (e) "Do year-round modified calendar schools spend less money per student than schools operating under a traditional

calendar? And if so, what expenditure categories differ?" All research questions are addressed using a quantitative research design set forth in the following section of this paper.

## Participants

Data for the dissertation covers a five-year period, the 2004-05 through 2008-09 school years. Assembled by the Carolina Institute for Public Policy, the original data was provided by the North Carolina Department of Public Instruction (NCDPI). The NCDPI supplied a wealth of detailed information on student course enrollment and attendance, student achievement on End-of-Grade examinations, teacher characteristics such as years of experience and National Board Certification status, and detailed school level information from local education agencies across the state. NCDPI's data on student course enrollment included class rosters in which students and teachers were identified, making it possible to link individual students to teachers. Student roster information is matched with current year test scores for over 95% of the students listed on rosters. The dataset contains information on over 200,000 students each year in all North Carolina Public schools. The target population for the dissertation is elementary and middle school-aged students. Sub-populations of interest include students of low socio-economic status, students in special education, and English-language learners, all of which are identified and included in the dataset.

## Measures

A number of variables are used in a series of models that look at the effects of calendar type on schooling outcomes. Besides the main variable of interest, whether a school is operating under a modified year-round calendar or a traditional calendar, a host of additional variables are included in models. The need for these additional control variables stems from the need to separate out the effects of the student and school characteristics in order to isolate the true effects of the calendar type from these other influences. Theory and prior research suggests that factors at multiple levels, including the individual and school level, help to explain differences in student achievement as well as teacher retention and school expenditure patterns (Coleman, 1966; Sewell, 1964). Controlling for

these variables does not *assume* that these other characteristics influence schooling outcomes, but simply adjusts for differences *if* they exerted an effect.

**Student Measures**

Individual characteristics account for much of the variation in student performance (Coleman, 1961, Jencks, 1972). Below is an explanation of student measures used in the dissertation analyses.

*Standardized test score.* Students in grades 3-8 take end of grade tests in reading and math. For comparability across years and across grade levels, I use the standardized test score to define my outcome variable.

Students in North Carolina in the study time frame take a pretest at the beginning of third grade and an end of grade test in reading and math at the end of third through eighth grade. They are assigned a scaled score (and a grade of I, II, III, or IV; where III and IV are a passing score) for both math and reading. All tests are vertically scaled, meaning that students' scores can be compared across the years. The reliability and validity information for the North Carolina End of Grade reading and math tests are available from the North Carolina Department of Public Instruction in technical reports based upon the most recent time each test was updated. The math test was updated in 2006 and the technical report, available online, shows evidence of high test reliability. In Appendix B, Table 1 shows the reliability estimates for the math end of grade tests and Table 2 shows the reliability estimates among student subgroups. The reading test was updated in 2004. Table 3 in Appendix B shows the reliability estimates for the reading end of grade tests. Unfortunately, reliability estimates among subgroups are unavailable for the reading tests.

Validity evidence in the technical reports includes information about percentages of items that match each of the major strands mentioned in the North Carolina Standard Course of Study to support content validity. Concurrent validity estimates are given based upon teachers expected grade for a student measured against a student's score on the math and reading tests. Predictive validity estimates are given based upon interpretations of scaled scores to predict final fourth quarter report card grades compared to the actual grade achieved. In addition, the technical report on the

mathematics test reviews the procedure for teachers to verify test items and their alignment to the standard course of study by specific objective.

*Days absent.* Absenteeism is cited as a barrier to learning (Ehrenberg, Ehrenberg, Rees, & Ehrenberg, 1991; Lamdin, 1996). The days absent variable is the raw number of days absent in the current academic year. Days absent is used as an outcome variable or a control variable depending upon the model.

*Prior student achievement.* To make a strong case for any educational intervention, it is important to account for pre-existing differences in academic ability between treatment and comparison groups. It would be unfair to attribute differences in academic ability at the end of an intervention to a treatment, if in fact the groups differed in academic ability at the onset. Additionally, controlling for prior achievement is essentially controlling for all the determinants of achievement that occurred prior to the treatment or control condition (Rivkin et al., 2005). In order to remove the effect of initial academic prowess, I use the average score on the prior year's end of grade exam or pretest (standardized). If students have exam scores in multiple subjects, such as math and reading, I use the average of all tests. Approximately 90% of students are matched to prior year test scores.

*Gender.* Whether or not a student is male is controlled for in the models.

*Racial designation.* Racial designations are developed based on the categories constructed by the initial data source, the NCDPI. Students are classified into the following mutually exclusive categories: black, white, Hispanic, Asian, American Indian, and multiracial.

*Socioeconomic status.* On average, poorer students tend to perform worse o achievement tests (for review, see Sirin, 2005). Socioeconomic status has been defined by a number of indices usually including measurements of a child's family income, parental educational attainment, and parental occupational prestige (Entwisle & Astone, 1994). Unfortunately, the database does not contain information on parental educational attainment for all study years and does not collect any information on parental occupational prestige. However, NCDPI does collect a proxy for family income. Families complete paperwork pertaining to their financial situation which is evaluated in

order to determine whether a student qualifies for free or reduced price lunch. I use two variables for examining an individual's socioeconomic status: whether a student qualified for free lunch and whether a student qualified for a reduced price lunch.

*English language learner status.* Students who enter the classroom without English as their first language may encounter additional challenges in learning within North Carolina public schools. In order to provide additional support to English language learners, North Carolina schools administer a test to see whether students qualify for supplementary language services. As a result, students with English language difficulties are classified as Limited English Proficient and given certain accommodations. Once students have progressed in English language proficiency they may be exited from the program and no longer classified as Limited English Proficient. In order to track students who have exited the program, NCDPI identifies these students as persons who were once classified as Limited English Proficient. I use two variables to capture English language learner states. One variable identifies current students who are classified as Limited English Proficient. The second variable identifies individuals who were classified as Limited English Proficient at one time.

*Special education classification.* NCDPI outlines a number of exceptionality codes for which a student may be classified in order to receive special education services. Diagnoses include cognitive, behavioral, and emotional disorders. I use a variable that aggregates all special education designations into one special education status variable. Only students who are classified as academically gifted are excluded from my special education status variable. Academically gifted students are coded in a separate special education variable.

*Underage.* Schools establish entry dates for Kindergarten each year. Using those cutoff dates, the Carolina Institute for Public Policy was able to create a window of birth dates for each grade that represents on-time grade progression. Students whose birth date occurs after the on-time window are considered underage.

*Overage.* Using the same windows described above, a student who is older than the on-time window predicts is considered overage.

*Structural move.* Transitional periods are often related to dips in academic performance (Eccles & Midgley, 1989). One type of transition is the movement from one school to another because a person has reached the final grade offered in a particular school. In this dataset, this is commonly reflected in the transition from an elementary school (housing grades K-5) to a middle school (housing grades 6-8). Grade configurations for each school are provided by NCDPI. Student are linked to each NC public school they have attended , therefore mandatory transitions between schools can be identified and recorded as a structural move.

*Moved in year.* Moved in year is based on a student's enrollment. First the median value of days enrolled is calculated per school. If a student is enrolled for less than 7 days from the median value they are considered "moved in year."

*Grade.* Students are designated into grades based on NCDPI test records. Only students in grades 3-8 are used for the current dissertation study.

**Classroom Measures**

Several characteristics are particular to the classroom unit of analysis. Next is an explanation of classroom measures used in some of the dissertation models.

*Number of students.* The number of students is the number of students in a student's classroom.

*Average peer test score.* The average standardized score of the peers in a student's classroom, excluding the individual.

*Average peer dispersion.* The average spread of scores of peers in the student classroom, excluding the individual.

**School Measures**

Apart from individual factors, environmental factors, such as school features, oftentimes affect student achievement (Greenwald, Hedges, & Laine, 1996). Below is an explanation of school measures used in the dissertation analyses.

*Calendar type.* NCDPI classifies schools as operating under a traditional, year-round, or traditional and year-round calendar. Only schools which operate solely on a traditional or year-round calendar schedule were selected for analysis. I used an indicator variable to denote year-round calendar status.

*Between year teacher retention*. Between year teacher retention is measured by comparing teachers in all pay periods for the prior academic year and observing whether they appear in the September pay file for the current year. Teachers who are paid at a school anytime in the prior academic year and appear in September of the current academic year are counted as a returned teacher. The teacher retention formula, calculated for each school, is the number of returning teachers divided by the number of teachers employed at a school. Between year teacher retention is used as an outcome measure.

*Within year teacher retention.* Within year retention is defined as a teacher who is paid at a school in September of an academic year and is not paid at the same school in May of the same school year. Within year turnover is also used as an outcome measure.

*Number of students in the school.* NCDPI reports the average daily membership (ADM) for each school every year. Some analyses include a log or squared ADM term in order to permit proper modeling of this count variable.

*Average proportion of black, white, Hispanic, Asian, American Indian, and multiracial students*. Proportions of the school population by racial designations are aggregated from school rosters. Rosters contain information on all students in tested subjects. Therefore this variable is limited somewhat in that it is based solely on tested grades. In using this variable construction, I make the assumption that the racial designation proportions in non-tested grades (grades K-2) look similar to proportions in tested grades (3-8) within the same school.

*Average proportion of students with low socioeconomic status*. Following the logic in the student measure of socioeconomic status, the school level variable of socioeconomic status is based on the proportion of students in the school who receive either free or reduced-price lunch.

*ABC status.* NCDPI uses a classification system to reward teachers based on student performance. The designations are determined based on the proportion of students performing on grade level and on whether students have learned the amount to be expected for one year of schooling. The designations, from most to least prestigious, are as follows: Honor School of Excellence, School of Excellence, School of Distinction, School of Progress, No recognition, Priority School, and Low Performing. An indicator variable for each designation is used in the analyses.

*Growth status.* Used in conjunction to determine ABC status, the growth status is based upon the proportion of students who have mastered material appropriate for one year of schooling. Schools are classified as: High Growth, Expected Growth, or Expected Growth Not Achieved. An indicator variable for each growth status is utilized.

*Percent of Adequate Yearly Progress targets met*. Another school characteristic that is part of the No Child Left Behind federal legislation is the percent of Adequate Yearly Progress targets met by each school. Adequate Yearly Progress targets are based on student subgroups: white, black, Hispanic, American Indian, Asian, multi-racial, Economically Disadvantaged (as determined by qualifying for free or reduced-price lunch), Limited English Proficient, and students with disabilities. The number of targets is based upon having a minimum number of students who are classified into individual categories. Looking at student performance within targets and then across targets, NCPDI calculates the percent of Adequate Yearly Progress targets met.

*Average teacher supplements.* Local supplements are added to teachers' base salaries, which are determined by the state based on teaching experience and credentials. Supplements vary across local education agencies (LEAs). A preponderance of evidence suggests that teacher salary influences the quality of teachers schools can attract and retain (e.g. Ferguson, 1991). I use the average teacher supplement in each LEA to adjust for local differences in teacher salaries.

*Teaching experience.* Within the teacher quality literature, many find that teaching experience is critical in predicting student achievement. On average, educators early in their teaching careers struggle to induce the same level of student performance as experienced educators (Clotfelter, Ladd,

& Vigdor, 2007). NCPDI reports each school's percentage of beginning teachers, defined as teachers with less than 3 years of experience.

*Teacher credentialing.* There is mixed evidence indicating that teachers with a masters or doctoral degree produce greater student learning (Goldhaber & Brewer, 2000; Rivkin et al., 2005). I use either the percent of teachers with advanced degrees or the percent of teachers with a supplemental masters degree to control for any differences in student learning related to teacher credentials.

*National Board Certification.* National Board Certification (NBC) is a process in which teachers assemble documentation to demonstrate their use of best practices and mastery teaching. In addition, North Carolina rewards National Board Certification with a 12% salary increase. For these reasons, NBC status could impact student achievement and is controlled for in analytical models.

*Short term suspension rate.* Disruptive behavior with a school can affect the schools learning environment (Bowen & Bowen, 1999). I use the number of suspensions per 100 students in my models.

*Violent acts rate.* Disruptive behavior includes violent acts that occur on school grounds during the school day. Therefore I also include a variable defined as the number of violent acts per 1,000 students.

*Per pupil spending.* Evidence of a link between access to resources and student performance is provided in the educational finance literature (Burtless, 1996; Ferguson & Ladd, 1996). Using the expenditure data, organized by purpose and object codes provided by NCDPI, I was able to classify expenditures into categories. The sum of all spending categories (with the exception of expenditures spent on capital outlay and community services) creates the overall spending total. The average per pupil spending by a school was determined by dividing the overall school spending total by the number of students in the school. Per pupil spending serves as a control in some models and an outcome measure in others.

During the categorization process, several expenditure classifications are hypothesized to be affected by calendar type. Used as outcome measures, specific expenditure categories are described below.

*Per pupil spending in Regular Instruction.* Regular instruction expenditures consist of annual teacher salary, benefits, local salary supplements, bonuses, classroom materials for instruction of regular students.

*Per pupil spending in Special Instruction.* Special instruction expenditures consist of annual teacher salary, benefits, local salary supplements, bonuses, classroom materials for instruction of students with special needs.

*Per pupil spending in Student Services.* Student services expenditures consist of salaries, benefits, and materials for guidance services, psychological services, speech, language pathology, media services, and some health services related to instruction.

*Per pupil spending in Instructional Support.* Instructional support spending consists of expenditures related to media services, technical support for teachers, salaries and benefits for technology support personnel not coded for school and district administration.

*Per pupil spending in Transportation.* Transportation expenditures consist of salaries and benefits for transportation personnel and other expenditures related to the daily transportation of pupils.

*Per pupil spending in School Maintenance & Utilities.* School maintenance and utilities expenditures consist of salaries, benefits, and supplies for activities related to cleaning, repairing, and maintaining school premises and the utility charges.

*Per pupil spending in School Leadership.* School leadership expenditures consist of salaries, benefits, and supplies related to the principal's office.

*Per pupil spending in Capital Outlay.* Capital outlay spending consists of expenditures related to the acquisition of property, renovations, replacement of furnishings and acquisition of buses, etc.

The variables described previously are collected for all North Carolina public school students by the NC Department of Public Instruction, but only a subset of students are involved in the data

analysis. Missing data is infrequent because of the state-wide initiative to track student demographics and achievement in order to meet the federal No Child Left Behind standards. See Table 1 for a list of variables.

Using the variables described in Table 1, I seek to understand the effect of the modified year-round calendar on several outcomes: student achievement, student absenteeism, teacher retention between years, teacher retention within the year, and per pupil spending over all as well as per pupil spending in regular instruction, special instruction, student services, instructional support, transportation, school maintenance and utilities, school leadership, and capital outlay. Next is a description of the analytic strategy.

## Dual Methodologies

Two separate methodologies are employed to address research question one which postulates a causal relationship between calendar type and student achievement (Research Question 1). The two methodologies are called (a) growth curve analysis and (b) a student fixed effects model. The growth curve analysis will detect if achievement trajectories differ between students enrolled in a year-round modified calendar school and students in a traditional calendar schools. Although this method maximizes power and external validity by including many North Carolina students in the analysis, it may suffer from selection bias. That is, students enrolled in year-round schools may be systematically different than students in traditional calendar schools. Although statistical methods are employed to match similar students in control and treatment groups (details forthcoming in this chapter) and to remove variation from a rich set of covariates, bias could remain (Cook, Shadish, & Wong, 2008; Glazerman, Levy, & Myer, 2003). A fixed-effects design would not be as vulnerable to the problem of selection bias because students serve as their own controls. In essence, each student is in both the treatment and comparison group, but at different times. A natural experiment in Wake County provides a unique opportunity to observe the same students under both traditional and year-round calendars.

**Table 1. List of Variables**

| Student-Level Variables | Classroom-Level Variables | School-Level Variables |
|---|---|---|
| Current math or reading score (standardized) | Number of students in the classroom | Year-round school status |
| Days absent | Average peer score (standardized) | Between year teacher turnover |
| Average score on the prior year's end of grade exams or pretest (standardized) | Average peer dispersion | Within year teacher turnover |
| Gender | | Number of students in the school |
| Racial designation | | Average proportion of black, white, Hispanic, Asian, American Indian, and multiracial students |
| Eligibility for free or reduced lunch | | Percent of students receiving free or reduced lunch |
| English language learner status | | ABC status |
| Special education classification | | Growth status |
| Underage | | Percent of Adequate Yearly Progress targets met |
| Overage | | Average teacher supplements |
| Structural move | | Percent of teachers with less than 3 years of experience |
| Within year move | | Percent of teachers with advanced degrees |
| Grade | | Percent of teachers with a supplemental masters |
| | | Percent who are National Board certified |
| | | Short term suspension rate (per 100 students) |
| | | Violent acts rate (per 1000 students) |
| | | Per pupil spending |
| | | Per pupil spending in regular instruction |
| | | Per pupil spending in special instruction |
| | | Per pupil spending in student services |
| | | Per pupil spending in instructional support |
| | | Per pupil spending in transportation |
| | | Per pupil spending in school maintenance & utilities |
| | | Per pupil spending in school leadership |
| | | Per pupil spending in capital outlay |

In the fixed effects analyses student test score gains under a traditional calendar, before the conversion to a modified year-round calendar, will be compared to student test score gains after the conversion to a modified year-round calendar. The fixed effects design however, uses fewer students and could lead to limited power and generalizeability. Therefore, both methodologies will be implemented in order to see if complementary findings can support a causal effect for the modified year-round calendar.

The second research question examines the link between calendar type and student absenteeism. The student fixed effects using the Wake County sample is utilized for obtaining an estimate of the effect of the modified year-round calendar on the number of days a student is absent. While a growth curve methodology is appropriate for an outcome with additive advantages or disadvantages over time (such as student achievement), the modified year-round calendar is not hypothesized to accrue additive benefits over time. That is, students may experience fewer days absent under a modified year-round calendar, but they are not expected to be absent even fewer days the following year, after exposure to the modified year-round calendar for an additional year. Plus, there is a greater concern for ceiling effects with the number of days absent variable, and therefore it was judged inappropriate to apply the growth curve methodology. Only the student fixed effects results will be applied in addressing Research Question 2.

While both methodologies can be used to examine differences in student achievement due to calendar type, research question three, which focuses on the effect of calendar type on specific subgroups, can only be addressed through the growth curve analysis. The subgroups of interest include students who qualify for free lunch, students who qualify for reduced lunch, students who are Limited English Proficient, students who were Limited English Proficient, and students with special needs. The growth curve analysis allows for interaction terms where a separate effect can be estimated for each subgroup within year-round calendar status. In the fixed effects analysis students serve as their own controls. Only characteristics that vary over time can be included in the analysis.

Coefficients on relatively stable characteristics, including the subgroups of interest, are based on a very small number of cases and should not be interpreted.

Research questions four and five deal with school level outcomes. These research questions address between year teacher retention, within year teacher retention, and a variety of per pupil school expenditures. Like the fixed effects design, I exploit the natural experiment in Wake County where schools convert from a traditional calendar to a modified year-round calendar in 2007-08. However, using a fixed effects design for a school level analysis with such a small number of schools would lack power (the ability to detect an effect if indeed an effect exists). Fortunately the difference-in-differences comparison group design gives me the ability to increase power and control for "history," an important threat to validity (Shadish et al., 2002). To increase power I will add a number of similar comparison schools chosen through a process called propensity score matching which will be detailed in the following section. The difference-in-differences method addresses the threat of history, that is, whether historical factors are responsible for differences in outcomes instead of the manipulation of calendar type by controlling for secular trends. In difference-in-differences there are two "differences" that are considered. First, the difference between pre and post values in the treatment schools is estimated. Second, differences in the dependent variable that occur over time in the comparison schools, the variation attributed to historical factors, is subtracted out. Thus, the effect of calendar change can be isolated more effectively. Please see Table 4 in Appendix B for summary of research questions and analytic strategies.

## Data Analysis

### Propensity Score Matching

As alluded to in the prior explanation of growth curve analysis and the difference-in-differences methodology, better estimates can be generated if a comparison sample of traditional calendar schools is matched to the modified year-round calendar treatment group. Selecting an appropriate matched sample is extremely important in order to limit selection bias. Since the treatment variable (calendar type) occurs at the school level, matching procedure, called "intact group

matching," were used to identify schools operating on a traditional calendar (comparison schools) that are similar to the schools operating on a modified year-round calendar (treatment schools). The idea is that similar schools will house similar students. Several procedures exist for selecting a matched group. Propensity score matching is one such technique used to select a control group sample that is similar to the treatment sample. Propensity score matching was developed by Paul Rosenbaum and Donald Rubin (1983) to address selection issues in identifying comparable treatment and comparison groups. Methodological research indicates that propensity score matching can be a suitable procedure for reducing bias in estimating treatment effects (Cook et al., 2008; Diaz & Handa, 2005; Glazerman et al., 2003; Henry & Yi, 2009). In propensity score matching, the probability of being in the treatment group is determined for each control and treatment unit based on a set of observed characteristics in the time period prior to the intervention (the switch to the modified year-round calendar) or the analysis period. A probability of being a modified year-round school score is generated from a rich set of covariates for every school in the sample thru logistic regression. Covariates are selected based on their ability to contribute to the prediction of the propensity score. Then, using the propensity scores, a matching procedure is conducted to select a subset of traditional calendar schools that will serve as the comparison group. Propensity score matching will be done with the intact groups one nearest neighbor without replacement method and stratified by urbanicity (city, suburb, town, or rural) and school grade configurations (middle or elementary school).

In order to ensure that comparison schools were well-matched, the logistic regression included several variables that could explain differences in student achievement, stratified by rural location and grade configuration. As mentioned previously, I used a no- replacement, stratified intact group without caliper, one-to-one nearest neighbor matching procedure. According to recent work by Henry and Yi (2009), the intact group without caliper design leads to the identification of the strongest comparison group. Cook and colleagues also find that intact group matching is the strongest matching design, and add that, if good matches are obtained, effect estimates are very similar to estimates obtained from randomized experiments (Cook et al., 2008). In addition to using

the strongest propensity score matching design, two additional checks on the matched sample were conducted to support the quality of the match between comparison and treatment groups. The first balancing test examines the difference of the means of the logits of the propensity scores between treatment and comparison groups. The value should be low, less than .5, between the two groups to demonstrate that the odds of being a treatment school are similar across treatment and comparison schools. The second balancing test requires an examination of the ratio of the variances of the treated to comparison logits. Optimally, the value should fall between .05 and 2.0 to confirm that the odds of being a treatment school vary within a reasonable range among treatment and comparison schools.

Another check of the propensity score matching procedure involves a falsification test. Appropriate for situations with a pre-treatment time period, a falsification test regresses the treatment variable on the dependent variable for the year prior to treatment. If the treatment has an effect before it was enacted it means there are great differences between the treatment and comparison groups and the matching was unsuccessful. In this case, a falsification test can be performed using the difference-in-differences sample. Before 2007-08 all sample schools were under a traditional calendar. If differences in outcomes exist between schools that stay on a traditional calendar and schools that eventually switch calendars before the even switch occurs then we would not be able to say outcomes were due to calendar reform, but more likely influenced by other unobserved factors.

A further check on the integrity of the matching procedure includes a descriptive comparison of school characteristics among treatment schools, comparison schools, and all schools. The matching procedure is considered successful if it shrinks the differences between treatment and non-treatment schools to the point that they can be ignored. Therefore treatment schools should be more similar to matched comparison schools than to the full sample of schools on most characteristics. However, the matching utilized *observed* characteristics. If treatment schools are qualitatively different due to an unobserved characteristic and not captured by the covariates in the propensity score matching, then the matching procedure may not have created equivalent treatment and comparison schools. Differences in outcomes between treatment and comparison groups should not

be attributed to the treatment, in this case calendar type, but instead to an unaccounted for additional factor, resulting in incorrect effect estimates. While we can be confident in our matching with observed characteristics, it is important to note that we cannot be certain that treatment and comparison schools are also similar on unobserved school characteristics.

## Models

Several models are employed to address each of the research questions. To address whether calendar arrangement affects student achievement (Research Question 1), a growth curve analysis and a student fixed effects design will be implemented. Student fixed effects models are also used to examine whether a causal relationship exists between student absenteeism and calendar type (Research Question 2). The effect of modified year-round calendars on specific subgroups' student achievement is analyzed using the growth curves model (Research Question 3). In order to estimate the impact of calendar arrangement on teacher retention (Research Question 4) and school expenditures (Research Question 5), a difference in differences method will be applied. Sample descriptions and analysis models are presented for each research method. Refer to Table 4 in Appendix B for a summary of research questions and analytic strategies.

### Growth Curve Sample

The treatment sample for the growth curve analysis involves all year-round modified calendar students in North Carolina. From the three-year period of 2005-06 to 2008-09 there were 141 elementary schools and 21 middle schools that operated under a year round calendar (see Table 5 in Appendix B). The year-round modified calendar schools are public non-charter schools whose calendar does not include extended learning time. The comparison sample will be pulled from the 1,676 regular traditional-calendar elementary and middle schools using propensity score matching on school level variables for the 2004-05 academic year. Through the logistic regression procedure to determine propensity scores, several matching variables were identified. The matching variables included the log of the number of students in the school, the average proportion of black, Hispanic, Asian, American Indian, and multiracial students, the percent of students receiving free or reduced

lunch, per pupil spending, school performance designation (aka ABC status), school achievement growth designation (aka growth status), percent of Adequate Yearly Progress targets met, the percent of teachers with less than 3 years of experience, the percent of teachers with a supplemental masters, the percent who are National Board certified, the short term suspension rate, and the violent acts rate.

**Growth Curve Analysis Model Structure**

The data will be analyzed with multi-level modeling using the statistical program SAS version 9.2. Nested data does not conform to the assumption of independence of observations and can underestimate standard errors. Multilevel modeling takes into account the degree of dependence in the data. In the growth curve analysis, time will be nested within individuals, and individuals will be nested within schools.

I follow the slopes as outcomes framework (Raudenbush & Bryk, 2002), because I am interested in the differences in achievement slopes based on school calendar. Time is placed at the first level. Individual student variables are at the second level. Second level individual variables will help address whether certain subgroups outperform others under year-round modified calendar conditions. The second level variables also control for individual student characteristics. The third level contains school level variables including the treatment variable: an indicator of modified year-round school status. Additional third level school variables will control for the contributions of other school level characteristics. The model below is appropriate for answering whether a modified calendar is beneficial to student achievement overall (Research Question 1), as well as examining differences in calendar arrangement for disadvantaged student subgroups; student of low socioeconomic status, students who are learning English, and students enrolled in special education programs (Research Question 3).

Growth curve analysis is a powerful statistical model that permits each case to vary over time according to individual trajectories. Relating the various trajectories to each other on variables of interest, growth curve analysis output provides an estimate of the fixed effect of model variables. In this instance, a coefficient is produced for year-round school status and four interaction terms

between year-round status and each student subgroup of interest (free or reduced lunch status, was limited English proficient, is limited English proficient or student with a disability). Below is the statistical model.

(1) Level 1 (time): where *Y* is the scaled score on the EOG

$$Y_{tij} = \pi_{0ij} + \pi_{1ij}\text{Grade}_{tij} + e_{tij}$$

(2) Level 2 (individual):

$$\pi_{0ij} = \beta_{00j} + \beta_{01j}\text{SESfree}_{ij} + \beta_{02j}\text{SESreduced}_{ij} + \beta_{03j}\text{IsLEP}_{ij} + \beta_{04j}\text{WasLEP}_{ij} + \beta_{05j}\text{SpecialEd}_{ij} + \beta_{06j}\text{VectorIndividualControls}_{ij} + r_{0ij}$$

$$\pi_{1ij} = \beta_{10j} + \beta_{11j}\text{SESfree}_{ij} + \beta_{12j}\text{SESreduced}_{ij} + \beta_{13j}\text{IsLEP}_{ij} + \beta_{14j}\text{WasLEP}_{ij} + \beta_{15j}\text{SpecialEd}_{ij} + \beta_{16j}\text{VectorIndividualControls}_{ij} + r_{1ij}$$

(3) Level 3 (school):

$$\beta_{00j} = \gamma_{000} + \gamma_{001}\text{YearRound}_j + \gamma_{002}\text{VectorSchoolControls}j + u_{00j}$$

$$\beta_{10j} = \gamma_{10j}$$

$$B_{01j}... \beta_{05j} = \gamma_{010}\ldots \gamma_{050} + \gamma_{011}\ldots \gamma_{015}\text{YearRound}j$$

$$B_{20j} = \gamma_{100}$$

Where Grade = school grade (coded 3-8), SESfree = whether the student is qualified to receive free, SESreduced = reduced lunch (coded 1 or 0), IsLEP = whether the student is limited English language proficient (coded 0 or 1), WasLEP = whether the student was limited English language proficient (coded 0 or 1), and SpecialEd = whether the student receives special education (coded 0 or 1). Vector of Individual Controls includes: Race = a vector of indicator variables for a student's race: black, white, Asian, American Indian, and Multiracial (coded 0 or 1), Underage (coded 0 or 1), Overage (coded 0 or 1), Structural move (coded 0 or 1), Moved within year (coded 0 or 1), and Days absent (numeric count). YearRound is a dichotomous variable indicating whether the schools operate under a modified year-round calendar (coded 0 or 1). The vector of school controls include school size, the school size squared, total per pupil expenditures, the average teacher supplement, the violent acts rate, the short-term suspension rate, the free and reduced lunch mean, the black mean, the Hispanic mean,

the Asian mean, the multiracial mean, the American Indian Mean, the percent of teachers with less than three years of experience, the percent of teachers with an advanced degree, and the percent who are National Board certified, and the school propensity score.

In order to provide more robust evidence for a causal link between calendar type and student achievement and to provide an estimate of the modified year-round calendar on student absenteeism, a student fixed effects approach was also conducted.

**Student Fixed Effects Sample**

The sample for this study involves the Wake County elementary school students who have experienced the transition in 2007-2008 from a traditional calendar to a year-round calendar within the same school. Since 1990, Wake County has been gradually increasing the number of schools operating under a year-round calendar. By 2006 there were 20 year-round elementary schools in Wake County. In 2007 20 additional elementary schools and 3 middle schools in Wake County were mandated to operate under a year-round multi-track calendar where students are on one of 4 tracks, each with 180 total days of instruction (45 days of instruction followed by a 3 week break). This policy change came from the projected influx of students to the Wake county area and the lack of time (and to some extent money) to construct additional buildings. See Appendix A for a historical account of the Wake County traditional to year-round calendar transition.

Using an indicator variable for each student and controlling for school characteristics, the student fixed effects estimation will detect whether the switch to a modified year-round calendar resulted in higher achievement for students of the schools that made the switch. Data from four years (2005-06, 2006-07, 2007-08, and 2008-09) will be included in the analysis. Student fixed effects analysis looks at within-person variation, only individual variables that change over time will be retained in the fixed effects model.

**Student Fixed Effects Model Structure**

The student fixed effects model assumes that students progress at the same rate between grades. For example, the slope, which reflects the rate of change over time, between grades 3 and 4

and grades 4 and 5 are expected to be similar in the model. Extrapolations from this model also assume that any calendar effect is constant across grade levels. See below for the fixed effects equation.

(4) $y_{it} = \beta_{00} + \beta_{01}\text{YearRound}_{it} + \beta_{02}\text{VectorofControls}_{ij} + \beta_i s_i + r_{it}$

Where $y_{it}$ is student $i$'s achievement score at time $t$, $\beta_{00}$ is a constant, $\text{YearRound}_{it}$ represents whether students attended a school that switched to a modified year-round calendar, $\beta_{02}$ is a vector of individual and school controls that may have fluctuated over time including: eligibility for free or reduced lunch (the proxy for SES), English language learner status, special education classification, grade, days absent, underage or overage, whether the student moved during the school year, students average score on the prior year's end of grade exam or pretest, the number of peers in the classroom, the average dispersion of peers' prior test score, school size, the school size squared, total per pupil expenditures, the average teacher supplement, the violent acts rate, the short-term suspension rate, the free and reduced lunch mean, the black mean, the Hispanic mean, the Asian mean, the multiracial mean, the American Indian Mean, the percent of teachers with less than three years of experience, the percent of teachers with an advanced degree, and the percent who are National Board certified. S is a dummy variable for every student, and $r_{it}$ represents the error term.

**Difference in Differences Sample**

The difference in differences method will be applied to research questions concerning teacher retention and school expenditures. The sample will be limited to Wake County to take advantage of the natural experiment Wake County provided when 23 schools switched from a traditional calendar to a multi-track modified year-round calendar. The matched sample consists of 23 other Wake County traditional-calendar elementary schools. Again, the propensity score matching using one nearest-neighbor without replacement technique is used to select the appropriate control units. Comparison schools were matched based on the number of students in the school, the average proportion of black, Hispanic, Asian, American Indian, and multiracial students, the percent of students receiving free or reduced lunch, per pupil spending, school performance designation (aka

ABC status), school achievement growth designation (aka growth status), percent of Adequate Yearly Progress targets met, the percent of teachers with less than three years of experience, the percent of teachers with supplemental masters degrees, the percent who are National Board certified, the short term suspension rate, and the violent acts rate in 2006-07.

**Difference in Differences Model Structure**

In the difference in differences method, the outcomes of interest (in this case, teacher retention rates and expenditures) are measured before and after treatment; the conversion of schools to a year-round modified calendar in 2007-2008. However, the change in outcome may be attributed to historical factors other than the variable of interest; calendar arrangement. As you recall, the difference in differences method takes an additional step by subtracting out the variation attributed to other historical factors. The variation attributed to historical factors is quantified by examining the change in outcomes for the control group. In this case, the teacher retention and expenditure changes in matched Wake County schools that remained on a traditional calendar throughout the time frame will be utilized. See equation 5 below.

(5) $Y_{it} = \beta_0 + \beta_1 \text{Switch}_i + \beta_2 \text{Year}_t + \delta \text{Switch}_i * \text{Year}_t + \varepsilon_i$

where $Y_{it}$ represents each school's per pupil total expenditures, regular instruction expenditures, special instruction expenditures, student services expenditures, instructional support expenditures, transportation expenditures, school maintenance and utilities expenditures, school leadership expenditures, or capital outlay expenditures. In teacher retention models, $Y_{it}$ represents the percent of teachers who have returned from the prior year or the percent of teachers retained within one year. $\beta_1$ is coded 1 for schools that switch from a traditional calendar to a modified calendar and 0 for schools that remain on a traditional calendar. $\beta_2$ represents the academic calendar year (2006-2007, 2007-2008, 2008-2009). The estimate of the effect is the coefficient $\delta$ on the interaction term which captures the effect of the treated schools before and after the reform.

## Limitations

All of the proposed analyses have some limitations. Although minimized by the propensity score matching, the growth curve is limited in the potential for unobserved selection issues. The fixed effects model is plagued by a small longitudinal time frame, which could inhibit the detection of an effect. Also, the fixed effects model does not include students who tend to move schools often or students who transferred out of their school likely motivated by the conversion to year-round. The difference in differences method requires some difficult assumptions, including the assumption that the all other factors that could be influencing teacher retention or expenditures were similar across treatment and control schools, in essence that the secular factors affected all sample schools in the same way. Additionally, none of the analyses directly measure summer school or intersession tutoring that may occur during non-schooling intervals. However, the large sample size should help with power and external validity and ultimately the multiple methodological approaches should instill confidence behind the study conclusions.

# CHAPTER 5

# RESULTS

In order to address the possible advantages a modified year-round calendar may offer over the traditional calendar, five research questions were addressed using three different methodologies. First, to address the overall question of whether student achievement is higher in modified year-round calendar schools a growth curve analysis and a student fixed effects design were utilized (research question #1). The student fixed effects design was also used to investigate whether calendar type affects student absenteeism (research question #2). Next, to probe whether a modified year-round calendar is advantageous for subsets of the population, a growth curve analysis was conducted which included interactions between student subgroups and the modified year round calendar (research question #3). Lastly, to speak to whether teacher retention or expenditure patterns relate to calendar type (research questions #4 and #5), I used a difference-in-differences methodology.

Each of the three methodologies required separate samples. The growth curve analysis included all modified year-round schools in North Carolina and a matched sample of similar traditional calendar schools. The fixed effects analysis focused solely on the 23 schools in Wake County which converted from a traditional calendar to a modified year round calendar. In the difference-in-differences models the 23 Wake County schools that changed calendar type were included in the model along with 23 other matched schools in Wake County that operated under a traditional calendar throughout the entire study period. I turn now to a description of the matching procedure results, sample characteristics, and model results.

## Propensity Score Matching

An essential component for creating an argument for causal inference requires attention to how the comparison group is selected. Next is an overview of the propensity score matching process used to select the comparison sample schools.

### Growth Curve Analysis

In order to examine differences between student achievement in modified year-round calendar schools versus traditional calendar schools, I needed to identify traditional calendar schools that were similar to the modified year-round calendar schools. Using propensity score matching stratified by school type (elementary or middle) and urbanicity (city, suburb, town, and rural) I used a no- replacement, stratified intact group without caliper, one-to-one nearest neighbor matching procedure. The first step in the matching procedure is conducting a logistic regression with observable characteristics that could distinguish traditional and modified year-round calendar schools. Using the specification outlined in chapter three, and the Stata version 11 statistical software package, I generated propensity scores for all elementary and middle schools in North Carolina. See Table 2 for the logistic regression results.

**Region of common support.** One preliminary check to see whether proceeding with the matching process is appropriate is to graph the propensity scores to examine the amount of overlap between the treatment and comparison samples. Figure 1 depicts the propensity score graphs for treatment and comparison groups. The area of overlap on both graphs is referred to as the region of common support. If there are comparison units with similar propensity scores to the treatment units, then it is suitable to move forward to select the set of comparison schools. Figure 1 shows a good deal of overlap between traditional and modified year-round schools. There are a few schools with propensity scores above .2 that do not have as many potential comparison school matches. However, at each propensity score value below .2 there appears to be at least one comparison school to match each treatment school.

**Table 2. Logistic Regression for Propensity Score Matching: Growth Curve Analysis**

| Variable Label | Odds Ratio | Standard Error | P-value |
|---|---|---|---|
| Proportion of Teachers with less than 3 years experience (*100) | 1.023 | 0.014 | 0.105 |
| Proportion of Teachers with a supplemental masters degree (*100) | 1.013 | 0.017 | 0.442 |
| Proportion of Teachers with National Board Certification (*100) | 1.019 | 0.021 | 0.377 |
| Log of Average Daily Membership | 1.262 | 0.543 | 0.588 |
| City indicator | 0.663 | 0.261 | 0.296 |
| Suburb indicator | 1.258 | 0.198 | 0.562 |
| Town indicator | 1.954 | 0.766 | 0.088 |
| Percent of Asian students | 0.937 | 0.051 | 0.229 |
| Percent of black students | 1.043 | 0.010 | <0.001 |
| Percent of Hispanic students | 1.067 | 0.020 | <0.001 |
| Percent of multiracial students | 1.127 | 0.068 | 0.047 |
| Percent of American Indian students | 1.048 | 0.015 | 0.001 |
| Percent of students receiving free or reduced price lunch | 0.959 | 0.010 | <0.001 |
| Total per pupil expenditures | 1.011 | 0.012 | 0.347 |
| ABC status | 1.332 | 0.194 | 0.049 |
| Growth status | 0.838 | 0.218 | 0.497 |
| Percent of AYP target met | 0.975 | 0.013 | 0.058 |
| Suspension rate per 100 students | 0.982 | 0.012 | 0.127 |
| Violent acts rate per 1000 students | 0.989 | 0.032 | 0.741 |

The results of the propensity score matching generated 65 traditional calendar schools (out of 1,484) matched to 65 modified year-round calendar schools. Once the matches were generated, several checks were necessary to ensure the quality of the matches.

**Balancing tests.** First, I conducted a test in the difference of the means of the logits of the propensity scores between treatment and comparison groups. The value should be less than .05. With a difference of .003, the sample met the first balancing test. The second balancing test requires an examination of the ratio of the variances of the treated to comparison logits. Optimally, the value should fall between .05 and 2.0. Again, the quality of the matching procedure was supported as the ratio was .995.

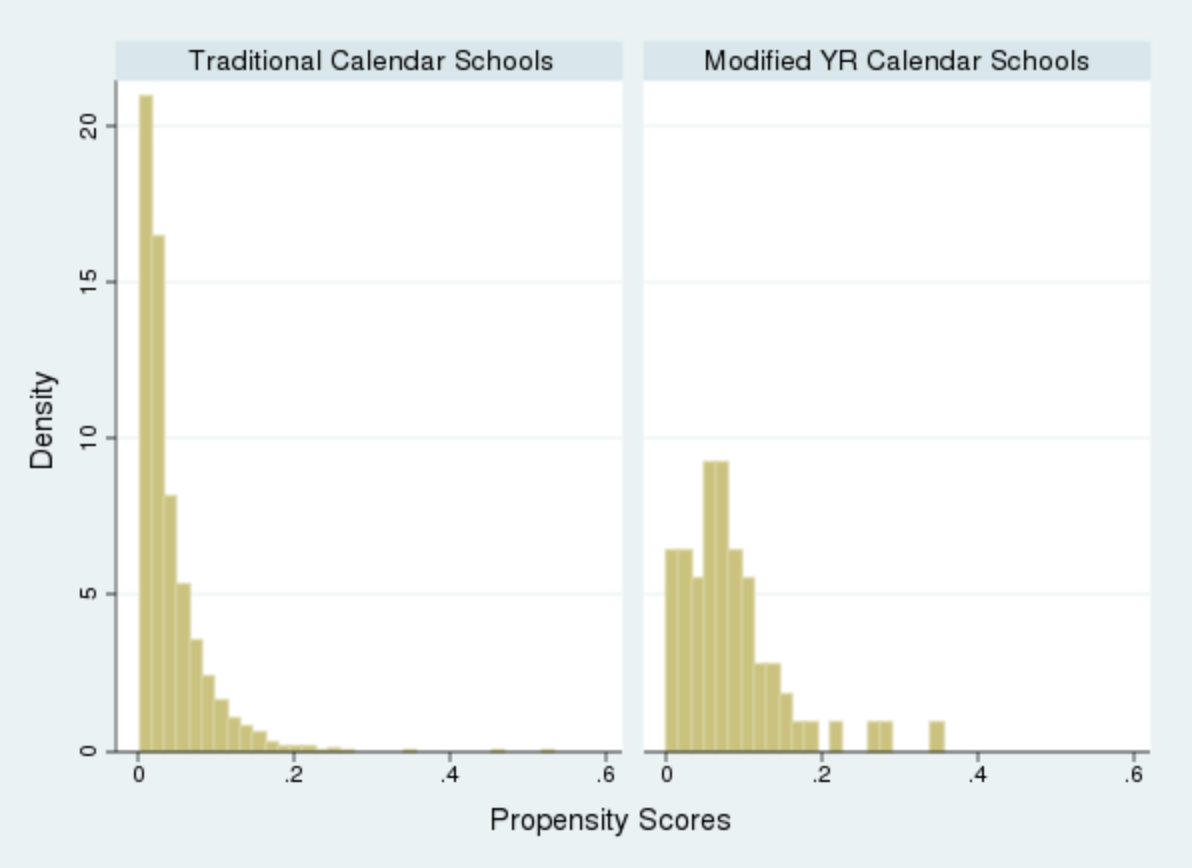


**Figure 1. Density of Propensity Scores in Traditional Calendar Schools and Modified Year-round Schools**

**Full vs. matched sample**. Additionally, it is important to look at a number of descriptors to see if the modified year-round calendar schools appear similar to the matched traditional calendar schools, or at the very least look more similar to the matched schools than the full sample. See Table 3 for descriptive statistics.

Table 3 shows that on average matched sample schools are more similar to modified year-round school compared to the full sample that includes all elementary and middle schools. There are a few variables where, on average, the matched sample is not more similar to treatment school: percent Asian, total per pupil expenditures, average teacher supplements, growth status, and the percent of AYP targets met. Of these variables, the difference was negligible except for average teacher supplements. The average teacher supplement difference between modified year-round

schools and the matched sample is larger than between the modified year round schools and the full sample. In sum, propensity score matching produced a matched sample that was more similar to the treatment group, modified year-round calendar schools, based on a number of characteristics.

**Table 3. Comparison of Matched and Full Traditional Calendar Samples to Modified Year-Round Schools**

| | Modified Year-Round School Sample ($n$ = 65) | | Matched Traditional Calendar School Sample ($n$ = 65) | | Full Traditional Calendar School Sample ($n$ = 1484) | |
|---|---|---|---|---|---|---|
| **Variable** | ***M*** | ***SD*** | ***M*** | ***SD*** | ***M*** | ***SD*** |
| Percent Asian | 2.04 | 3.58 | 2.22 | 2.16 | 1.92 | 2.99 |
| Percent Black | 36.74 | 26.71 | 35.42 | 25.07 | 30.17 | 25.38 |
| Percent Hispanic | 9.61 | 9.51 | 8.73 | 6.50 | 7.51 | 7.84 |
| Percent Multiracial | 3.40 | 1.89 | 3.33 | 2.38 | 2.88 | 2.11 |
| Percent American Indian | 2.77 | 8.82 | 3.53 | 15.74 | 1.43 | 7.19 |
| Percent white | 45.44 | 28.47 | 46.77 | 27.84 | 54.89 | 28.46 |
| Percent free/red. Lunch | 52.63 | 31.64 | 50.18 | 25.68 | 55.76 | 23.29 |
| Total per pupil expenditures | 78.27 | 14.58 | 76.64 | 17.61 | 77.39 | 18.18 |
| Suspension rate | 8.25 | 13.10 | 5.86 | 9.50 | 13.09 | 22.18 |
| Violent acts rate | 2.88 | 4.62 | 2.73 | 4.10 | 3.89 | 7.62 |
| Avg. teacher supplement | 32.74 | 15.01 | 125.39 | 727.98 | 95.99 | 573.76 |
| ABC status (1-6) | 4.06 | 1.68 | 4.23 | 1.58 | 3.77 | 1.59 |
| Growth status (0-2) | 0.98 | 0.74 | 1.08 | 0.74 | 0.91 | 0.78 |
| Percent AYP targets met | 92.10 | 13.24 | 94.52 | 8.70 | 94.13 | 10.35 |
| Avg. Daily membership | 6.21 | 2.71 | 6.66 | 3.03 | 5.55 | 2.35 |
| Percent of teachers with less than 3 yrs of experience | 25.85 | 11.12 | 25.54 | 11.66 | 23.43 | 10.86 |
| Percent of teachers with a supplemental masters | 19.84 | 7.90 | 19.44 | 8.99 | 19.20 | 8.41 |
| Percent of teachers with an NBC | 7.98 | 8.60 | 7.42 | 6.95 | 7.15 | 6.59 |
| Propensity score | 0.08 | 0.06 | 0.08 | 0.06 | 0.04 | 0.04 |

The final sample included all 65 middle and elementary modified year-round schools in operation during the baseline year. Additionally the final sample included 65 matched comparison schools. Comparison and modified year-round samples each contained 54 elementary schools and 11 middle schools. All schools in the sample were in operation and maintained a consistent calendar type for the entire study period.

**Difference-in-Differences Analysis**

The difference-in-differences analyses required a matched sample of schools similar to the 23 Wake County schools which converted to a modified year-round calendar in 2007-2008. Again I used the no- replacement, stratified intact group without caliper, one-to-one nearest neighbor matching procedure, stratified by school type (elementary or middle) and urbanicity (city, suburb, town, and rural) in Stata version 11. Using school level data from 2006-07, a matched Wake County sample was selected from all Wake County elementary schools to serve as a comparison group. As the preliminary step for performing propensity score matching, a logistic regression estimated the probability of conversion to a year-round calendar for every traditional calendar elementary school. The logistic regression contained a number of variables outlined in Chapter 3. Due to a small treatment sample (23 schools) and subsequently less variation than the state as whole, ABC status was broken down into indicator variables for several designations: School of distinction, School of progress, and no recognition. Table 4 displays the output for Wake Count logistic regression.

**Table 4. Logistic Regression for Propensity Score Matching: Difference-in-Differences Analysis**

| Variable | Odds Ratio | Standard Error | *p* |
|---|---|---|---|
| Average Daily Membership | 1.014 | 0.105 | 0.892 |
| Percent of Asian students | 0.886 | 0.073 | 0.142 |
| Percent of black students | 0.934 | 0.041 | 0.116 |
| Percent of Hispanic students | 1.043 | 0.078 | 0.576 |
| Percent of multiracial students | 1.291 | 0.358 | 0.356 |
| Percent of American Indian students | 0.414 | 0.423 | 0.388 |
| Percent of students receiving free or reduced price lunch | 1.014 | 0.043 | 0.746 |
| Total per pupil expenditures | 0.890 | 0.055 | 0.058 |
| School of Progress | 0.565 | 0.725 | 0.656 |
| School with No Recognition | 0.128 | 0.225 | 0.242 |
| School of Distinction | 0.202 | 0.223 | 0.148 |
| Growth status | 0.382 | 0.284 | 0.195 |

| Variable | Odds Ratio | Standard Error | *p* |
|---|---|---|---|
| Percent of AYP target met | 1.072 | 0.055 | 0.177 |
| Suspension rate per 100 students | 0.991 | 0.046 | 0.848 |
| Violent acts rate per 1000 students | 0.931 | 0.086 | 0.433 |
| Proportion of Teachers with less than 3 years experience (*100) | 1.018 | 0.034 | 0.602 |
| Proportion of Teachers with a supplemental masters degree (*100) | 0.990 | 0.052 | 0.851 |
| Proportion of Teachers with National Board Certification (*100) | 0.992 | 0.053 | 0.886 |

**Region of common support.** Again it was important to check the region of overlap. Histograms depicting the frequency of propensity scores for the possible comparison schools and treatment schools show some overlap, indicating that is it appropriate to precede with the propensity score matching procedure (see Figure 2).

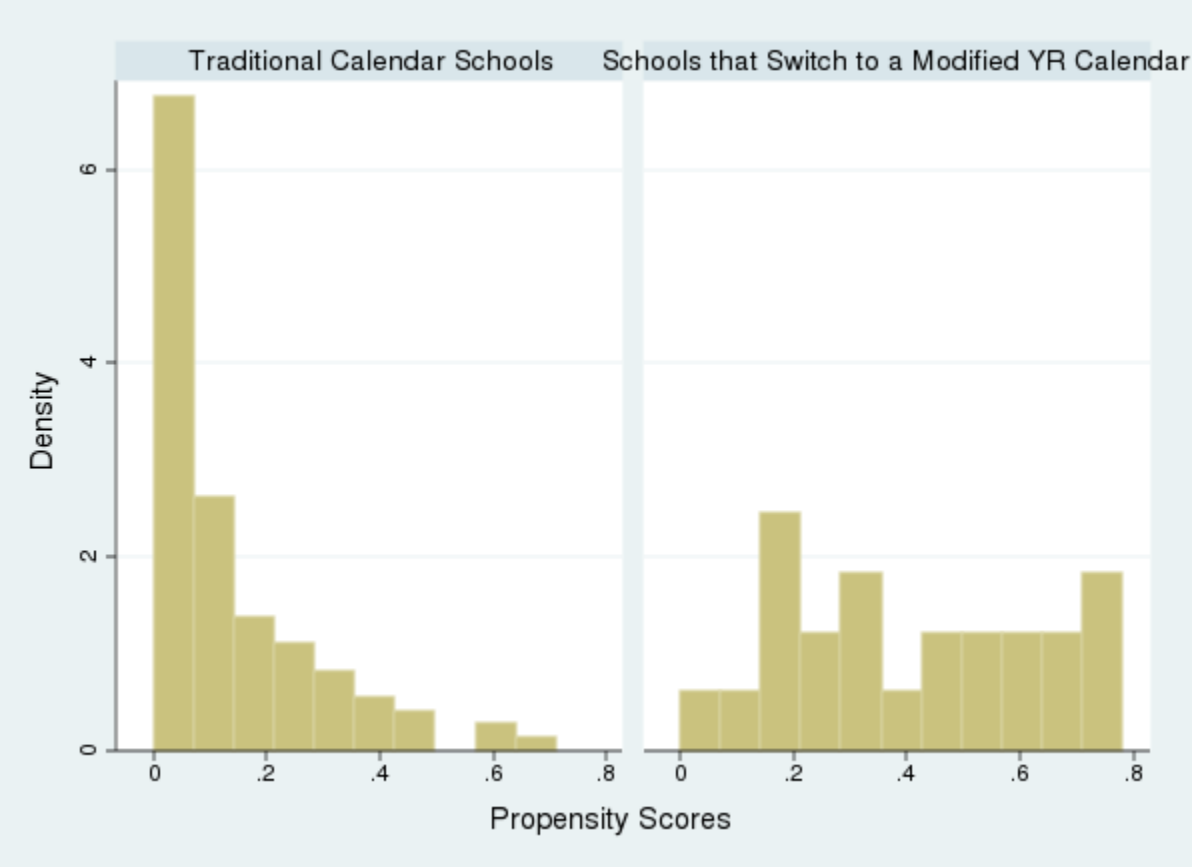


**Figure 2. Density of Propensity Scores in Traditional Calendar Wake County Schools and Wake County Schools that Switch to the Modified Year-round**

**Balancing tests.** The balance test of the difference of the means of the logits of the propensity scores between treatment and comparison groups yielded a value of .245, somewhat higher than the ideal range which less than .05. The second balancing test compares the ratio of the variances of the treated group to comparison group logits. With a value of .381, the second balancing test result falls inside of the optimal range; between .05 and 2.0. Since the quality of the matching procedure is not supported by both of the balancing tests,a series of alternative matching procedures were conducted including radius matching, full matching, and making changes in the stratification specification. The original matching specification yielded balancing test values that most closely resembled the ideal values and therefore the original matching specification was deemed the most appropriate matching procedure for the current data.

**Full vs. matched sample.** Finally, an examination of the matched and treatment sample a number of observable characteristics provides support for integrity of the matching process. See Table 5 for a description of the schools which were converted (treatment schools) to the matched sample and the full sample of elementary schools in Wake County.

**Table 5. Comparison of Matched and Full Traditional Calendar Samples to Modified Year-Round Schools**

| | **Modified Year-Round School Sample (*n* = 23)** | | **Matched Traditional Calendar School Sample (*n* = 23)** | | **Full Traditional Calendar School Sample (*n* = 102)** | |
|---|---|---|---|---|---|---|
| **Variable** | ***M*** | ***SD*** | ***M*** | ***SD*** | ***M*** | ***SD*** |
| Percent Asian | 4.80 | 4.66 | 3.21 | 2.71 | 4.80 | 5.62 |
| Percent Black | 22.18 | 15.82 | 28.54 | 13.99 | 32.24 | 14.94 |
| Percent Hispanic | 12.16 | 8.70 | 13.08 | 6.34 | 10.68 | 6.27 |
| Percent Multiracial | 4.24 | 1.36 | 4.80 | 1.31 | 4.14 | 1.49 |
| Percent American Indian | 0.28 | 0.26 | 0.24 | 0.28 | 0.29 | 0.30 |
| Percent white | 56.34 | 20.97 | 50.14 | 17.76 | 47.85 | 15.72 |
| Percent freed/red. Lunch | 32.58 | 22.12 | 40.42 | 17.22 | 38.60 | 17.80 |
| Total per pupil expenditures | 75.16 | 6.52 | 77.26 | 7.07 | 81.66 | 15.24 |
| Suspension rate | 8.00 | 14.10 | 7.88 | 7.10 | 17.08 | 16.54 |
| Violent acts rate | 3.49 | 7.85 | 4.10 | 5.02 | 8.15 | 8.37 |
| Avg. teacher supplement | 58.59 | 0.00 | 58.59 | 0.00 | 58.59 | 0.00 |
| ABC status (1-6) | 3.96 | 1.27 | 3.52 | .73 | 3.32 | 0.96 |

| | Modified Year-Round School Sample ($n$ = 23) | | Matched Traditional Calendar School Sample ($n$ = 23) | | Full Traditional Calendar School Sample ($n$ = 102) | |
|---|---|---|---|---|---|---|
| **Variable** | ***M*** | ***SD*** | ***M*** | ***SD*** | ***M*** | ***SD*** |
| Growth status (0-2) | 1.22 | 0.60 | 1.17 | .57 | 1.18 | 0.67 |
| Percent AYP targets met | 95.92 | 7.25 | 95.59 | 5.24 | 90.02 | 13.21 |
| Avg. Daily membership | 8.18 | 1.60 | 7.26 | 2.38 | 8.99 | 5.45 |
| Percent of teachers with less than 3 yrs of experience | 26.83 | 10.64 | 25.04 | 10.74 | 23.36 | 9.14 |
| Percent of teachers with a supplemental masters | 17.18 | 5.37 | 17.19 | 7.32 | 20.38 | 6.76 |
| Percent of teachers with an NBC | 9.75 | 6.65 | 9.17 | 5.57 | 10.33 | 5.66 |
| Propensity score | 0.41 | 0.22 | 0.30 | 0.18 | 0.13 | 0.15 |

The matched schools do appear more similar to the schools that switched than the full sample of schools on most characteristics. The only characteristics in which the matched sample is more dissimilar than the full sample to the treatment group are percent Asian, percent multiracial, percent American Indian, the growth status, and school size, although the differences were relatively small. Ultimately, the 23 matched schools seem, on average, more similar to the treatment schools when compared to all elementary and middle schools in Wake County.

**Falsification test.** All of the schools in the sample were on a traditional calendar in the baseline matching year of 2006-07. Running a model to predict the dependent variable (in this case teacher retention and per pupil expenditure variables) with the treatment assignment variable (labeled "switch") in the year before the switch occurred will tell us whether the two groups differed prior to the calendar reform. In essence, if the matching was successful, it will have produced a comparison group that is similar enough to the treatment group that the future decision for some of the schools to convert to a modified year-round calendar will not be significant. See Table 6 for falsification test results run using Stata 11.

As Table 6 indicates, treatment assignment is not significant in predicting teacher turnover or school expenditures, with the exception of per pupil school maintenance and utilities expenditures.

Analyses of per pupil school maintenance and utilities expenditures may not reveal differences entirely due to the calendar switch since schools that eventually switch to a modified year-round calendar already spend significantly less in the baseline year than schools that remain on a traditional calendar. All other outcome variables were not predicted based on assignment to treatment or comparison conditions and provides additional support for the integrity of the matching.

**Table 6. Falsification Tests Results**

| Outcomes | Estimate on Switch Variable | Standard Error |
|---|---|---|
| Between year teacher retention | -0.062 | 0.042 |
| Within Year teacher retention | -0.0022 | 0.010 |
| Per pupil spending | -210.032 | 200.539 |
| Per pupil spending in regular instruction | 10.754 | 98.912 |
| Per pupil spending in special instruction | -23.752 | 51.347 |
| Per pupil spending in student services | -33.337 | 17.914 |
| Per pupil spending in instructional support | 7.460 | 11.430 |
| Per pupil spending in transportation | -1.426 | 1.110 |
| Per pupil spending in school maintenance & utilities | -46.520* | 22.126 |
| Per pupil spending in school leadership | -25.091 | 29.575 |
| Per pupil spending in capital outlay | 16.462 | 23.347 |

* $p < .05$

The final sample included all 23 middle and elementary Wake County schools that switched from a traditional calendar in 2006-07 to a modified year-round calendar in 2007-08. Additionally the final sample included 23 matched comparison Wake County schools. Comparison and modified year-round samples each contained 20 elementary schools and 3 middle schools.

## Descriptive Statistics

### Growth Curve Analysis

The sample for the growth curve analyses contained 130 schools, 65 schools with a modified year-round calendar and 65 schools with a traditional school calendar. Of the 130 schools there were

22 middle schools and 108 elementary schools. The sample included students in grades 3-8. Tables 6-8 in Appendix B provide a description of the school and student level characteristics for all samples used in the analyses; math achievement, reading achievement, and days absent. Missing data was minimal due to requirements from the Department of Public Instruction. Since missing data was minimal, cases with missing data (< 5% of cases) were excluded from the analysis. Outliers were accounted for in two ways. First, students with test score values considered out of range by DPI were changed to missing and excluded from the analyses (< 1% of cases). School level outliers are handled through the matching procedure, where treatment schools are matched to comparison schools based on their similarity in observed characteristics.

**Fixed Effects**

The student fixed effects analysis involved students who attended Wake County schools that went through the conversion from a traditional calendar to a modified year-round calendar in 2007-2008. The analysis sample includes students who went to traditional calendar schools and then switched to one of the modified year-round calendar schools in 2007-2008. Only students who switched from a traditional to a year-round calendar contribute to the calendar effect estimate. See tables 9-11 in Appendix B for the sample descriptions for math, reading, and days absent models. Missing data was minimal due to requirements from the Department of Public Instruction. Since missing data was minimal, cases with missing data (< 5% of cases) were excluded from the analysis. Students with test score values considered out of range by DPI were changed to missing and excluded from the analyses (< 1% of cases).

**Difference-in-Differences Analysis**

Schools are the unit of analysis in the difference-in-differences sample. The sample includes 46 schools: 23 schools that converted to the modified year-round calendar and 23 matched schools. Table 12 in Appendix B displays descriptive information on the full difference-in-differences sample over time. Missing data was minimal due to requirements from the Department of Public Instruction.

Schools that were missing data (< 5%) were excluded from the propensity score matching and thus excluded from subsequent analyses.

Two separate statistical packages were used in the analysis. Stata version 11 was used for all fixed effects and difference-in-difference models. SAS 9.2 was used for modeling growth curves. The following section details the results of each analysis, organized by research question.

## Inferential Statistics

### Research Question 1. Do Students Perform Better on Achievement Tests under a Year-Round Modified Calendar compared to a Traditional Calendar?

To address Research Question 1 two different methodologies were employed: growth curves analysis and student fixed effects. The results of each analysis strategy are discussed below.

**Fixed effects.** In the student fixed effects model, each student's achievement score under the traditional calendar is compared to the same student's achievement score after the school converted to a modified year-round calendar. Students demonstrated a trend of higher performance in math and reading under the modified year-round calendar, but the result was not significant. Additional test for robustness included separate analysis of elementary and middle schools and limiting the time frame. All robustness tests yielded similar results. Table 7 displays the student fixed effects model results.

**Table 7. Student Fixed Effects Results for Math and Reading Achievement**

| | Math | | Reading | |
|---|---|---|---|---|
| **Variables** | **Estimate** | **Standard Error** | **Estimate** | **Standard Error** |
| Constant | -0.59 | 0.41 | 0.34 | 0.30 |
| Modified Year-Round Calendar | 0.03 | 0.03 | 0.02 | 0.02 |
| Asian Mean | 0.00 | 0.01 | 0.00 | 0.01 |
| Black Mean | 0.01 | 0.00 | 0.01* | 0.00 |
| Hispanic Mean | -0.01 | 0.01 | -0.01* | 0.01 |
| Multiracial Mean | -0.02 | 0.02 | 0.00 | 0.01 |
| American Indian Mean | 0.00 | 0.06 | 0.02 | 0.05 |
| Percent Free or Reduced Price Lunch | 0.00 | 0.00 | 0.00 | 0.00 |
| School Size | 0.07 | 0.08 | -0.10+ | 0.06 |
| School Size Squared | 0.00 | 0.01 | 0.01 | 0.00 |
| Suspension Rate (per 100) | 0.00 | 0.00 | 0.01* | 0.00 |
| Violent Acts Rate (per 1000) | -0.01* | 0.00 | 0.00 | 0.01 |

| | Math | | Reading | |
|---|---|---|---|---|
| **Variables** | **Estimate** | **Standard Error** | **Estimate** | **Standard Error** |
| Teacher Supplements | 0.01 | 0.00 | 0.01 | 0.00 |
| Number of Students in Class | 0.00 | 0.00 | 0.00 | 0.00 |
| Average Peer Dispersion | 0.03 | 0.06 | -0.05 | 0.05 |
| Prior Achievement Score (standardized) | -0.19** | 0.02 | -0.23** | 0.03 |
| Average Peer Achievement Score (standardized) | -0.03 | 0.03 | 0.01 | 0.03 |
| Asian | 0.13 | 0.18 | -0.05 | 0.21 |
| Black | -0.09 | 0.21 | 0.08 | 0.23 |
| Hispanic | 0.13 | 0.32 | -0.35 | 0.24 |
| Multiracial | 0.03 | 0.14 | -0.17 | 0.21 |
| American Indian | 0.14* | 0.05 | -0.08 | 0.05 |
| Male | 0.20 | 0.25 | 0.01 | 0.23 |
| Days absent | 0.00* | 0.00 | 0.00* | 0.00 |
| Move in Year | -0.01 | 0.03 | 0.00 | 0.05 |
| Underage | 0.05 | 0.04 | 0.92** | 0.03 |
| Overage | 0.64** | 0.13 | 0.70** | 0.16 |
| Gifted | -0.07** | 0.02 | 0.02 | 0.02 |
| Special Needs | 0.06 | 0.05 | 0.01 | 0.07 |
| Is LEP | 0.17 | 0.12 | -0.06 | 0.10 |
| Was LEP | 0.29** | 0.10 | 0.01 | 0.11 |
| Free Lunch | 0.01 | 0.04 | 0.05 | 0.05 |
| Reduced Price Lunch | -0.02 | 0.04 | 0.03 | 0.05 |
| Free or Reduced Price Lunch Missing | 0.39 | 0.19 | 0.37 | 0.27 |

** $p < .01$
* $p < .05$
+ $p < .10$

**Growth curve analysis.** The growth curve analysis examines student achievement over 4 time points from 2005-06 to 2008-09. Comparing the rate of change in student outcomes, I test whether students in a modified year-round calendar outperform students in traditional calendar schools. First, I use a simplified model to examine the proportion of variance (also known as the Intraclass Correlation Coefficient) which signifies the amount of variance at each of the three levels; time, student, and school. In the reading model, the proportion of variance within a student ($\sigma^2$) is calculated in equation 6, across students ($\tau_\pi$) in equation 7, and across schools ($\tau_\beta$) in equation 8.

(6) $\rho_{time} = \sigma^2 / (\sigma^2 + \tau_\pi + \tau_\beta) = .175$

(7) $\rho_{student} = \tau_\pi / (\sigma^2 + \tau_\pi + \tau_\beta) = .670$

(8) $\rho_{school} = \tau_\beta / (\sigma^2 + \tau_\pi + \tau_\beta) = .155$

The equations about indicate that about 17% of the variation in reading scores is due to changes within an individual across time. Sixty-seven percent of the variation in reading scores is due to differences between individuals within schools. Only about 16% of the variance in reading achievement is due to differences between schools. In the math model, the proportion of variance within a student ($\sigma^2$) is calculated in equation 9, across students ($\tau_\pi$) in equation 10, and across schools ($\tau_\beta$) in equation 11.

(9) $\rho_{time} = \sigma^2 / (\sigma^2 + \tau_\pi + \tau_\beta) = .159$

(10) $\rho_{student} = \tau_\pi / (\sigma^2 + \tau_\pi + \tau_\beta) = .469$

(11) $\rho_{school} = \tau_\beta / (\sigma^2 + \tau_\pi + \tau_\beta) = .172$

The variances in mathematic scores are portioned in a similar fashion. Fifteen percent of the variance in mathematics scores is due to differences over time, while differences across students capture about 47% of the variance. Seventeen percent of the variance in mathematics scores is due to school differences.

Next, the full models for math and reading are estimated. The estimate of modified year-round calendar status was not significant for math or reading. See model results in Table 8.

**Table 8. Growth Curve Model Results for Math and Reading Achievement**

| | Math | | Reading | |
|---|---|---|---|---|
| **Variables** | **Estimate** | **Standard Error** | **Estimate** | **Standard Error** |
| Constant | 0.071 | 0.053 | 0.218** | 0.047 |
| Time | -0.015** | 0.002 | -0.005* | 0.002 |
| Modified Year-Round Calendar | -0.021 | 0.018 | -0.012 | 0.012 |
| Prior Achievement Score (standardized) | 0.690** | 0.002 | 0.711** | 0.002 |
| Asian | 0.156** | 0.009 | -0.047** | 0.009 |
| Black | -0.188** | 0.004 | -0.130** | 0.004 |
| Hispanic | -0.029** | 0.007 | -0.055** | 0.007 |
| Multiracial | -0.067** | 0.008 | -0.024** | 0.008 |
| American Indian | -0.116** | 0.013 | -0.130** | 0.013 |
| Grade | 0.004** | 0.002 | 0.007** | 0.002 |
| Male | 0.088** | 0.003 | -0.079** | 0.003 |
| Is LEP | -0.021** | 0.008 | -0.213** | 0.008 |

| | Math | | Reading | |
|---|---|---|---|---|
| Variables | Estimate | Standard Error | Estimate | Standard Error |
| Was LEP | 0.055** | 0.011 | 0.000 | 0.012 |
| Gifted | 0.274** | 0.004 | 0.166** | 0.004 |
| Special Needs | -0.096** | 0.005 | -0.172** | 0.005 |
| Free or Reduced Price Lunch | -0.054** | 0.004 | -0.078** | 0.004 |
| Free or Reduced Price Lunch Missing | -0.006 | 0.006 | -0.019** | 0.006 |
| Underage | 0.067** | 0.013 | 0.049** | 0.013 |
| Overage | -0.116** | 0.004 | -0.091** | 0.004 |
| Days absent | -0.009** | 0.000 | -0.003** | 0.000 |
| Forced Move | 0.005 | 0.005 | 0.015** | 0.005 |
| Move in Year | -0.092** | 0.006 | -0.032** | 0.006 |
| Per pupil spending (per $100) | 0.002** | 0.000 | 0.000 | 0.000 |
| Percent of teachers with National Board Certification | 0.139** | 0.052 | 0.093+ | 0.050 |
| Percent of teachers with an Advanced Degree | -0.025 | 0.040 | -0.158** | 0.038 |
| Percent of teachers with 3 or less years of Experience | 0.011 | 0.030 | -0.116** | 0.030 |
| Asian Mean | -0.002 | 0.002 | -0.002 | 0.002 |
| Black Mean | -0.003** | 0.000 | -0.001** | 0.000 |
| Hispanic Mean | -0.003** | 0.001 | -0.002** | 0.001 |
| Multiracial Mean | -0.002 | 0.002 | 0.005** | 0.002 |
| American Indian Mean | -0.002* | 0.001 | -0.002** | 0.001 |
| Percent Free or Reduced Price Lunch | 0.001** | 0.000 | 0.001+ | 0.000 |
| School Size | -0.001 | 0.006 | 0.000 | 0.005 |
| School Size Squared | 0.000 | 0.000 | 0.000 | 0.000 |
| Suspension Rate (per 100) | -0.001** | 0.000 | -0.001* | 0.000 |
| Violent Acts Rate (per 1000) | 0.001** | 0.000 | 0.000 | 0.000 |
| Teacher Supplements | 0.000 | 0.000 | 0.001** | 0.000 |
| Propensity Score | 0.212 | 0.179 | 0.065 | 0.127 |

** $p < .01$
* $p < .05$
+ $p < .10$

**Research Question 2. Are Student Absenteeism Rates Lower under a Year-Round Modified Calendar compared to a Traditional Calendar?**

Student absenteeism was addressed through student fixed effects.

**Fixed effects.** According to the student fixed effects results, students do exhibit lower rates of student absenteeism when enrolled under a modified year-round calendar compared to a traditional calendar ($\beta(hat) = -1.35$ , $p < .01$). A coefficient of -1.35 on the modified year-round calendar is

interpreted as follows: on average students on a modified year-round calendar are absent 1.35 days less than when students are on a traditional calendar, holding all other variables constant. (Note: Weighting by reading or math observations yielded similar results. Reported results are weighted by math observations.) Additional test for robustness included separate analysis of elementary and middle schools and limiting the time frame. All robustness tests yielded similar results. See Table 9 for model results.

**Table 9. Student Fixed Effects Results for Student Absenteeism**

| Variables | Estimate | Standard Error |
|---|---|---|
| Constant | -7.53** | 1.85 |
| Modified Year-Round Calendar | -1.35** | 0.17 |
| Asian Mean | 0.03 | 0.04 |
| Black Mean | 0.04 | 0.03 |
| Hispanic Mean | 0.00 | 0.05 |
| Multiracial Mean | 0.10 | 0.08 |
| American Indian Mean | 0.30 | 0.27 |
| Percent Free or Reduced Price Lunch | -0.03* | 0.01 |
| School Size | 0.49 | 0.41 |
| School Size Squared | -0.04 | 0.02 |
| Suspension Rate (per 100) | -0.04 | 0.02 |
| Violent Acts Rate (per 1000) | 0.07** | 0.02 |
| Teacher Supplements | 0.22** | 0.02 |
| Number of Students in Class | -0.03 | 0.03 |
| Average Peer Dispersion | 0.32 | 0.36 |
| Prior Achievement Score (standardized) | -0.23* | 0.11 |
| Average Peer Achievement Score (standardized) | 0.07 | 0.20 |
| Asian | 1.63 | 2.05 |
| Black | 0.76 | 3.29 |
| Hispanic | -0.15 | 1.69 |
| Multiracial | 0.64 | 1.69 |
| American Indian | 4.44** | 0.27 |
| Male | 0.47 | 1.06 |
| Move in Year | -2.98** | 0.55 |
| Underage | 2.26** | 0.21 |
| Overage | 0.39 | 1.27 |
| Gifted | -0.29* | 0.12 |
| Special Needs | -0.59 | 0.38 |
| Is LEP | -0.41 | 0.91 |
| Was LEP | -0.80 | 0.82 |
| Free Lunch | -0.15 | 0.38 |
| Reduced Price Lunch | -0.03 | 0.43 |
| Free or Reduced Price Lunch Missing | -0.76 | 1.18 |

** $p < .01$
* $p < .05$

**Research Question 3. Do Certain Student Subgroups Perform Better on Achievement Tests under a Year-Round Modified Calendar compared to a Traditional Calendar?**

Research Question 3 was only able to be addressed through the growth curve analysis. The interaction terms reflect the effect of the modified year-round calendar on specific student subgroups. The growth curve analysis shows a positive effect of the modified year-round calendar in reading for students who qualify for free or reduced price lunch, ($\hat{\beta}$ = 0.023, $p < .05$). In mathematics, there is a general positive trend for students on free or reduced priced lunch who attend a modified year-round calendar, but the results are not significant at the .05 level, ($\hat{\beta}$ = 0.012, $p < .10$). There is a negative trend for modified year-round calendar students who are limited English proficient in mathematics ($\hat{\beta}$ = -0.021, $p < .10$) and a statistically significant negative effect of the modified year-round calendar for current ($\hat{\beta}$ = -0.040, $p < .01$) and past ($\hat{\beta}$ = -0.064, $p < .01$) English language learners in reading. There is a positive effect of the modified year-round calendar on reading achievement for students with disabilities ( $\hat{\beta}$ = 0.019, $p < .01$). Math and reading results are presented in Table 10. [Note that robustness checks confirmed a negative coefficient on students who are limited English proficient both past and present.]

**Table 10. Growth Curve Interactions Results for Student Achievement**

| | Math | | Reading | |
|---|---|---|---|---|
| **Variables** | **Estimate** | **Standard Error** | **Estimate** | **Standard Error** |
| Constant | 0.073 | 0.053 | 0.224** | 0.047 |
| Wave | -0.015** | 0.002 | -0.005* | 0.002 |
| Modified Year-Round Calendar | -0.023 | 0.018 | -0.020+ | 0.012 |
| Modified Year-Round Calendar* Free or Reduced Price Lunch | 0.012+ | 0.006 | 0.023** | 0.006 |
| Modified Year-Round Calendar * Is Limited English Proficient | -0.021+ | 0.012 | -0.040** | 0.012 |
| Modified Year-Round Calendar * Was Limited English Proficient | -0.019 | 0.021 | -0.064** | 0.021 |
| Modified Year-Round Calendar * Student with Special Needs | -0.004 | 0.009 | 0.019** | 0.009 |
| Prior Achievement Score (standardized) | 0.690** | 0.002 | 0.711** | 0.002 |
| Asian | 0.156** | 0.009 | -0.048** | 0.009 |
| Black | -0.187** | 0.004 | -0.130** | 0.004 |

| | Math | | Reading | |
|---|---|---|---|---|
| Variables | Estimate | Standard Error | Estimate | Standard Error |
| Hispanic | -0.029** | 0.007 | -0.054** | 0.007 |
| Multiracial | -0.067** | 0.008 | -0.024** | 0.008 |
| American Indian | -0.117** | 0.013 | -0.132** | 0.013 |
| Grade | 0.004* | 0.002 | 0.007** | 0.002 |
| Male | 0.088** | 0.003 | -0.079** | 0.003 |
| Is LEP | -0.012 | 0.010 | -0.195** | 0.010 |
| Was LEP | 0.063** | 0.015 | 0.030* | 0.015 |
| Gifted | 0.274** | 0.004 | 0.166** | 0.004 |
| Special Needs | -0.094** | 0.006 | -0.182** | 0.007 |
| Free or Reduced Price Lunch | -0.059** | 0.005 | -0.089** | 0.005 |
| Free or Reduced Price Lunch Missing | -0.005 | 0.006 | -0.018** | 0.006 |
| Underage | 0.066** | 0.013 | 0.049** | 0.013 |
| Overage | -0.116** | 0.004 | -0.091** | 0.004 |
| Days Absent | -0.009** | 0.000 | -0.003** | 0.000 |
| Forced Move | 0.005 | 0.005 | 0.015** | 0.005 |
| Move in Year | -0.092** | 0.006 | -0.032** | 0.006 |
| Per pupil spending (per $100) | 0.002** | 0.000 | 0.000 | 0.000 |
| Percent of teachers with National Board Certification | 0.140** | 0.052 | 0.095+ | 0.050 |
| Percent of teachers with an Advanced Degree | -0.026 | 0.040 | -0.161** | 0.038 |
| Percent of teachers with 3 or less years of Experience | 0.011 | 0.030 | -0.116** | 0.030 |
| Asian Mean | -0.002 | 0.002 | -0.001 | 0.002 |
| Black Mean | -0.003** | 0.000 | -0.001** | 0.000 |
| Hispanic Mean | -0.003** | 0.001 | -0.002** | 0.001 |
| Multiracial Mean | -0.002 | 0.002 | 0.004** | 0.002 |
| American Indian Mean | -0.002* | 0.001 | -0.002** | 0.001 |
| Percent Free or Reduced Price Lunch | 0.001** | 0.000 | 0.001+ | 0.000 |
| School Size | -0.001 | 0.006 | -0.001 | 0.005 |
| School Size Squared | 0.000+ | 0.000 | 0.000 | 0.000 |
| Suspension Rate (per 100) | -0.001** | 0.000 | -0.001* | 0.000 |
| Violent Acts Rate (per 1000) | 0.001* | 0.000 | 0.000 | 0.000 |
| Teacher Supplements | 0.000 | 0.000 | 0.001** | 0.000 |
| Propensity Score | 0.213 | 0.179 | 0.065 | 0.127 |

** $p < .01$
* $p < .05$
+ $p < .10$

**Research Question 4. Is There Less Teacher Turnover In Year-Round Modified Calendar Schools Compared to Traditional Calendar Schools?**

The difference-in-differences methodology is utilized for examining changes in teacher retention based on calendar type. The model uses three waves of data; 2006-07 (pre-conversion) and 2007-08 and 2008-09 (post-conversion). Recall that the difference-in-differences design provides an estimate of the differences in teacher retention pre and post calendar conversion, netting out the effect of time. Thus, interaction terms indicate the modified year-round calendar and the post-conversion year. Two types of teacher retention were analyzed; retention between years and retention with an academic year. The modified year-round calendar increases teacher retention compared to the traditional calendar ($\beta(hat) = 0.084$, $p < .05$) in 2008-09. However, retention within the year appears unaffected by calendar type. See Table 11 for teacher retention model results.

**Table 11. Difference-in-differences Results for Teacher Retention**

| Variables | Estimate | Standard Error |
|---|---|---|
| *Between Year Teacher Retention* | | |
| Constant | 0.918** | 0.021 |
| School Year 2007-08 | -0.034 | 0.030 |
| School Year 2008-09 | -0.028 | 0.030 |
| Converted Calendar School | -0.062* | 0.030 |
| Converted Calendar School * School Year 2007-08 | 0.037 | 0.042 |
| Converted Calendar School * School Year 2008-09 | 0.084* | 0.042 |
| *Within Year Teacher Retention* | | |
| Constant | 0.974** | 0.005 |
| School Year 2007-08 | -0.005 | 0.007 |
| School Year 2008-09 | 0.010 | 0.007 |
| Converted Calendar School | -0.002 | 0.007 |
| Converted Calendar School * School Year 2007-08 | 0.013 | 0.011 |
| Converted Calendar School * School Year 2008-09 | 0.004 | 0.011 |

** $p < .01$
* $p < .05$
+ $p < .10$

**Research Question 5. Do Year-Round Modified Calendar Schools Spend Less Money Per Student than Schools Operating under a Traditional Calendar? And If So, What Expenditure Categories Differ?**

Again, the difference-in-differences design provided the structure for the expenditure and calendar type analysis. The modified year-round calendar provided only one effect on expenditures. Per pupil spending on student services actually increased in 2008-09 under a year-round calendar compared to a traditional calendar, netting out the effects of time ($\beta$(hat) = -62.299, $p < .05$). See Table 12 for all expenditure model results.

**Table 12. Difference-in-differences Results for Expenditures**

| Variable (*Dependent Variable*) | Estimate | Standard Error |
|---|---|---|
| *Per Pupil Spending* | | |
| Constant | 7726.354** | 141.275 |
| School Year 2007-08 | 27.034 | 199.794 |
| School Year 2008-09 | 541.666** | 199.794 |
| Converted Calendar School | -210.032 | 199.794 |
| Converted Calendar School * School Year 2007-08 | 374.943 | 282.551 |
| Converted Calendar School * School Year 2008-09 | 45.331 | 282.551 |
| *Per Pupil Spending in Capital Outlay* | | |
| Constant | 11.540 | 13.199 |
| School Year 2007-08 | 17.783 | 18.666 |
| School Year 2008-09 | 8.069 | 18.666 |
| Converted Calendar School | 16.462 | 18.666 |
| Converted Calendar School * School Year 2007-08 | -41.894 | 26.398 |
| Converted Calendar School * School Year 2008-09 | -28.127 | 26.398 |
| *Per Pupil Spending in Regular Instruction* | | |
| Constant | 3817.369** | 78.148 |
| School Year 2007-08 | 294.793 | 110.517 |
| School Year 2008-09 | 514.592** | 110.517 |
| Converted Calendar School | 10.754 | 110.517 |
| Converted Calendar School * School Year 2007-08 | 70.424 | 156.295 |
| Converted Calendar School * School Year 2008-09 | -136.243 | 156.295 |
| *Per Pupil Spending in Special Instruction* | | |
| Constant | 760.403** | 42.356 |
| School Year 2007-08 | 229.954** | 59.901 |
| School Year 2008-09 | 312.234** | 59.901 |
| Converted Calendar School | -23.752 | 59.901 |
| Converted Calendar School * School Year 2007-08 | 67.565 | 84.712 |
| Converted Calendar School * School Year 2008-09 | 22.135 | 84.712 |

| Variable (*Dependent Variable*) | Estimate | Standard Error |
|---|---|---|
| *Per Pupil Spending in Student Services* | | |
| Constant | 414.145** | 14.800 |
| School Year 2007-08 | -78.172** | 20.930 |
| School Year 2008-09 | -75.184** | 20.930 |
| Converted Calendar School | -33.337 | 20.930 |
| Converted Calendar School * School Year 2007-08 | 62.299* | 29.600 |
| Converted Calendar School * School Year 2008-09 | 47.544 | 29.600 |
| *Per Pupil Spending in Instructional Support* | | |
| Constant | 208.745** | 9.956 |
| School Year 2007-08 | 197.436** | 14.080 |
| School Year 2008-09 | 202.901** | 14.080 |
| Converted Calendar School | 7.460 | 14.080 |
| Converted Calendar School * School Year 2007-08 | 23.269 | 19.912 |
| Converted Calendar School * School Year 2008-09 | 8.811 | 19.912 |
| *Per Pupil Spending in Transportation* | | |
| Constant | 219.678** | 0.860 |
| School Year 2007-08 | 14.413** | 1.217 |
| School Year 2008-09 | 10.302** | 1.217 |
| Converted Calendar School | -1.425 | 1.217 |
| Converted Calendar School * School Year 2007-08 | 0.972 | 1.721 |
| Converted Calendar School * School Year 2008-09 | 1.601 | 1.721 |
| *Per Pupil Spending in Maintenance* | | |
| Constant | 579.508** | 16.373 |
| School Year 2007-08 | -13.933 | 23.155 |
| School Year 2008-09 | -12.676 | 23.155 |
| Converted Calendar School | -46.519* | 23.155 |
| Converted Calendar School * School Year 2007-08 | 2.682 | 32.747 |
| Converted Calendar School * School Year 2008-09 | -10.146 | 32.747 |
| *Per Pupil Spending in School Leadership* | | |
| Constant | 486.923** | 18.185 |
| School Year 2007-08 | -13.642 | 25.717 |
| School Year 2008-09 | -1.396 | 25.717 |
| Converted Calendar School | 31.244 | 25.717 |
| Converted Calendar School * School Year 2007-08 | -10.041 | 36.367 |
| Converted Calendar School * School Year 2008-09 | -14.514 | 36.367 |

** $p < .01$
* $p < .05$

## Summary and Conclusion

Overall there is mixed support for the modified year-round calendar. In terms of student achievement, the student fixed effects models and the growth curve models do not support a direct link between calendar type and student achievement. However, the growth curve models allowed me

to probe interactions between student subgroups and calendar type. In these growth models, the modified year-round calendar was significant and positively associated with higher student achievement for students on free or reduced-price lunch and for students with disabilities. Interestingly, the modified year-round calendar was associated with lower reading achievement for students who are or were classified as Limited English proficient. The student fixed-effects model showed a link between the modified year-round calendar and fewer days absent. However, the Wake County sample is small which makes generalizations more tenuous.

There are some indications that teacher turnover may be impacted by calendar type. There are lower rates of between year teacher turnover in converted calendar Wake County schools for the second year of the modified year-round calendar implementation (2008-09). Within year teacher turnover was unrelated to calendar type.

Lastly, there is little evidence that calendar change impacted school spending. In fact, spending was higher for student services in converted calendar schools (net secular changes). Note that coefficients for the modified year-round calendar effect may appear small. In fact, the size of the coefficients is consistent with similar effect sizes found in other educational policy research (Boyd, Grossman, Lankford, Loeb, & Wycoff, 2006; Clotfelter, Ladd, & Vigdor, 2007; Ferguson & Ladd, 1996). In the final chapter, Chapter 6, the results are examined in detail along with study implications, limitations, and areas for further research.

# CHAPTER 6

# DISCUSSION

Since the inception of American schools the arrangement of schooling time has served to meet the needs of contemporary family life (Gold, 2002). As a primarily agrarian society has shifted largely to concentrated urban areas and suburban offshoots, the needs of families has changed over the years. Rising expectations for student achievement and the desire to remain globally competitive has led school reformers to suggest two possible alternatives to the allocation of time in US public schools: extend the amount of time spend in school or redistribute schooling days throughout the calendar year (Ballinger, Kirschenbaum, & Poimbeauf, 1987; Elam, Rose, & Gallop, 1996).

Response from academics and the public scrutinized the effect of calendar change on school and family life. Those in favor of the traditional calendar argue that potential benefits resulting from changes in the school calendar would not outweigh the disruption of family time (Sardo-Brown & Rooney, 1992). Private sector industries such as the travel industry worry about the economic impact of such a change (Waddell, 1997). For schools, the challenge of scheduling of sporting events and for parents, the coordinating activities when children in the same family are operating on different schooling schedules, is difficult (Orellana & Thorne, 1998).

Proponents of calendar change have argued that increases in schooling time leads to increases in student learning. Others believe rearranging schooling time is a better fit for families (Gandara & Fish, 1994). The quality of professional life for school personnel is also an important factor. Studies of teacher absenteeism and teacher attitudes on a year-round calendar versus a traditional calendar show lower rates of absenteeism and improved morale, although some were descriptive rather than analytical in nature (Gandara & Fish, 1994; Pelavin, 1979). Concerns about rising costs have led to

further examination of the two alternatives. Extending time, whether it be adding schooling days or extending the number of schooling hours per day, would require longer hours for school staff and greater use of schooling resources (i.e. heating or cooling the building; serving additional meals). For this reason, extended learning time has been largely ignored by policy makers as a potential schooling reform. In contrast, redistributing schooling days without increasing the total schooling time might be a better option financially. Redistributing schooling days in such a fashion where students are rotated in and out of schooling times provides an opportunity to serve more children without spending additional monies on building schools (Daneshvary & Clauretie, 2001). Although schools would incur the additional financial responsibility of keeping the school building operational throughout the entire calendar year, the number of students served could be maximized. Referred to as multi-tracking, four classes of students could be served in three classrooms, as long as one class of students was on a break. Rotating the classes by giving each class several breaks throughout the school year could satisfy the requirement of educating students for 180 days, but maximizing the resources in an existing school building.

The redistribution of schooling days is often referred to as the modified year-round calendar. The academic benefits of a modified year-round calendar may be rooted in two theories of learning: the spacing effect and time-on-task. The spacing effect focuses on patterns of learning sessions where researchers are finding that learning trials spaced over time leads to greater long-term retention (see Dempster, 1989 for a review). The time-on-task literature connects student engagement to student achievement (Carroll, 1989). Forces that influence student engagement may be directly impacted by the schooling calendar (Karweit, 1984). Frequent breaks could help students re-engage with material, and absorb learning concepts more easily. Additionally, frequent breaks could influence teacher energy, translating into more promising teaching methods leading to higher student engagement again and ultimately greater student achievement.

Besides the support of the educational literature in explaining the possible benefits of a modified year-round calendar, academics concerned with equity in schooling point to evidence that

the traditional calendar, defined by a long summer break, disadvantages vulnerable populations (Heyns, 1978). Examined most thoroughly with students of low socio-economic status, a series of work has concentrated on the gap in learning that occurs between the end of one school year and the beginning of the next (Alexander et al., 2001, Entwisle et al., 1997). During the summer, researchers find that poorer students experience greater learning loss than wealthier students. Due to a variety of proposed explanations including differential exposure to learning activities such as visiting a museum or the library, students of low socio-economic status enter the new school year with a distinct and measureable disadvantage than their peers, compounded every summer, leading to a sizeable gap over the time (Borman et al., 2005). Eliminating the long summer break may alter the cycle of disadvantage and lead to greater equity in achievement among students of varying socio-economic status.

The current study addressed many of the potential advantages for the modified year-round calendar including the relationship between: calendar type and student achievement, student achievement in specific student subgroups, student absenteeism (a proxy for student engagement), teacher retention, and school spending. This chapter will review the dissertation findings, interpret the implications of the findings, discuss the study limitations, and provide suggestions for future research.

## Summary of Findings

There were five major findings in this dissertation. First, the modified year-round calendar leads to improved student achievement for students of low socio-economic status. Second, the modified year-round calendar also appears advantageous for students with special needs. Additionally, the modified year-round calendar is detrimental to student performance for students who are English language learners. Fourth, the link between the modified year-round calendar and lower rates of student absenteeism is supported in the student fixed effects methodology. Fifth, higher teacher retention rates are correlated with a change from a traditional calendar to a modified year-round calendar.

Some of the findings were puzzling given what we know from the literature. There was no benefit of the modified year-round calendar detected for all students' achievement. Somewhat surprisingly, the notion that multi-tracked modified year-round schools are a panacea for rapid growth and limited resources was incongruent with study results. Instead, the current research finds that schools, on average, spent at least as much money per pupil on a modified year-round calendar than on a traditional calendar and in at least one case, the modified year-round calendar lead to greater spending (student services).

## Interpretation and Implications

The modified year-round calendar does appear to be a viable option for improving student performance for at least some student subgroups. On average, students of low socio-economic status and students with special needs could specifically benefit from calendar reform. Students who are learning English would be negatively affected from a calendar change. Understanding why some subgroups seem to benefit, while students overall are not more likely to see a benefit from the modified year-round calendar requires revisiting the theoretical framework and policy implications.

First, isolating any school reform that impacts student achievement is difficult. Since Coleman's landmark study of school resource inequality researchers have been investigating schooling inputs and debating whether or not schools matter (Coleman, 1966; Greenwald et al., 1996; Hanushek, 1997). Heyns's work on comparing school-time learning rates with out-of-school learning rates to isolate the effects of schooling reinforced the notion that schools do impact student learning (Heyns, 1978, 1987). Since this time, researchers have focused on identifying policy actions that could lead to improved student performance (Nye, Hedges, & Konstantopoulos, 2002). The empirical results from the current study support the notion that school reforms can be directly linked to student achievement. Understanding why the reform is successful at improving student learning for some subgroups, but not others requires further thought.

The theoretical foundations of the spacing effect and time-on-task literatures may shed some information as to the reasons why the modified year-round calendar increases student performance for

a subset of students. First, the main principal behind the spacing effect says that when time between learning sessions is lengthened, information is better remembered (Dempster, 1989). Frequent testing of information also appears to increase retention (Cepeda et al., 2009). As explained in the theoretical literature review, it is the lag time plus repeated testing that requires more sophisticated cognitive strategies to retain information and opportunities to test the robustness of those strategies. Modified year-round calendar schools were thought to allow more opportunities to put these cognitive strategies into practice by requiring more frequent review sessions upon each return from a break. Initially, the literature suggests that all students would benefit for the modified year-round calendar reform. However, students of social and economic advantage may not need the extra rehearsal and retention tests provided by modified year-round teachers to improve their learning. Instead, parents may be providing those opportunities during out-of-school time, regardless of whether that time occurs over short breaks or during a long summer recess. In contrast, students of low socio-economic status may not receive the additional opportunities to review information and the extra retention tests present in modified year-round schools fill the gap. Additionally, students with special needs whom often need additional help in the learning process may be helped by receiving the extra opportunities to recall information given in modified year-round calendar schools.

Yet, it remains puzzling that all students did not reap the benefits of the modified year-round calendar reform which warrants a revisiting of the spacing effect and time-on-task literatures. Perhaps the reform effort was not as closely aligned with the theoretical literatures cited, as originally thought. For example, much attention was paid to the distinction between the arrangements of in-school and out-of-school time in traditional versus modified-year round calendar schedules. The modified year-round calendar was viewed as an example of spaced learning trials, with a rotation of nine weeks of learning followed by a three week break. Subsequently, the robust spacing effect was expected to result in pronounced differences in achievement outcomes compared to the traditional calendar "massed" learning approach. However, I neglected to attend to differences in the length of the learning trials referred to in the literature compared to the length of learning trials that occur

across the school year. Most of the research on the spacing effect contained learning trials that lasted less than a single day, oftentimes less than a hour (Bahrick, 1979). Generalizing from the relatively short learning episodes reported in the literature to a nine week learning episode may be an over-application of the spacing effect. It may be inappropriate to apply the spacing effect to calendar reform because the underlying strategies on which the spacing effect is explained in the literature such as encoding variability theory or the attention hypothesis may be inapplicable when learning sessions last over a longer period of time (Dempster, 1989, Hintzman, 1974).

Another reason the spacing effect was not present in the current study on calendar reform may be due to environmental factors. Most of the research on the spacing effect was conducted within a laboratory session. As Bahrick (1979) points out, most of the learning material used in memory studies is restricted to relatively straightforward information and devoid of the complex cognitive learning commonplace in the classroom. While robust across subject area (Dempster, 1989), the spacing effect may be less consistent in dealing with the acquisition and maintenance of higher order skills. Using the spacing effect to explain differences in outcomes across calendar arrangements would be more appropriate if traditional memory research trials were conducted using more complex learning tasks and the spacing effect remained.

Remarkably, students who speak English as a second language are disadvantaged by a modified year-round calendar. Perhaps the breaks for these students are too disruptive in the learning acquisition process. Or perhaps students take extended vacations to their homeland that result in higher rates of absenteeism. Much research finds that language acquisition is best acquired through full immersion (Genesee, 1985). Perhaps the frequent breaks detract from the immersion process and result in lower rates of learning.

The empirical results provide some evidence that students are less frequently absent under a year-round calendar. Student absenteeism is commonly used as a proxy for student engagement. The relationship between student engagement and student achievement is fleshed out in the time-on-task literature. In sum, the time-on-task theory emphasizes the quality of schooling time over the quantity

of time. It is the quality of the time that influences student engagement and ultimately student achievement. The quality of the time may be affected by a number of factors, in this case, the school calendar. The modified year-round calendar manipulates time in ways that may contribute to higher quality in-school sessions. If the quality of schooling time is enhanced, then presumably students in a modified calendar would exhibit higher student engagement measured by fewer days absent. Students in Wake County experienced both types of calendar arrangements and were absent less (more than a day on average) on a modified year-round calendar than a traditional calendar. The current study provides preliminary evidence that features, however defined, of the learning context which accompany the modified year-round calendar contribute to higher rates of engagement.

While the study results indicate less student absenteeism under the modified year-round calendar, there is no evidence that student achievement on standardized tests is improved. While time-on-task researchers theorize that greater student engagement leads to higher student achievement (Carroll, 1963; Cobb, 1972), the presence of other mediating and moderating variables could inhibit the strength of the engagement-achievement relationship. In the case of the current study, the lack of evidence connecting student engagement to student achievement does not necessarily require a revision of the theory. In fact, the large database of information collected in the current study could be used in a structural equation modeling analysis to more directly test the time-on-task theoretical framework.

The simplest explanation for the lack of an observed relationship between student engagement and student achievement is the use of student absenteeism as a measure of engagement. While student absenteeism is the best available source for information on student engagement in the current study, and not an uncommon variable used to measure engagement, it lacks a strong direct correlation between the measurement variable and the underlying construct. Finally, it is also possible that a threshold effect may be at play. Perhaps the modified year-round calendar improves student engagement, but does not reach a threshold in which the extra engagement is translated into higher student achievement.

One feature that may be affected by the restructuring of the schooling calendar is teacher retention. Individuals leave the teaching profession for a variety of reasons (Ingersoll, 2003). However, I find an increase in teacher retention after the conversion to the modified year-round calendar. Establishing a link between a specific school reform and teacher retention is important because teacher retention is connected to a number of important issues including the teacher shortage and importance of teaching experience in improving student learning gains (Clotfelter et al., 2007; Darling-Hammond, 1994). Interestingly, teacher attitudes appear more favorable to year-round calendars over time after they had experienced the modified year-round calendar (Fardig, 1992). Directly aligned with this notion is the finding that teacher retention was significantly increased two years after the calendar conversion. As teachers adjust to the modified year-round calendar, they begin to prefer the new schedule. Similarly, parents and students also favor the modified year-round calendar once exposed (Fardig, 1992; Shields, 1996). The only exception is parents who have children in both traditional and year-round calendar schools.

Teacher support is also essential if calendar change is to become a politically viable reform. The wave of school reform called for by the current administration brings calendar reform to the forefront. Catching on in the larger policy arena, extended learning time is on the agenda for Secretary of Education Arne Duncan (Quinn & Negron, 2009). Increasing pressure to compete globally, beginning with the report A Nation at Risk (National Commission on Excellence in Education, 1983), and supported by conclusions from the National Commission on Time and Learning, the desire to increase time in-school is motivated by lagging test scores of US students when compared to other national counterparts. Wiseman and Baker (2004) report that national differences in supplemental learning opportunities during out-of-school time may account for sub-par international ratings of US students.

Examinations on the effectiveness of time in a schooling environment are split into two camps; either increasing the amount or distribution of schooling time through calendar changes or by supplementing in-school learning with programs during the summer recess. Historically, the federal

governments have avoided the more controversial calendar change, instead promoting a variety of summer programs intended to increase academic learning opportunities, which left local district leaders as the primary initiators of novel school calendars (Shepard & Reed, 1975). Yet, the study of summer learning programs has mixed results (Cooper et al., 2000). Perhaps the summer learning experiences, which are not always mandatory and regardless have lower attendance rates, are not of the same quality as the supplemental learning experiences of students in other nations. Regardless of the reason, policy-makers are shifting their attention to calendar reform as evidenced by the proliferation of year-round schools. The current study supports the notion that calendar reform may be a vehicle for accelerating academic achievement, however the results indicate that conversion to a modified year-round calendar seems more appropriately marketed as a measure to promote school equity than overall school effectiveness.

Besides raising student achievement, the rationale behind the movement to a modified year-round calendar is also motivated by fiscal concerns. In a national review, Nygaard (1974) includes utilizing facilities and resources more effectively and alleviating overcrowding as examples of district rational for instituting calendar change. In my study, the conversion to a modified year-round calendar did not reap financial benefits. Modified year-round schools did not have lower per pupil expenses. A number of reasons might explain the results. First, the largest area for cost-saving in switching to a multi-track modified year-round calendar is in real estate savings. Capital outlay expenses are often allocated at the district level. In the dataset, capital outlay expenses at the district level are reallocated to schools based on a per-pupil basis. This means that schools in the treatment and comparison groups, all from Wake County, received capital outlay expenses proportionate to their student population. A stronger analysis of changes in capital outlay expenses would be an inter-district rather than intra-district analysis. Second, other research suggests that cost-savings for modified year-round schools are not evident in the first few years of implementation because of specific start-up costs (Fardig, 1992). For example, the need to rotate classrooms results in the desire to purchase special rolling cabinets for teachers and rolling cubbies for students to use to stow desk

materials and roll into storage during track-out periods. Additionally, Fardig (1992) noted that modified year-round calendar schools purchased fans and additional refrigerators for use during the hot summer months. Third, schools just beginning a year-round calendar often do not operate at full capacity during the first few years (Fardig, 1992). When schools do not operate at full capacity, per-pupil expenditures rise. In order to garner more conclusive evidence on the potential cost-savings of the modified year-round calendar, longer longitudinal designs should be employed.

## Limitations

The current study has several limitations that may affect the internal validity, external validity, or both. While every attempt was made to select the strongest research design to address the research questions, some additional issues remain. In addition, my design utilized several analysis techniques, each with their own limitations. Next is a discussion of the study limitations.

The research design uses two techniques (growth curve analysis and student fixed effects) for addressing whether a link exists between calendar type and student achievement or days absent. Although each technique has its pros and cons, the design centered on isolating agreement between analyses to further the line of inquiry on the effect of a modified year-round calendar.

The growth curve models (as well as the fixed effects models) used standardized student test scores as the dependent variable. The conversion to standardized test scores creates independent variable estimates that are based on movement within the student distribution over time. Changes within the distribution, as opposed to the original scale score metric, masks absolute growth over time. The decision to use standardized scores was based primarily on the difficulty of using the scale score to create a consistent metric across grade levels and across time. If a reliable metric can capture scale score on End of Grade tests in reading and math for grades 3-8, it may warrant a reanalysis of the growth curve procedure.

Another drawback to the growth curve models is the paucity of information about why specific schools in North Carolina adopted a modified year-round calendar. As prior research suggests, some school districts are motivated by cost savings since there is evidence that multi-track

modified calendar year-round schools is a financial conservative measure when faced with rapid student population growth (Daneshvary & Clauretie, 2001). However, it is equally possible that struggling schools turned to the modified year-round calendar as part of a school improvement plan. If schools select into the modified year-round calendar condition for reasons other than the observed characteristics used in the matching procedure, then the treatment and comparison schools cannot be considered equivalent resulting in erroneous effect estimates. A thorough understanding of the motivations behind a school's assignment to a traditional or modified year-round calendar would be a good step in increasing confidence in the matched sample. Along the same lines, individual families may intentionally relocate or petition schools so that their children will attend a modified year-round school. If the families selecting into year-round schools are qualitatively different from families selecting into traditional calendar schools, then the current study is vulnerable to selection bias.

The difference-in-differences estimation of teacher retention and school expenditures was based on propensity score matching within Wake County. Choosing schools within Wake County helped to eliminate sources of variation at the district level between treatment and matched schools. However, the matching procedure yielded less than ideal matching on observed characteristics. Opening up the matching sample to schools throughout the state would likely have resulted in a comparison group that more closely approximated the treatment sample on observed characteristics. However, it could have introduced heterogeneity between the treatment and comparison groups with respect to district level factors. Additional analyses using an alternate matched sample may prove fruitful in providing further support for the study conclusions.

The study used grades 3-8 with the intention that results would be generalized to elementary and middle schools. However, grades K-2 are not currently subjected to mandatory state-wide testing in reading and mathematics. The assumption that elementary grade students in grade K-2 would perform similarly on traditional or a modified year-round calendar to students in grades 3-5 is somewhat tenuous and not without question.

Additionally, I chose the metric of free or reduced price lunch eligibility as representative of low socio-economic status. While there is a long line of research supporting a link between student achievement and socio-economic status, using free or reduced price lunch only incorporates parental income and ignores parental education, occupational prestige, and access to resources at home, which all play a part in the conceptual understanding of socioeconomic status (White, 1982). It is also true that using a dichotomized indicator for the otherwise continuous variable of socio-economic status may downwardly bias the effect of socioeconomic status on achievement (Sirin, 2005).

Finally, the current study largely ignored other non-mandatory school-sponsored learning opportunities such as summer school and intersession study. Lack of available data on non-mandatory learning programs was the primary reasons summer school and intersession study was not addressed. Presumably some student in the sample experienced either summer school or intersession remediation. This would be especially problematic if there is a qualitative difference between summer school and intersession learning conditions. For purposes of this study, all sources of learning outside of mandatory in-school time were considered randomly distributed across calendar conditions.

## Suggestions for Future Research

As a research community, we have only scratched the surface of understanding calendar change and its impact on teachers, students, families, and communities. While many schooling outcomes of interest exist, understanding the influence of school calendar reform on student achievement may be the most consequential. Yet, direct estimation of students performance under a modified year-round calendar compared to a traditional calendars face remains difficult because of the non-random sorting of students into schools of varying calendar arrangements. Whether schools are electing to adopt a modified year-round calendar or parents are choosing to place their child in a modified year-round calendar, those choices are related to factors that are not always observed and therefore often unaccounted for in statistical models. Understanding whether students who purposefully select into modified year-round schools is another powerful tool for developing

knowledge on the pathways in which the school calendar influences student achievement. In an alternate specification of the Wake County student fixed effects model, students who attended schools other than the converted schools were retained in the analysis. Using this sample, the coefficient of the modified year-round calendar was positive and statistically significant, suggesting that students and families who intentionally move from a school under a traditional calendar to a school under a modified year-round calendar are seeing an academic benefit to the switch. Further modeling is necessary in order to fully investigate whether student, rather than school, switchers experience greater academic success in a modified year-round calendar.

Along similar lines, unpacking the relationship between calendar arrangement and student achievement is vital for understand the sequence of events or conditions that promote greater student learning. Qualitative investigations are often the best way to explore the process of reform. Understanding how the modified year round calendar impacts the quality of schooling time will be beneficial in generating knowledge about the learning process through the identification of mediating factors such as teacher attitudes, student engagement, and community support. Examining teacher absenteeism and student retention rates are other possible factors for investigation. Besides enhancing our conceptual knowledge of the learning process, researchers must also explore ways in which the community context responds to calendar reform. Policy makers must consider financial constraints and local opinion along with student achievement in order to execute sweeping changes in calendar arrangement.

Lastly, the modified year-round calendar is not without its faults. Administrative burnout may be a major factor in the success of calendar change. In the move from a ten-month academic year to a twelve-month calendar year, the amount of time required by administrative personnel increases. Additionally, with multi-tracking the size of the school increases, creating a heavier load for school administrators. In one study, researchers examined administrator stress (French, 1992). Believing that the prolonged startup of the school year created increased stress for school principals and assistant principals, French (1992) surveyed administrators in year-round and traditional calendar

schools and compared their responses. French found that there were no significant differences in stress, measured as exhaustion or accomplishments. Still, more work should be done to investigate the way calendar change affects school administrators and other support staff (including student services professionals) whose workload increases when transitioned to a year-round schedule.

## Conclusion

The current study supports the use of a modified year-round calendar to raise achievement for students of low socio-economic status and students with special needs. Students attend more days of schooling and teachers are retained at a higher level under the modified year-round calendar condition. More work needs to be done to see whether the multi-tracking modified year-round calendar is a viable cost-savings measure.

## Appendix A

### Background of Wake County Public Schools Year-Round Schooling Plan

Wake County Public Schools are governed by a nine-member board of education whose members are elected for four year terms from nine separate districts in the county. While the school board determines policy, they are not responsible for the system's day-to-day operation and are not able to raise taxes, but instead submit annual budget requests to the Wake County Board of Commissioners which have power over the allocation of funds to schools. During the 2006-2007 school year the Board of Education and County Commissioners, worried about student enrollment increases and the long school construction process, developed a plan and voted to transition 19 elementary schools and 3 middle schools to multi-track year-round calendar for the 2007-08 academic year (Wake County Public School System, 2008). This policy change came from the projected influx of students to the Wake county area and the lack of time (and to some extent money) to construct additional buildings. As discussed previously, year-round schools are thought of as a fiscally sound alternative to dealing with growing school-age populations and limited resources.

Since 1990 Wake County in North Carolina has been gradually increasing the number of schools operating under a year-round calendar. In 2006 there were 20 existing year-round elementary schools in Wake County. Unlike most other year-round schools which opened up from the outset as a year-round school, this plan would convert schools that were on a traditional calendar the year before to a year-round calendar the next year. No specific provisions were put in place for parents to transfer their children to other traditional calendar schools, except for the existing transfer request process.

On February 13, 2007 a group called WakeCARES presented a letter from Robert Hunter Jr., the attorney for Wake CARES, to the Wake County Public School System. The letter requested that Wake County strip the year-round conversions from the Board of Education's plan. Noncompliance resulted in a lawsuit filed by WakeCARES against the Wake School Board in 2007. The court case began in April and on May 3, 2007 Judge Howard Manning, Jr. ruled in favor of WakeCARES and

decided that the Wake County School System may not assign pupils to a calendar other than traditional (WakeCARES, 2008). The parents/guardians of pupils attending year-round or modified calendar schools were required to sign a consent form agreeing to participate in year-round schooling, or be assigned to a traditional-calendar school. The next day the Wake County Board of Education decided to continue with the conversion and filed an expedited appeal of Judge Manning's ruling. The Wake County school board asked for a stay of Judge Manning's decisions mandating parental consent for year-round assignment, but later in June the request was denied. Wake County complied with the ruling for the 2007-2008 school year, but pursued the appeal. In January 2008 the appeal hearing began. On May 1, 2009 the North Carolina Supreme Court upheld the earlier court's decision which states that the NC State Board of Education has the authority to assign students to year-round calendar schools without parental consent.

## Appendix B

## Data Tables

**Table B.1. Reliability of End of Grade Math Tests (2006): Gender and Ethnicity**

| | Gender | | Ethnicity | | | | | |
|---|---|---|---|---|---|---|---|---|
| Grade | Male | Female | Asian | Black | Hispanic | Native American | Multi-Racial | White |
| 3 | .96 | .96 | .97 | .95 | .97 | .95 | .95 | .95 |
| 4 | .95 | .95 | .97 | .94 | .97 | .95 | .94 | .95 |
| 5 | .96 | .95 | .97 | .94 | .97 | .94 | .95 | .95 |

**Table B.2. Reliability of End of Grade Math Tests (2006): Disability Status, Limited English Proficiency**

| | Disability Status | | English Proficiency Status | |
|---|---|---|---|---|
| Grade | No Disability | Disability | Limited English | Not Limited English |
| 3 | .94 | .97 | .98 | .96 |
| 4 | .94 | .97 | .98 | .96 |
| 5 | .94 | .97 | .97 | .95 |

**Table B.3. Reliability of End of Grade Reading Tests (2004)**

| Grade | Reliability |
|---|---|
| 3 | .92 |
| 4 | .92 |
| 5 | .91 |

**Table B.4. Linkages between Research Questions and Specific Analyses**

| Research Question | Analysis Used to Probe Question |
|---|---|
| (1) Do students perform better on achievement tests under a year-round modified calendar compared to a traditional calendar | (a) Growth Curve Analysis<br>(b) Student Fixed Effects |
| (2) Do students exhibit lower rates of absenteeism under a modified year-round calendar compared to students under a traditional calendar? | (b) Student Fixed Effects |
| (3) Do specific student subgroups (low SES, English Language Learners, students in special education) perform better on student achievement tests under a year-round modified calendar compared to a traditional calendar? | (a) Growth Curve Analysis |
| (4) Is there less teacher turnover in year-round modified calendar schools compared to traditional calendar schools? | (c) Difference-in-differences |
| (5) Do year-round modified calendar schools spend less money per student than schools operating under a traditional calendar? And if so, what expenditure categories differ? | (c) Difference-in-differences |

**Table B.5. Year-Round schools in North Carolina**

| Year | Number of Year-Round Schools |
|---|---|
| 2005-2006 | 64 |
| 2006-2007 | 62 |
| 2007-2008 | 86 |
| 2008-2009 | 91 |

**Table B.6. Description of Growth Curve Sample for Math Achievement Outcome (*n* = 180, 196)**

| Variable | *M* | *SD* |
|---|---|---|
| Math Achievement Score (standardized) | 0.092 | 1.018 |
| Wave of time | 2.491 | 1.121 |
| Modified Year-Round Calendar | 0.486 | 0.500 |
| Prior Achievement Score (standardized) | 0.081 | 0.950 |
| Asian | 0.026 | 0.159 |

| Variable | *M* | *SD* |
|---|---|---|
| Black | 0.318 | 0.466 |
| Hispanic | 0.111 | 0.315 |
| Multiracial | 0.038 | 0.191 |
| American Indian | 0.025 | 0.156 |
| Grade | 5.127 | 1.670 |
| Male | 0.506 | 0.500 |
| Is LEP | 0.063 | 0.243 |
| Was LEP | 0.018 | 0.132 |
| Gifted | 0.167 | 0.373 |
| Special Needs | 0.114 | 0.318 |
| Free or Reduced Price Lunch | 0.378 | 0.485 |
| Free or Reduced Price Lunch Missing | 0.067 | 0.250 |
| Underage | 0.012 | 0.110 |
| Overage | 0.211 | 0.408 |
| Days absent | 6.394 | 6.244 |
| Forced Move | 0.139 | 0.346 |
| Move in Year | 0.050 | 0.217 |
| Per pupil spending (per $100) | 80.979 | 13.204 |
| Percent of teachers with National Board Certification | 0.103 | 0.084 |
| Percent of teachers with an Advanced Degree | 0.300 | 0.094 |
| Percent of teachers with 3 or less years of Experience | 0.248 | 0.104 |
| Asian Mean | 2.660 | 2.647 |
| Black Mean | 31.855 | 21.340 |
| Hispanic Mean | 12.184 | 9.944 |
| Multiracial Mean | 3.891 | 1.973 |
| American Indian Mean | 2.450 | 10.666 |
| Percent Free or Reduced Price Lunch | 46.041 | 25.461 |
| School Size (per 100 students) | 8.017 | 3.041 |
| School Size Squared | 73.524 | 53.886 |

| Variable | *M* | *SD* |
|---|---|---|
| Suspension Rate (per 100) | 15.264 | 20.289 |
| Violent Acts Rate (per 1000) | 4.216 | 5.135 |
| Teacher Supplements | 40.891 | 15.788 |
| Propensity Score | 0.081 | 0.052 |

**Table B.7. Description of Growth Curve Sample for Reading Achievement Outcome (*n* = 180, 445)**

| Variable | *M* | *SD* |
|---|---|---|
| Math Achievement Score (standardized) | 0.068 | 1.006 |
| Wave of time | 2.490 | 1.121 |
| Modified Year-Round Calendar | 0.486 | 0.500 |
| Prior Achievement Score (standardized) | 0.084 | 0.949 |
| Asian | 0.026 | 0.160 |
| Black | 0.317 | 0.465 |
| Hispanic | 0.111 | 0.314 |
| Multiracial | 0.038 | 0.191 |
| American Indian | 0.025 | 0.156 |
| Grade | 5.129 | 1.670 |
| Male | 0.505 | 0.500 |
| Is LEP | 0.063 | 0.242 |
| Was LEP | 0.018 | 0.132 |
| Gifted | 0.167 | 0.373 |
| Special Needs | 0.112 | 0.315 |
| Free or Reduced Price Lunch | 0.377 | 0.485 |
| Free or Reduced Price Lunch Missing | 0.067 | 0.250 |
| Underage | 0.012 | 0.110 |
| Overage | 0.209 | 0.407 |
| Days absent | 6.392 | 6.249 |

| Variable | *M* | *SD* |
|---|---|---|
| Forced Move | 0.139 | 0.346 |
| Move in Year | 0.049 | 0.217 |
| Per pupil spending (per $100) | 80.956 | 13.191 |
| Percent of teachers with National Board Certification | 0.103 | 0.084 |
| Percent of teachers with an Advanced Degree | 0.300 | 0.094 |
| Percent of teachers with 3 or less years of Experience | 0.248 | 0.104 |
| Asian Mean | 2.661 | 2.648 |
| Black Mean | 31.844 | 21.328 |
| Hispanic Mean | 12.173 | 9.928 |
| Multiracial Mean | 3.891 | 1.972 |
| American Indian Mean | 2.447 | 10.652 |
| Percent Free or Reduced Price Lunch | 46.013 | 25.450 |
| School Size (per 100 students) | 8.021 | 3.040 |
| School Size Squared | 73.574 | 53.880 |
| Suspension Rate (per 100) | 15.280 | 20.307 |
| Violent Acts Rate (per 1000) | 4.221 | 5.138 |
| Teacher Supplements | 40.898 | 15.787 |
| Propensity Score | 0.081 | 0.052 |

**Table B.8. Weighted Description of Student Fixed Effects Sample for Math Achievement Outcome (*N* = 15, 917)**

| Variable | *M* | *SD* |
|---|---|---|
| Math Achievement Score (standardized) | 0.380 | 0.971 |
| Modified Year-round calendar | 0.428 | 0.495 |
| Asian Mean | 5.116 | 4.914 |
| Black Mean | 20.934 | 15.459 |
| Hispanic Mean | 11.143 | 7.657 |
| Multiracial Mean | 4.299 | 1.151 |

| Variable | *M* | *SD* |
|---|---|---|
| American Indian Mean | 0.308 | 0.279 |
| Percent Free or Reduced Price Lunch | 28.891 | 19.546 |
| School Size (per 100 students) | 8.391 | 1.447 |
| School Size Squared | 72.501 | 23.821 |
| Suspension Rate (per 100) | 8.449 | 15.338 |
| Violent Acts Rate (per 1000) | 4.379 | 8.785 |
| Teacher Supplements | 58.129 | 3.391 |
| Number of Students in Class | 23.335 | 3.764 |
| Average Standard Deviation of Peers Prior Achievement Score | 0.841 | 0.172 |
| Prior Achievement score (standardized) | 0.355 | 0.922 |
| Average Peer Prior Achievement Score (standardized) | 0.354 | 0.489 |
| Asian | 0.048 | 0.214 |
| Black | 0.204 | 0.403 |
| Hispanic | 0.104 | 0.306 |
| Multiracial | 0.041 | 0.199 |
| American Indian | 0.003 | 0.053 |
| Male | 0.514 | 0.500 |
| Days absent | 6.364 | 5.687 |
| Move in Year | 0.044 | 0.205 |
| Underage | 0.015 | 0.123 |
| Overage | 0.212 | 0.409 |
| Gifted | 0.186 | 0.389 |
| Special Needs | 0.128 | 0.334 |
| Is LEP | 0.056 | 0.229 |
| Was LEP | 0.036 | 0.187 |
| Free Lunch | 0.186 | 0.389 |
| Reduced Price Lunch | 0.051 | 0.219 |
| Free or Reduced Price Lunch Missing | 0.126 | 0.332 |

**Table B.9. Weighted Description of Student Fixed Effects Sample for Reading Achievement Outcome (*N* = 12,626)**

| Variable | *M* | *SD* |
|---|---|---|
| Reading Achievement Score (standardized) | 0.348 | 0.944 |
| Modified Year-round calendar | 0.440 | 0.496 |
| Asian Mean | 5.394 | 5.019 |
| Black Mean | 18.647 | 13.685 |
| Hispanic Mean | 10.798 | 7.656 |
| Multiracial Mean | 4.371 | 1.236 |
| American Indian Mean | 0.280 | 0.284 |
| Percent Free or Reduced Price Lunch | 26.776 | 17.781 |
| School Size (per 100 students) | 8.109 | 1.392 |
| School Size Squared | 67.696 | 22.272 |
| Suspension Rate (per 100) | 3.173 | 4.307 |
| Violent Acts Rate (per 1000) | 1.350 | 1.833 |
| Teacher Supplements | 58.197 | 3.378 |
| Number of Students in Class | 23.009 | 2.952 |
| Average Standard Deviation of Peers Prior Achievement Score | 0.886 | 0.147 |
| Prior Achievement score (standardized) | 0.388 | 0.911 |
| Average Peer Prior Achievement Score (standardized) | 0.386 | 0.392 |
| Asian | 0.051 | 0.221 |
| Black | 0.176 | 0.381 |
| Hispanic | 0.102 | 0.303 |
| Multiracial | 0.043 | 0.202 |
| American Indian | 0.003 | 0.051 |
| Male | 0.511 | 0.500 |
| Days absent | 6.015 | 5.107 |
| Move in Year | 0.039 | 0.193 |
| Underage | 0.015 | 0.121 |

| Variable | *M* | *SD* |
|---|---|---|
| Overage | 0.195 | 0.396 |
| Gifted | 0.179 | 0.384 |
| Special Needs | 0.122 | 0.327 |
| Is LEP | 0.055 | 0.227 |
| Was LEP | 0.040 | 0.196 |
| Free Lunch | 0.163 | 0.369 |
| Reduced Price Lunch | 0.045 | 0.208 |
| Free or Reduced Price Lunch Missing | 0.154 | 0.361 |

**Table B.10. Weighted Description of Student Fixed Effects Sample for Days Absent Outcome (*N* = 15,971)**

| Variable | *M* | *SD* |
|---|---|---|
| Days absent | 0.348 | 0.944 |
| Modified Year-round calendar | 0.440 | 0.496 |
| Asian Mean | 5.394 | 5.019 |
| Black Mean | 18.647 | 13.685 |
| Hispanic Mean | 10.798 | 7.656 |
| Multiracial Mean | 4.371 | 1.236 |
| American Indian Mean | 0.280 | 0.284 |
| Percent Free or Reduced Price Lunch | 26.776 | 17.781 |
| School Size (per 100 students) | 8.109 | 1.392 |
| School Size Squared | 67.696 | 22.272 |
| Suspension Rate (per 100) | 3.173 | 4.307 |
| Violent Acts Rate (per 1000) | 1.350 | 1.833 |
| Teacher Supplements | 58.197 | 3.378 |
| Number of Students in Class | 23.009 | 2.952 |
| Average Standard Deviation of Peers Prior Achievement Score | 0.886 | 0.147 |
| Prior Achievement score (standardized) | 0.388 | 0.911 |

| Variable | *M* | *SD* |
|---|---|---|
| Average Peer Prior Achievement Score (standardized) | 0.386 | 0.392 |
| Asian | 0.051 | 0.221 |
| Black | 0.176 | 0.381 |
| Hispanic | 0.102 | 0.303 |
| Multiracial | 0.043 | 0.202 |
| American Indian | 0.003 | 0.051 |
| Male | 0.511 | 0.500 |
| Move in Year | 6.015 | 5.107 |
| Underage | 0.039 | 0.193 |
| Overage | 0.015 | 0.121 |
| Gifted | 0.195 | 0.396 |
| Special Needs | 0.179 | 0.384 |
| Is LEP | 0.122 | 0.327 |
| Was LEP | 0.055 | 0.227 |
| Free Lunch | 0.040 | 0.196 |
| Reduced Price Lunch | 0.163 | 0.369 |
| Free or Reduced Price Lunch Missing | 0.045 | 0.208 |

**Table B.11. Description of Difference-in-differences Sample**

| | 2006-07 | | 2007-08 | | 2008-09 | |
|---|---|---|---|---|---|---|
| **Variables** | ***M*** | ***SD*** | ***M*** | ***SD*** | ***M*** | ***SD*** |
| Asian Mean | 4.00 | 3.85 | 4.31 | 4.35 | 4.62 | 4.53 |
| Black Mean | 25.36 | 15.12 | 25.22 | 14.98 | 23.47 | 13.71 |
| Hispanic Mean | 12.62 | 7.54 | 12.70 | 7.22 | 14.62 | 8.95 |
| Multiracial Mean | 4.52 | 1.35 | 4.78 | 1.12 | 5.44 | 1.24 |
| American Indian Mean | 0.26 | 0.27 | 0.28 | 0.26 | 0.25 | 0.20 |
| Percent Free or Reduced Price Lunch | 36.50 | 20.00 | 35.28 | 20.10 | 35.76 | 18.69 |

| Variables | 2006-07 | | 2007-08 | | 2008-09 | |
|---|---|---|---|---|---|---|
| | ***M*** | ***SD*** | ***M*** | ***SD*** | ***M*** | ***SD*** |
| Suspension Rate (per 100) | 7.94 | 11.04 | 7.69 | 9.77 | 7.23 | 8.21 |
| Violent Acts Rate (per 1000) | 3.79 | 6.52 | 4.23 | 6.17 | 4.85 | 6.68 |
| Per Pupil Spending (in $100) | 76.21 | 6.81 | 78.36 | 6.84 | 81.86 | 6.63 |
| Per Pupil Spending in Regular Instruction | 3822.75 | 331.72 | 4152.75 | 377.91 | 4269.22 | 406.08 |
| Per Pupil Spending in Special Instruction | 748.53 | 172.60 | 1012.26 | 222.14 | 1071.83 | 206.25 |
| Per Pupil Spending in Student Services | 397.48 | 62.39 | 350.45 | 82.03 | 346.06 | 68.60 |
| Per Pupil Spending in Instructional Support | 212.48 | 38.51 | 421.55 | 52.75 | 419.78 | 52.39 |
| Per Pupil Spending in Transportation | 218.97 | 3.79 | 233.87 | 3.28 | 230.07 | 5.04 |
| Per Pupil Spending in School Maintenance and Utilities | 556.25 | 77.83 | 543.66 | 82.83 | 538.50 | 83.86 |
| Per Pupil Spending in School Leadership | 512.88 | 99.98 | 498.11 | 79.70 | 513.19 | 86.12 |
| Per Pupil Spending in Capital Outlay | 19.77 | 78.73 | 16.61 | 57.25 | 13.78 | 50.47 |
| School Size (per 100 students) | 7.72 | 2.06 | 7.96 | 2.01 | 8.00 | 2.10 |
| Percent of teachers with National Board Certification | 0.09 | 0.06 | 0.10 | 0.06 | 0.11 | 0.06 |
| Percent of Teachers with an Advanced Degree | 0.29 | 0.09 | 0.30 | 0.08 | 0.31 | 0.09 |
| Percent of Teachers with 3 or less years of Experience | 0.26 | 0.11 | 0.26 | 0.09 | 0.24 | 0.08 |
| Percent Passed End of Grade Examinations | 0.75 | 0.09 | 0.61 | 0.14 | 0.70 | 0.12 |

| Variables | 2006-07 | | 2007-08 | | 2008-09 | |
|---|---|---|---|---|---|---|
| | ***M*** | ***SD*** | ***M*** | ***SD*** | ***M*** | ***SD*** |
| Percent of Teachers Retained Between Years | 0.89 | 0.15 | 0.87 | 0.08 | 0.90 | 0.06 |
| Percent of Teachers Retained Within a Year | 0.97 | 0.03 | 0.97 | 0.03 | 0.98 | 0.02 |

## REFERENCES


Aldridge, B. G. (1983). A mathematical model for mastery learning. *Journal of Research in Science Teaching, 20,* 1-17.

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. (2007). Lasting consequences of the summer learning gap. *American Sociological Review, 72*(2), 167-180.

Anderson, L. W. (1981). Instruction and time-on-task: A review. *Curriculum Studies, 13*(4), 289-303.

Anderson, L. W., & Scot, C. C. (1978). The relationship among teaching methods, student characteristics, and student involvement in learning. *Journal of Teacher Education, 29*(3), 52-56.

Ascher, C. (1988). Summer school, extended school year, and year-round schooling for disadvantaged students. *ERIC Clearinghouse on Urban Education Digest, 42,* 1-2.

Austin, S. D. M. (1921). A study in logical memory. *American Journal of Psychology, 32,* 370-403.

Bahrick, H. P. (1979). Maintenance of knowledge: Questions about memory we forgot to ask. *Journal of Experimental Psychology, 108*(3), 296-308.

Ballinger, C. (1988). Rethinking the school calendar. *Educational Leadership, 45*(5), 57-61.

Ballinger, C. E., Kirschenbaum, N., & Poimbeauf, R. P. (1987). *The year-round education: Where learning never stops.* Bloomington, IN: Phi Delta Kappa.

Beaton, A. E., Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1996). *Mathematics achievement in the middle school years: IEA's Third International Mathematics and Science Study (TIMSS)*. Chestnut Hill, MA: Center for the Study of Testing, Evaluation, and Educational Policy, Boston College.

Becker, G. S. (1965). A theory of the allocation of time. *Economic Journal, 75,* 493-517.

Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General, 127,* 55-68.

Berliner, D. (1990). What's all the fuss about instructional time? In M. Ben Peretz & R. Bromme (Ed.), *The Nature of Time in Schools: Theoretical concepts, practitioners perceptions* (pp. 3-35). New York: Teacher College Press.

Bloom, B. S. (1968). Learning for mastery. *Evaluation Comment, 1*(2), 1-10.

Bloom, B. S. (1984). The new direction in educational research: Alterable variables. *Asia Pacific Journal of Education, 6* (1), 1-6.

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

Borman, G. D., & Dowling, N. M. (2006). Longitudinal achievement effects of multiyear summer school: Evidence from the Teach Baltimore randomized field trial. *Educational Evaluation and Policy Analysis, 28*(1), 25-48.

Bowen, N. K., & Bowen, G. L. (1999). Effects of crime and violence in neighborhoods and schools on the school behavior and performance of adolescents. *Journal of Adolescent Research, 14,* 319-342.

Boyd D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2006.) How changes in entry requirements alter the teacher workforce and affect student achievement. *Education Finance and Policy 1*(2), 176-216.

Brueckner, L. J., & Distad, H. W. (1924). The effect of the summer vacation on the reading ability of first grade children. *The Elementary School Journal, 24,* 698-707.

Bums, R. B., & Mason, D. A. (1995). Organizational constraints on the formation of elementary school classes. *American Journal of Education, 103,* 185-212.

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 specification and estimation. *Sociology of Education, 77*(1), 1-31.

Burtless, G. (1996). Introduction and summary. In G. Burtless (Ed.), *Does money matter? The effect of school resources on student achievement and adult success*. Washington, DC: Brookings Institution.

Carroll, J. B. (1963). A model of school learning. *Teachers College Record, 64,* 723-733.

Carroll, J. B. (1989). The Carroll model: A 25-year retrospective and prospective view. *Educational Researcher, 18*(2), 26-31.

Cepeda, N., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C., & Pashler, H. (2009). Optimizing distributed practice: Theoretical analysis and practical implications. *Experimental Psychology, 56*(4), 236-246.

Cepeda, N., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*(3), 354-480.

Clotfelter, C., Ladd, H., & Vigdor, J. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review, 26*(6), 673-682.

Cobb, J. A. (1972). Relationship of discrete classroom behaviors to fourth-grade academic achievement. *Journal of Educational Psychology, 63*(1), 74-80.

Cobb, J., A., & Hops, H. (1973). Effects of academic survival skills training on low achieving first graders. *Journal of Educational Research, 67*(3), 108-113.

Cochran, W. G., & Rubin, D. B. (1973). Controlling bias in observational studies: A review. *Sankhya, Series A, 35,* 417-446.

Coleman, J. S. (1966). *Equality of Educational Opportunity (COLEMAN) Study (EEOS).* [Computer file]. ICPSR06389-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2007-04-27. doi:10.3886/ICPSR06389

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, F., Mood, A. M., Weinfeld, F. D., et al. (1966). *Equality of educational opportunity* . Washington, DC: U.S. Government Printing Office.

Coleman, J. S. (1961). *The adolescent society: The social life of the teenager and its impact on education.* New York: Free Press.

Cook, T. D., Shadish, W. R., & Wong, V. C. (2008). Three conditions under which experiments and observational studies produce comparable causal estimates: New findings from within-study comparisons. *Journal of Policy Analysis and Management, 27*(4), 724-750.

Cooper, H. (1979). Pygmalion grows up: A model for teacher expectation communication and performance influence. *Review of Educational Research, 49*(3), 389-410.

Cooper, H., Charlton, K., Valentine, J. C., & Muhlenbrock, L. (2000). Making the most of summer school: A meta-analysis and narrative review. *Monographs of the Society for Research in Child Development, 65,* 1-130.

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.

Cooper, H., Valentine, J. C., Charlton, K., & Melson, A. (2003). The effects of modified school calendars on student achievement and on school and community attitudes. *Review of Educational Research, 73*(1), 1-52.

Cornell, E. H. (1980). Distributed study facilitates infants' delayed recognition memory. *Memory and Cognition, 8,* 539-542.

Daneshvary, N., & Clauretie, T. M. (2001). Efficiency and costs in education: year-round versus traditional schedules. *Economics of Education Review, 20*(3), 279-287.

Darling-Hammond, L. (1994). *Beyond the commission reports: The coming crisis in teaching,* Santa Monica, CA: RAND.

Dellarosa, D., & Bourne, L. E. (1985). Surface form and the spacing effect. *Memory and Cognition, 13,* 529-537.

Dempster, F. N. (1986). *Spacing effects in text recall: An extrapolation from the laboratory to the classroom.* Manuscript submitted for publication.

Dempster, F. N. (1987). Effects of variable encoding and spaced presentations on vocabulary learning. *Journal of Educational Psychology, 79,* 162-170.

Dempster, F. N. (1989). Spacing effects and their implications for theory and practice. *Educational Psychology Review, 1*(4), 309-330.

Diaz, J. J., & Handa, S. (2005). An assessment of propensity score matching as a nonexperimental impact estimator: Evidence from Mexico's PROGRESA program. *Journal of Human Resources, 41*(2), 319-345.

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology, 84,* 795-805.

Downey, D., von Hippel, P., Broh, B., & Powell, M. (2004). Are schools the great equalizer/School and non-school of inequality in cognitive skills. *American Sociological Review, 69*(5), 613-635.

Eccles, J. S., & Midgley, C. (1989). Stage/environment fit: Developmentally appropriate classrooms for early adolescents. In R. E. Ames & C. Ames (Eds.), *Research on motivation in education* (Vol. 3). New York: Academic Press.

Ehrenberg, R. G., Ehrenberg, R. A., Rees, D. I., & Ehrenberg, E. L. (1991). School district leave policies, teacher absenteeism, and student achievement. *The Journal of Human Resources, 26,* 72-105.

Elam, S. M., Rose, L. C., & Gallop, A. M., (1996). The 28th annual Phi Delta Kappa/Gallop poll of the public's attitudes toward the public schools. *Phi Delta Kappan, 78,* 41-59.

Elmes, D. G., Dye, C. G., & Herdelin, N. J. (1983). What is the role of affect in the spacing effects? *Memory and Cognition, 11,* 144-151.

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (1997). *Children, schools, and inequality.* Boulder, CO: Westview Press.

Entwisle, D. R., & Astone, N. M. (1994). Some Practical Guidelines for Measuring Youth's Race/Ethnicity and Socioeconomic Status. *Child Development, 65*(6), 1521-1540.

Estes, W. (1955). Statistical theory of distributional phenomena in learning. *Psychological Review, 62*(5), *369*-377.

Fardig, D. (1992). *Year-round education: Program evaluation report.* Orlando, FL: Orange County Public Schools. ERIC Document Reproduction Service No. ED 357 047.

Ferguson, R. F. (1991). Paying for public education: New evidence on how and why money matters. *Harvard Journal on Legislation, 28*(2), 465-498.

Ferguson, R. F., & Ladd, H. F. (1996). How and why money matters: An analysis of Alabama schools. In H. Ladd (Ed.), *Holding schools accountable* (pp. 265-298). Washington, DC: Brookings Institution.

Fisher, C., Berliner, D., Filby, N., Narliave, R., Cohen, L., Dishaw, M., & Moore, J. (1978). *Teaching and learning in elementary school: A summary of the beginning teacher evaluation study.* San Francisco, CA: Far West Laboratory of Educational Research and Development.

Fishman, E., Keller, L., & Atkinson, R. (1968). Masses versus distributed practice in computerized spelling drills. *Journal of Educational Psychology, 59*(4), 290-296.

French, D. A. (1992). *California administrator burnout and year-round schools.* San Diego, CA: ERIC Document Reproduction Service.

Gandara, P., & Fish, J. (1994). Year-round schooling as an avenue to major structural reform. *Educational Evaluation and Policy Analysis, 16*(1), 67-85.

Genesee, F. (1985). Second language learning through immersion: A review of U.S. programs. *Review of Educational Research, 55,* 541-561.

Glazerman, S., Levy, D. M., & Myer, D. (2003). Nonexperimental versus experimental estimates of earning impacts. *Annals of the American Academy of Political and Social Science, 589,* 63-93.

Glenberg, A. (1979). Components-level theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition, 7 (2),* 95-112.

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

Goldhaber, D. D., & Brewer, D. J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis,* 22 (2), 129-145.

Greenwald, R., Hedges, L., & Laine, R. (1994). When reinventing the wheel is not necessary: A case study in the use of meta-analysis in educational finance. *Journal of Education Finance, 20*(1), 1-20.

Greenwald, R., Hedges, L., & Laine, R. (1996). The effect of school resources on student achievement. *Review of Educational Research, 66,* 361-396.

Gump, P. V. (1967). *The classroom behavior setting: Its nature and relation to student behavior.* Lawrence, KS: University of Kansas Department of Psychology. ERIC ED 015 515.

Haertel, G. D., Walberg, H. J., Weinstein, T. (1983). Psychological models of educational performance: A theoretical synthesis of constructs. *Review of Educational Research, 53*(1), 75-91.

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis, 19*(2), 141-164.

Harris, A. J. (1979). The effective teacher of reading, revisited. *The Reading Teachers, 33*(2), 135-140.

Haser, S., & Nasser, I. (2003). Teacher job satisfaction in a year-round school. *Educational Leadership, 60*(8), 65-67.

Hawley, W. D., Rosenholtz, S., Goodstein, H., & Hasselbring, T. (1984). Good schools: What research says about improving student achievement. *Peabody Journal of Education, 61*(4), 1-178.

Henry, G. T. (2010). Comparison group designs. In J. S. Wholey, H. P. Hatry, & K. E. Newcomer (Eds.), Handbook of practical program evaluation (3rd ed.). San Francisco, CA: Jossey-Bass Publishers.

Henry, G. T., & Yi, P. (2009, November 7). *Design matters: A within-study assessment of propensity score matching designs.* Unpublished paper presented at the meeting of the Association for Public Policy Analysis and Management, Washington, DC.

Henry, G. T., Thompson, C. L., Fortner, C. K., Rickman, D. K., & Zulli-Lowe, R. (2008). *The impact of the disadvantaged student supplemental fund on high school student performance in pilot districts.* Chapel Hill, NC: Carolina Institute for Public Policy.

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

Heyns, B. (1987). Schooling and cognitive development: Is there a season for learning? *Child Development, 58,* 151-160.

Hintzman, D. L. (1974). Theoretical implication of the spacing effect. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium* (pp. 77-99). Potomac, MD: Erlbaum.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association, 81,* 945-960.

Hops, H., & Cobb, J. A. (1974). Initial investigations into academic survival-skill training, direct instruction, and first grade achievement. *Journal of Educational Psychology, 66*(4), 548-553.

Ingersoll, R. (2003). *Is there really a teacher shortage? A report co-sponsored by the Center for the Study of Teaching and Policy and the Center for Policy Research in Education.* Seattle, WA: University of Washington Center for the Study of Teaching and Policy.

Jencks C. (1972). *Inequality, a reassessment of the effect of family and schooling in America*. New York: Basic Books.

Johnston, K. L., & Aldridge, B. G. (1985). Examining a mathematical model of mastery learning in a classroom setting. *Journal of Research in Science Teaching, 22,* 543-554.

Karpicke, J. D., & Roediger, H. L. (2010). Is expanding retrieval a superior method for learning text materials? *Memory & Cognition, 38*(1), 116-124.

Karpicke, J. D., & Roediger, H. L. (2006). Repeated retrieval during learning in the key to long-term retention. *Journal of Memory and Language, 57,* 151-162.

Karweit, N. L. (1984). Time-on task reconsidered: Synthesis of research on time and learning. *Educational Leadership, 41*(8), 32-35.

Karweit, N. L. (1988). Time-on-task: The second time around. *National Association of Secondary School Principals, 72,* 31-39.

Karweit, N. L. (1993). *Summer achievement growth and the effectiveness of summer school.* Baltimore, MD: Johns Hopkins University for Research on Effective Schooling for Disadvantaged Students.

Karweit, N. L., & Slavin, R.E. (1982). Time-on-task: Issues of timing, sampling, and definition. *Journal of Educational Psychology, 74*(6), 844-851.

Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language, 32,* 1-24.

Kim, J. S. (2007). The effects of a voluntary summer reading intervention on reading activities and reading achievement. *Journal of Educational Psychology, 99*(3), 505-515.

Kneese, C. (1996). Review of research on student learning in year-round education. *Journal of Research and Development in Education, 29,* 60-67.

Kneese, C. (2000). *Year-round learning: A research synthesis relating to student achievement*. San Diego, CA: NAYRE.

Kornell, N. (2009). Optimising learning using flashcards: Spacing is more effective than cramming. *Applied Cognitive Psychology, 23,* 1297-1317.

Kounin, J. S. (1970). *Discipline and group management in classrooms.* New York: Holt, Rinehard and Winston.

Kounin, J. S., & Doyle, P. H. (1975). Degree on continuity of a lesson's signal system and task involvement of children. *Journal of Educational Psychology, 67*(2), 159-164.

Kounin, J. S., & Gump, P. V. (1974). Signal systems of lesson settings and the task-related behavior of preschool children. *Journal of Educational Psychology, 66*(4), 554-562.

Lamdin, D. J. (1996). Evidence of student attendance as an independent variable in education production functions. *Journal of Educational Research, 89,* 155-162.

Landauer, T. K. (1969). Reinforcement as consolidation. *Psychological Review, 76,* 82-96.

Levin, H. M. (1986). Are longer school sessions a good investment? *Contemporary Policy Issues, 4,* 63-75.

Macdonald, F. J., & Elias, P. J. (1976). *The effects of teaching performance on pupil learning: Beginning teacher evaluation study* (Phase II, Final report, Vol 1). Princeton, NJ: Educational Testing Service.

Magliero, A. (1983). Pupil dilations following pairs of identical words and related to-be-remembered words. *Memory and Cognition, 11,* 609-615.

McMillen, B. (2001). A statewide evaluation of academic achievement in year-round schools. *Journal of Educational Research, 95*(2), 67-74.

Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior, 9,* 596-606.

Mitchell, R., & Mitchell, D. E. (2005). Student segregation and achievement tracking in year-round schools. *Teachers College Record, 107*(4), 529-562.

Morgan, S. L., & Winship, C. (2007). *Counterfactuals and causal inference.* New York: Cambridge University Press.

Moss, V. D. (1996). The efficacy of massed versus distributed practice as a function of desired learning outcomes, and grade levels of the student. (Doctoral dissertation, Utah State University, 1995) *Dissertation Abstracts international, 56,* 5204.

National Association for Year-Round Education. (2007). *Statistical summaries of year-round education programs: 2006-2007.* Retrieved from http://www.nayre.org/STATISTICAL%20SUMMARIES%20OF%20YRE%202007.pdf.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform.* Washington, DC: U.S. Government Printing Office.

National Education Commission on Time and Learning. (1993). *Research findings.* Washington, DC: Author. ED 372 491.

National Education Commission on Time and Learning. (1994). *Prisoners of time.* Washington, DC: U.S. Government Printing Office.

National Institute of Education. (1978.) *Compensatory education study.* Washington, DC: U.S. Department of Health, Education, and Welfare.

Naus, M. J., Ornstein, P. A., & Aivano, S. (1977). Developmental changes in memory: The effects of processing time and rehearsal instructions. *Journal of Experimental Child Psychology, 23*(2), 237-251.

Nye, B., Hedges, L., & Konstantopoulos, S. (2002). Do low-achieving students benefit more from small classes? Evidence from the Tennessee class size experiment. *Educational Evaluation and Policy Analysis, 24*(3), 201-217.

Nygaard, D. (1974). Evaluations of year-round school programs. Washington, DC: Educational Research Service. ERIC Document Reproduction Services No. ED 170 926.

Orellana, M., & Thorne, B. (1998). Year-round schools and the politics of time. *Anthropology and Education Quarterly, 29*(4), 446-472.

Patterson, M. V., & Rensselaer, N. Y. (1925). The effect of the summer vacation on children's mental ability and on their retention of arithmetic and reading. *Education, 46,* 222-228.

Pelavin, S. H. (1979). *A study of year-round schools* (Final report, Vol. I). Menlo Park, CA: Stanford Research Institute.

Pelavin, S. H., & David, J. L. (1977). *Evaluating long-term achievement: An analysis of longitudinal data from compensatory education programs* (SRI Project No. 4537). Washington, DC: Educational Policy Research Center.

Pipho, C. (1999). Summer School: Rx for low performance. *Phi Delta Kappan, 81*(1), 7-8.

Pittman, R., Cox, R., & Burchfiel, G. (1986). The extended school year: Implication for student achievement. *The Journal of Experimental Education, 45*(4), 211-215.

Postman, L., & Knecht, K. (1983). Encoding variability and retention. *Journal of Verbal Learning and Verbal Behavior, 22,* 133-152.

Pyc, M. A., & Rawson, K. A. (2007). Examining the efficiency of schedules of distributed retrieval practice. *Memory and Cognition, 35*(8), 1917-1927.

Quinn, J., & Negron, R. (2009). *Expanded learning opportunities.* Children's Aid Society. Retrieved April 2010, from http://www.childrensaidsociety.org/community-schools/partnership-press/winter-2011/leadership-message.

Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed., Ch. 1-8). Thousand Oaks, CA: Sage.

Rea, C. P., & Modigliani, V. (1987). The spacing effect in 4- to 9-year-old children. *Memory and Cognition, 15,* 436-443.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417-458.

Roderick, M., Jacob, B. A., & Bryk, A. S. (2004). Summer in the city: Achievement gains in Chicago's summer bridge program. In G. Bornam & M. Boulay (Eds.), *Summer learning: Research, policies, and programs* (pp.73-102). Mahwah, NJ: Lawrence Erlbaum.

Rohrer, D., & Pashler, H. (2007). Increasing retention without increasing study time. *Current Directions in Psychological Science, 16,* 183-186.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrica, 70*(1), 41-55.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology, 66*(5), 688-701.

Sardo-Brown, D., & Rooney, M. (1992). The vote on all-year schools. *American School Board Journal, 179,* 25-27.

Schacter, J., & Jo, B. (2005). Learning when school is not in session: a reading summer day-camp intervention to improve the achievement of exiting First-Grade students who are economically disadvantaged. *Journal of Research in Reading, 28*(2), 158-169.

Sewell, W. H. (1964). Community of residence and college plans. *American Sociological Review, 29,* 24-38.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*(3), 417-453.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin Company.

Shepard, M. A., & Reed, M. (1975). *A research agenda for year-round schools: Executive Summary, Volume 1.* (Department of Health, Education, and Welfare. No. AAI-75-159). Cambridge, MA: Abt Associates.

Shields, C. M. (1996, January). *Year-round education: Is it worth the hassle?* Paper presented at the meeting of the University of British Columbia Robson Square Lecture Series, Vancouver, BC.

Shulman, L. S., & Carey, N. B. (1984). Psychology and the limitations of individual rationality: Implications for the study of reasoning and civility. *Review of Educational Research, 54,* 501-524.

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis.* New York: Dr. Appleton-Century Company.

Stallings, J. (1980). Allocated academic learning time revisited, or beyond time on task. *Educational Researcher, 9,* 11-16.

Stallings, J. (1975). Implementation and child effects of teaching practices in follow through classrooms. *Monographs of the Society for Research in Child Development, 40*(7/8), 1-133.

Stallings, J., Corey, R., Fairweather, J., & Needels, M. (1977). *Early Childhood Education classroom evaluation.* Menlo Park, CA: SRI International.

Stevenson, H. W., Lee, S., Lewis, C., & Rohlen, T. (1998). An examination of American student achievement from an international perspective. *Brookings Papers on Educational Policy, 1,* 7-52.

St. Gerard, V. (2007). Year round schools look better all the time. *Education Digest: Essential Readings Condensed for Quick Review, 72*(8), 56-58.

Toppin, T. C., & DiGeorge, W. (1984). The spacing effect in free recall emerges with development. *Memory and Cognition, 12,* 118-122.

Waddell, R. (1997). How to stop 'shrinking summer' topic of panel at WWA symposium. (longer school years decrease summer season; World Waterparks Assn). *Amusement Business, 109*(43), 3-4.

WakeCARES. (2008). *Working to provide equal education opportunities to all Wake County children.* Retrieved September 21, 2008, from http://www.wakecares.com/?page_id=20.

Wake County Public School System. (2008). *Year round education in Wake County public school system.* Retrieved September 29, 2008 from http://www.wcpss.net/history/year_round/index.html.

Walberg, H. J. (1986). Syntheses of research on teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 214-229). New York: Macmillan.

Walker, H. M., & Hops, H. (1976). Increasing academic achievement by reinforcing direct academic and/or facilitating nonacademic responses. *Journal of Educational Psychology, 68*(2), 218-225.

Wentzel, K. (1998). Social relationships and motivation in middle school: The role of parents, teachers, and peers. *Journal of Educational Psychology, 90*(2), 202-209.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91*(3), 461-481.

Wiley, D. E., & Harnischfeger, A. (1974). Explosion of a myth: Quantity of schooling and exposure to instruction, major educational vehicles. *Educational Researcher, 3*(7), 7-12.

Wiseman, A. W., & Baker, D. P. (2004). The American summer learning gap from an international perspective. In G. Bornam & M. Boulay (Eds.), *Summer learning: research, policies, and programs* (pp. 3-24). Mahwah, NJ: Lawrence Erlbaum.

Wooldridge, J. M. (2009). *Introductory Econometrics, 4th edition*. Masson, OH: South-Western Cengage Learning.

Zechmeister, E. B. & Shaughnessy, J. J. (1980). When you know that you know and when you think that you know but you don't. *Bulletin of the Psychonomic Society, 15,* 41-44.

Zykowski, J. L., Mitchell, D. E., Hough, D., & Gavin, S. E. (1991). *A review of year round education research*. Riverside, CA: California Educational Research Cooperative. (ERIC Document Reproduction Service No. ED330040)