PUTTING 'PRE' IN 'SCHOOL':
THE INSTITUTIONALIZATION OF PRESCHOOL IN ELEMENTARY SCHOOLS

Michael H. Little

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 (Policy, Leadership, and School Improvement).

Chapel Hill
2018

Approved by:

Lora Cohen-Vogel

Thurston Domina

Eric Houck

Kirsten Kainz

Matthew Springer

ABSTRACT

Michael Little: Putting ‘Pre’ in ‘School’:
The Institutionalization of Preschool in Elementary Schools
(Under the direction of Lora Cohen-Vogel)

This study examines whether the physical location of Pre-K programs—whether in elementary school buildings or stand-alone centers—leads to differences in student outcomes in elementary school. Over the past several decades, public investment in Pre-K programming has burgeoned. According to the National Institute for Early Education Research, 32 percent of all four-year-olds in the United States attend state-funded Pre-K programs (Barnett et al., 2016). As Pre-K programs are taken to scale across the United States, a key concern is ensuring that programs are high-quality and provide significant and persistent effects on children’s school readiness and early schooling outcomes. Researchers are working to identify the components that predict high-quality and effective Pre-K programs, including components such as teacher credentials and measures of classroom quality, so that policies can better promote high-quality programs. Unfortunately, research to date has revealed few consistent and reliable proxies for high-quality Pre-K programs. My dissertation seeks to further investigate potential components of high-quality Pre-K programs by studying the *physical location of Pre-K settings*—namely, whether or not Pre-K programs are located in elementary school buildings or stand-alone centers.

In order to provide evidence on the role of the physical location of Pre-K programs on differences in student outcomes in elementary school, I use a concurrent, explanatory mixed-methods design that combines nationally-representative, quantitative data with in-depth, qualitative interview data from school administrators and teachers in North Carolina. The

combination of quantitative and qualitative data enable me to provide a holistic analysis of the phenomena of the physical location of Pre-K programs by providing estimates of the effects of setting type on a range of student outcomes and also providing evidence on the potential reasons for the observed relationships.

In the quantitative portion of this dissertation, I used nationally-representative data from the Early Childhood Longitudinal Study- Kindergarten Cohort of 2011 to estimate the impacts of school-based versus non-school-based Pre-K on a range of academic and social-emotional outcomes from kindergarten entry through the third grade. Enabled by the robust set of covariates available in the dataset, I use new propensity score weighting methods that ensure balance on observables between treatment and control groups. In the qualitative portion of this dissertation, I collected data from a convenience sample of elementary school administrators and teachers in central North Carolina to provide in-depth information about the physical location of Pre-K programs. Finally, I concluded my analysis by considering the merged quantitative and qualitative data to explore areas of convergence and divergence.

From the quantitative analysis, I find little evidence that school-based Pre-K is predictive of differences in student outcomes in kindergarten through third grade. However, I do find suggestive evidence that co-location—wherein students who attend Pre-K in an elementary school building and remain in that building—have superior outcomes when compared to students who move to attend another elementary school after Pre-K. These findings largely cluster in the academic achievement domains. From the qualitative analysis, I find that there is significant variability between elementary schools in the extent to which schools engage with Pre-K programs in their buildings and support collaborative, vertically aligned environments. Together, these results indicate that the physical location of Pre-K programs, alone, is insufficient for

differentiating program quality. I conclude this dissertation with a discussion of my findings in relation to the existing literature, highlight the limitations of my study, and discuss directions for future research in the area of Pre-K location and program quality.

ACKNOWLEDGEMENTS

Many individuals and institutions have supported me in my academic career and in the completion of this dissertation. First, I would like to thank my husband, Dexter, who has stood by me throughout this process. Thank you for always offering support and encouragement. I would also like to thank my family—your steadfast interest and engagement in my work kept me motivated and focused on my goals. To my advisor, Lora, I cannot thank you enough. You put your students first, you advocate for them, and you let them know that you care. Beyond teaching me about education policy and developing my writing skills, you taught me how to be a good citizen in the academic community. Finally, to my other committee members, Eric, Kirsten, Matt, and Thad, thank you. Your varied areas of expertise have been immeasurably helpful in ensuring that all aspects of my dissertation are sound and something that I can be proud of.

In addition to these individuals, I would like to thank the Royster Society of Fellows for the extremely generous financial assistance provided throughout my graduate studies. The Royster Society of Fellows program enabled me to focus intently on my research and accomplish goals that would not have been possible without the support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Research shows that high-quality Pre-K, where Pre-K is defined as center-based education for three and four year olds, is a promising intervention that boosts children's school readiness (see Kholoptseva, 2016, for a comprehensive meta-analysis), ameliorates school readiness gaps (Reardon & Portilla, 2016), improves health outcomes in later life (e.g., Reynolds et al., 2007), and pays dividends to society through reduced incarceration rates, for example (Deming, 2009; Heckman et al., 2010). Many of these promising findings, however, come from a series of small experimental interventions that took place over 40 years ago, most notable, the Perry Preschool Project in Ypsilanti, Michigan and the Carolina Abecedarian Study in Chapel Hill, North Carolina. Since that time, publicly-funded Pre-K programs have been brought to scale in states across the country; today, nearly 32 percent of 4-year-olds are served through publicly-funded Pre-K programs (Barnett et al., 2017). Unfortunately, these programs often do not provide the promising impacts of the smaller, experimental interventions (Bailey et al., 2017; Yoshikawa et al., 2013). While many have shown initial, positive impacts on child outcomes, the impacts do not reliably persist into elementary school—a pattern referred to as the *fade-out effect*—and their long-term effects are unknown because states and localities have delivered these programs for a relatively short time (Bailey et al., 2017).

Researchers have been working in recent years to help identify the components that constitute high-quality Pre-K programs—components that help to overcome fade-out and instead sustain Pre-K effects. Researchers have focused on a range of components—including *structural* components of Pre-K, such as teacher degrees and class sizes (Early et al., 2006), *process*

components, such as the quality of teacher-child interactions (Claessens et al., 2014), and *temporal* components, such as the number of years or hours per day that a child attends Pre-K (Marcus-Jenkins et al., 2016; Reynolds et al., 2014). One understudied component relates to the *physical location of Pre-K settings*—namely, whether or not Pre-K programs are located in school buildings or stand-alone centers.

Statement of Purpose

The purpose of this dissertation is to shed light on the relative effectiveness of school-based versus non-school-based Pre-K on a range of child outcomes, including academic achievement and social-emotional learning, in elementary school as well as to provide information about the potential reasons for the observed relationships. In addition to the relative effectiveness of school-based versus non-school-based Pre-K, I also examine the role of co-location for Pre-K attenders, where co-location refers to children who attended school-based Pre-K and remained in that same school for kindergarten or longer. Drawing on theoretical and empirical literatures that suggest that high-quality Pre-K and alignment across the early grades are important for early learning, I hypothesize that school-based Pre-K is more beneficial for child outcomes than non-school-based Pre-K. I further hypothesize that Pre-K benefits will be most persistent for co-locaters.

Theoretical and empirical literatures point to the importance of aligned and coherent education in the early grades. As highlighted in the introduction, Pre-K effects often fade rapidly once children move into elementary school. The reasons for fade out, elaborated upon in subsequent chapters, can be attributed to either the quality of the Pre-K experience itself or to the quality of subsequent educational experiences. School-based Pre-K and co-location may play a role in mitigating fadeout through both of these pathways. First, there is evidence that school-

based Pre-K programs may be of higher-quality than non-school-based Pre-K settings and that school-based Pre-K programs are more aligned to elementary school in terms of instruction. Second, evidence suggests that co-location may facilitate practices that promote vertical alignment and better sustainment of Pre-K effects, such as easing the transition to kindergarten for children and enabling teacher collaboration across the Pre-K to kindergarten grade span. In sum, I expect school-based Pre-K and co-location to help facilitate a more aligned, coherent educational experience for children that helps promote consistent learning gains and mitigate the fade-out effect.

Significance

As the first study to formally investigate the relative effectiveness of school-based versus non-school-based Pre-K, this dissertation makes a number of important contributions to the field of early childhood education. These contributions can be seen as research, methodological, and policy contributions, respectively. I detail each in the sections below.

Research Contributions

From a research perspective, this study helps to extend the field's current work to articulate what exactly constitutes high-quality Pre-K. Research has demonstrated that some Pre-K programs can generate positive student outcomes that persist throughout the lifespan (Duncan & Magnuson, 2013). However, less research has been able to identify components within these programs that reliably predict positive student outcomes. This dissertation explores the extent to which the physical location of Pre-K programs—school-based or non-school-based— and co-location are predictors of Pre-K effectiveness, as measured by student academic and social-emotional outcomes throughout elementary school.

The study is also one of the first to qualitatively examine the practical, on-the-ground implications of providing Pre-K education within elementary school contexts (Desimone et al., 2004). This study probes the extent to which Pre-K is actively integrated into elementary school cultures, potential divisions that exist, and the factors that may promote or inhibit children from experiencing a high-quality, aligned early learning experience. It is this focus on alignment and coordination in the early grades that is the third key research contribution of this study— a focus on school-based-Pre-K as a means to mitigate the fadeout of Pre-K effects.

It is common for Pre-K program effects to rapidly fade— especially on cognitive and academic outcomes— once children enter elementary school (Bailey et al., 2017; Duncan & Magnuson, 2013). In addition to examining the relative effectiveness of school-based versus non-school-based Pre-K, this study explores the potential benefits of co-location. Here, the theory is that co-location will foster aligned and coordinated instruction that continually pushes children's development (Vgotsky, 1997). While the role of subsequent schooling experiences has been theoretically noted as a potential key factor in the sustainment of Pre-K treatment effects (Bailey et al., 2017), this study will be one of the first to empirically test the theory.

Methodological Contributions

There are two key areas where this study breaks methodological ground: (1) the mixed methods design and (2) the use of novel propensity score analysis techniques in the quantitative portion of the study. While mixed methods research is not new, it remains a relatively under-utilized research paradigm with which to probe research questions (Creswell & Clark, 2011). As the review of the literature in Chapter 2 will demonstrate, few studies use a mixed approach to answer research questions in the realm of Pre-K education. By using a convergent parallel mixed-methods design, where quantitative and qualitative data are analyzed in a parallel fashion

and then integrated at the end of the analysis, I demonstrate the utility of a mixed design for the early childhood education field. This design allows me, in a single study, to not only probe questions about the *effects* of school-based Pre-K attendance on student outcomes in elementary school, but also *why* these observed relationships may exist.

The second key methodological contribution is my use of propensity score analysis methods in the quantitative portion of the study. Since the 1980s, propensity score analysis has been a popular methodological tool to construct treatment and control groups as a function of observable covariates (Kainz et al., 2017). In order for propensity score analysis to be successful, the propensity score procedures must yield high-levels of balance on covariate measures between the treatment and control groups (Ho et al., 2007). In recent years, new estimation techniques, including machine learning, have helped to nearly ensure that analysts can achieve high levels of covariate balance. I explore a range of these techniques in my analysis in order to select the method that achieves the highest levels of covariate balance. Specifically, I use Entropy Balancing to derive my propensity score weights.

Policy Contributions

This study provides the first guidance to a simple question: What type of Pre-K setting is most beneficial for children on both academic and non-academic outcomes in elementary school? As Pre-K programs continue to be expanded by policy makers in states and municipalities, this research will help to inform basic decisions regarding where to locate Pre-K classrooms in order to maximize their benefit. Across the county, states and localities are working to expand the provision of Pre-K programs to more and more children. The evidence generated from this study helps to inform these decisions.

Similarly, this study also explores the benefits of co-location. The findings related to this second research question help inform policymakers as they consider the logistics of assigning children to Pre-K programs. For example, if remaining in the same school building yields more persistent Pre-K effects, policymakers should consider assigning children to Pre-K programs that are located in schools where they are most likely to attend elementary school. Evidence from North Carolina suggests that assignment to Pre-K programs often does not consider the residence of children's families and their likelihood of attending a certain elementary school (Cohen-Vogel, 2017). This study helps to clarify the extent to which co-location matters and illuminate the ways in which policymakers can respond in order to promote the sustainment of Pre-K effects.

Beyond basic evidence about "what works" with respect to the relative effectiveness of school-based versus non-school-based Pre-K and the role of subsequent attendance patterns, this study also provides evidence about the practical realities of locating Pre-K programs within elementary schools. The qualitative portion of this study explores potential "active ingredients" that help explain why school-based Pre-K and co-location may provide beneficial experiences for young children. As I find quantitatively, location of Pre-K programs in and of itself is likely not enough. Rather, the location of programs is a facilitator (or inhibitor) of practices within schools that yield differential student outcomes. This study helps illuminate what these specific elements are. In sum, this study not only provides evidence about where to locate Pre-K programs physically, but how to ensure that programming is most effective within a given setting.

Research Questions

As I will detail in subsequent chapters, my dissertation seeks to answer to the following five research questions:

1) Do children who attend school-based Pre-K have different academic and social-emotional outcomes in elementary school than children who attend non-school-based Pre-K?

2) Do children who co-locate—those who remain in the same elementary school where they attended Pre-K—have different academic and social-emotional outcomes than children who switch to a different elementary school after Pre-K or attend non-school-based Pre-K?

3) For both Research Questions One and Two, do the observed relationships vary across subgroups of interest? Specifically, do the results vary by socioeconomic status, Pre-K funding type, urbanicity, school quality, and measures of P-3 vertical alignment?

4) In what ways are Pre-K programs different when located in elementary schools versus stand-alone centers and how might co-location foster alignment and continuity?

5) What are the areas of alignment and misalignment between the quantitative findings for Research Questions One, Two, and Three and the qualitative findings for Research Question Four?

Roadmap

In the sections that follow, I begin in Chapter 2 with a review of the empirical literature related to Pre-K education, including consideration of the fadeout of Pre-K treatment effects and current evidence on school-based Pre-K. In Chapter 3, I formally outline my conceptual framework for the study, detailing how a combination of increased program quality and alignment between Pre-K and the early grades supports my hypothesis that school-based Pre-K and co-location will yield positive student outcomes in elementary school. In Chapter 4, I discuss

the study methodology. In Chapter 5, I detail my findings for research questions one through five. In Chapter 5, I discuss how my findings relate to the existing literature, limitations of my study, and directions for future research in the area of Pre-K location and program quality.

CHAPTER 2: REVIEW OF THE LITERATURE

Introduction

The purpose of this chapter is to provide an overview of the literature that has relevance for my proposed study of the effects of school-based Pre-K and co-location on child outcomes in elementary school. I begin by providing a brief overview of the early childhood education landscape in the United States and situate Pre-K education within this landscape. Next, I provide a review of the research on the effectiveness of Pre-K education on student outcomes. Here, I include meta-analytic research findings as well as research on early model Pre-K programs, Head Start, and state and local Pre-K programs. I then provide a review of research on how the effects of these programs varies based on different subgroups.

Having put Pre-K into context and reviewed research on its effectiveness, I then review literature on the "fadeout" of Pre-K treatment effects and the common explanations for why fadeout occurs. Next, I detail what research has revealed about predictors of Pre-K program quality (e.g., structural factors, such as teacher degrees; and process factors, such as teacher-child interactions). Finally, I consider research directly focused on the topic of school-based Pre-K. I begin by providing an overview of programs (past and present) that include school-based Pre-K and then follow by reviewing research on the effectiveness of these programs, where applicable.

The Early Childhood Education Landscape

Most children in the United States attend some form of center-based Pre-K in the year or two before entering kindergarten. As of 2015, 67 percent of four-year-olds and 38 percent of three-year-olds attended center-based Pre-K, whether public or private (Barnett et al., 2016). A

recent analysis by the Child Care Resource Center (CCRC) found that the average cost of center-based Pre-K in the United States per year is approximately $8,000 (CCRC, n.d.). Enrollment rates are highest for children from high-socioeconomic families, due in large part to the high costs of these programs. About half of the four-year-olds who attend center-based Pre-K are served by private programs and the other half are served by public programs. The three primary types of public programs include: Head Start, State/Local Pre-K, and childcare subsidies funded through the Child Care Development Block Grant.

Head Start

Founded in 1965 as part of President Lyndon B. Johnson's War on Poverty, Head Start is a federal Pre-K program that serves low-income children and their families (Kalifeh et al., 2011). Head Start is a holistic program that generally serves three- and four-year-old children in center-based settings and focuses on (1) early learning, (2) health, and (3) family well-being (Office of Head Start, n.d.). In fiscal year 2015, the Office of Head Start was appropriated approximately $8.6 billion to run the program. In the 2015-16 academic year, Head Start served 8 percent of both three- and four-year-olds in the United States. As state and local Pre-K programs have been taken to scale in recent decades, which primarily serve four-year-olds, Head Start has shifted to serve a higher proportion of three-year-olds. Children are eligible for Head Start based on family income below poverty thresholds that vary by state, homelessness, use of public assistance through Temporary Assistance for Needy Families (TANF) or Supplemental Security Income (SSI), and placement in foster care (Office of Head Start, n.d.).

State/Local Pre-K

There are currently 59 public state or local Pre-K programs operating in 43 states and Washington D.C. Unlike Head Start, state and local Pre-K programs tend to offer care only to four-year-old children in the year before kindergarten entry. Most programs are targeted—meaning they serve a subset of children, not all—based on family income and other risk factors, such as disability. But, differences across communities as well as differences in where income targets are set mean that access varies significantly across programs. For example, the Pre-K program in Washington D.C. serves over 80 percent of four-year-olds while the state-funded Pre-K program in Minnesota serves approximately one percent of four-year olds, representing the highest and lowest enrollment rates, respectively. Notably, the most recent expansions of Pre-K have taken place at the local level in urban areas, including universal Pre-K in New York City and Seattle in 2014 (Barnett, et al., 2016). Approximately 32 percent of all four-year-olds and five percent of all three-year-olds in the United States are served by state and local Pre-K programs for a total cost of approximately $7.4 billion (Barnett et al., 2016).

Child Care Subsidies

The final major public program supporting Pre-K education are child care subsidies funded through the Child Care Development Fund (CCDF). In fiscal year 2015 total funding for CCDF was $8.5 billion, which supported subsidies for approximately 1.4 million children in the United States (Office of Child Care, 2017). These subsidies are provided to parents of children to support attendance at private, center-based preschools, family day care, and other forms of informal care. Additionally, CCDF subsidies can be used to support children from birth through age 12. As a result, few of the studies highlighted in the subsequent sections focus on this program as it is so broad and variable in scope; however, I note it here given its large scale.

The Effects of Pre-K on Student Outcomes

Meta-Analytic Findings

I begin this section by focusing on aggregate findings across studies of the impacts of Pre-K on child outcomes. The scientific literature on the effectiveness of Pre-K interventions has become mature enough to enable meta-analysis, a method of combining effect estimates across studies to draw broader conclusions about impacts. The most recent and comprehensive meta-analytic studies come from a database of 84 studies of Pre-K interventions from 1960 to 2007 assembled by the National Forum on Early Childhood Policy and Programs (Duncan & Magnuson, 2013).

Averaging across 84 studies of Pre-K effects that met high-methodological standards for inclusion in the meta-analytic database, Duncan and Magnuson (2013) estimate the average effect of Pre-K treatment on cognitive and achievement outcomes at the end of Pre-K is 0.35 standard deviations, a magnitude that is equal to nearly half of the race-based school readiness gap (Duncan & Magnuson, 2011). In Figure 2.1 below, I reproduce a summary graphic of all of the 84 studies arrayed by year (X-axis) and effect sizes (Y-axis). Additionally, the plotted point for each study in the figure is sized to reflect its weighted contribution, as measured by the inverse of the squared standard error of the impact estimate.

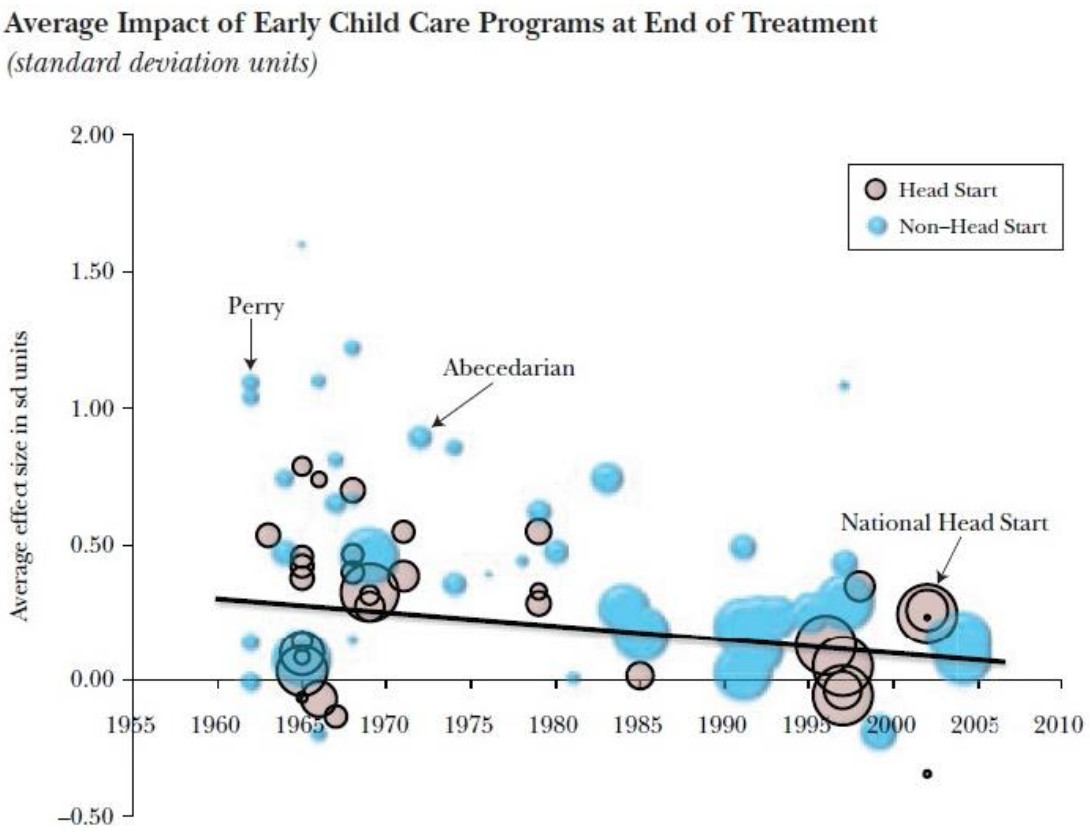

Figure 2.1 Average Impact of Early Child Care Programs at End of Treatment.

Reproduced from Duncan, G. J., & Magnuson, K. (2013). Investing in preschool programs. ***The Journal of Economic Perspectives*****, 27(2), 109-132.**

There are a few key lessons that Figure 2.1 reveals. The largest impact programs tend to be smaller in terms of the number of children served and include model programs such as Perry Preschool and Carolina Abecedarian. Relatedly, the smallest program impact estimates come from larger programs, such as state-level Pre-K programs and Head Start.

Another key lesson is that the magnitude of effect sizes has decreased over time. This latter finding is likely attributable to the shifting counterfactual conditions over time. A counterfactual represents what would be true if some treatment were not provided. In the context of Pre-K treatment, the counterfactual condition is what the child experiences if they did not receive Pre-K. Over the intervening decades covered in this meta-analysis, a number of broader

shifts may have improved the counterfactual experience for children, including use of other forms of childcare, higher-quality home environments, and more comprehensive social safety net programs (Duncan & Magnuson, 2013; Kline & Walters, 2016).

All of the aforementioned impact estimates are focused on the effects of treatment at the end of the program, often measured in the Spring of Pre-K or the Fall of kindergarten. The meta-analytic findings for longer-term schooling outcomes suggest persistent impacts of Pre-K on cognitive and achievement outcomes, but the magnitude of these effects fades over time. Specifically, Duncan and Magnuson (2013) find that program effect sizes, on average, decrease by approximately 0.03 standard deviations per year, and positive, statistically-significant effects remain for 10 years after the program ends. While statistically significant impacts may persist for 10 years, the largest drops in impact estimates occur in the first few years following treatment and the remaining impacts are small.

This finding of attenuation and eventual elimination of Pre-K treatment effects on cognitive and achievement outcomes, while also leading to long-term life outcomes such as improved educational attainment and lower crime incidence (Barnett et al. 2011), raises the question of whether or not Pre-K interventions are generating long-term effects through other outcomes, such as special education placement or conscientiousness, for example. Unfortunately, the studies in the meta-analytic database vary too extensively in the alternative outcomes measured and timing of measurement to generate estimates beyond cognitive and achievement outcomes. I return to the topic of Pre-K fadeout and the explanations for its existence in a later section of this chapter. I now move beyond aggregate, meta-analytic results to highlight research within specific Pre-K programs.

Model Programs

Perry Preschool

The Perry Preschool Project began in 1962 in Ypsilanti, Michigan with 162 high-risk African-American children. Of the 162 children, 58 received the program and 65 were in the control condition. The treatment was the provision of a high-quality center-based educational approach focused on academic and social development. Children attended the program five days a week for two and a half hours a day for two years. Additional services included weekly home visits by teachers and monthly parent group meetings. Data collection waves, which included a combination of child assessments and parent/participant surveys, occurred annually between the ages of four and 11, and then at ages 14, 15, 19, 27, and 40 (Belfield et al., 2006).

In terms of cognitive and achievement effects, the Perry Preschool program generated immediate, positive effects for treatment participants that quickly faded away. On the Stanford-Binet IQ test, treatment subjects performed nearly a standard deviation above control subjects at the end of the program, but this benefit faded and became statistically insignificant by the time subjects were eight years old (Schweinhart et al., 1980, 2005). Achievement outcomes, such as the Peabody Picture Vocabulary Test (PPVT), generally followed a similar pattern, where there were short-term benefits that faded to insignificance by six to nine years of age. A notable exception is the California Achievement Test (CAT) scores, which initially faded but positive, significant effects re-emerged for treatment subjects at the age of nine and ten and this benefit persisted through the age of 14 (Schweinhart et al., 1980, 2005).

In terms of non-cognitive or non-achievement outcomes, the benefits of the intervention were much more consistent and persistent. In terms of measures of academic success, treatment participants completed more years of total school (1 year), spent less time in special education

(1.3 years), and were 44 percent more likely to graduate from high school than control participants. While not directly a schooling outcome, the effects of treatment on crime involvement are particularly notable and generate the vast majority of returns to the program in cost-benefit studies (Belfield et al., 2006). Treatment participants were 46 percent less likely than control participants to have served time in jail or prison and were 33 percent less likely to be arrested for a violent crime (Schweinhart et al., 2005).

Carolina Abecedarian

The Carolina Abecedarian study, like Perry Preschool, is one of the oft-cited early Pre-K experiments. Abecedarian included a total of 111 participants, with 57 children receiving the treatment condition and 54 children receiving the control condition. The program began in 1972 and treatment consisted of comprehensive early education services from infancy through age five (excluding the K-2 follow-on supplement that was re-randomized in kindergarten). The program provided year-round, full-time, center-based care that included transportation, low child-teacher ratios, health and social services, and parental outreach services. Compared to Perry, Abecedarian was a more comprehensive and intensive treatment program (Ramey & Campbell, 1972).

Since the study began in 1972, there have been numerous follow-ups with study participants. A distinguishing factor of the Abecedarian study is that its impacts on cognitive and achievement outcomes were more prominent and persistent than in Perry Preschool. At kindergarten entry, the difference between the treatment and control group on an IQ assessment was nearly an entire standard deviation. Even by age 21, impacts of Abecedarian treatment on IQ scores was 0.38 standard deviations (Campbell & Ramey, 1994). These persistent effects were also present for academic achievement, as measured by the Woodcock-Johnson

Psychoeducational Battery. At age 12, achievement differences between treatment and control groups was approximately 0.40 standard deviations for both mathematics and reading achievement (Campbell & Ramey, 1994).

The Carolina Abecedarian program also generated a host of impacts on non-cognitive or non-academic outcomes during the schooling years (Campbell et al., 2012). In middle adolescence, children who participated in the treatment condition were less likely to be retained a grade and were less likely to have been assigned to special education (Campbell & Ramey, 1996). Children who participated in the treatment condition, by 21 years of age, completed more years of school and were more likely to enroll in college (Campbell et al., 2002). Other outcomes include higher likelihood of holding a job and lower likelihood of teen parenthood and drug use (Campbell et al., 2002; Pungello et al., 2010).

Chicago Child-Parent Centers (CPC)

The CPC is often grouped as an early model program but it is distinct in that it is (1) not a randomized control trial and (2) operated and continues to operate at a much larger scale than both Perry and Abecedarian. While the primary focus of the CPC program is the Pre-K component for four-year olds, for some participants, the program extended through age nine and included social and health services, as well as parental outreach initiatives. The Pre-K portion of the intervention was a part-day program that lasted, on average, 1.5 years before kindergarten entry. The curricula focused on language and classrooms had teacher-child ratios of 8.5:1, on average. Longitudinal research on the CPC comes from a sample of approximately 1600 low-income children who began attending the program in 1983-1984. While not randomized, 989 participants received the CPC program and a similar comparison group of 550 children was included (Reynolds et al., 2007).

A number of longitudinal evaluations of the CPC have been conducted over the years and have revealed positive effects on a range of academic and non-academic outcomes. In terms of cognitive and academic outcomes, Reynolds (1994) finds that children who participated in the program scored higher on assessments of math and reading achievement in the fifth grade. By the seventh grade, however, only the positive achievement effects for reading (0.43 SD) remained, though the insignificant effect size for math was 0.28 standard deviations (Reynolds & Temple, 1998). Summarizing the effects on non-cognitive and non-achievement outcomes from previous evaluation studies for a cost-benefit analysis, Reynolds et al. (2007) finds that children who participated in CPC were less likely to receive special education services in school (-11 percent), be retained (-15 percent), be arrested (-8 percent), and were more likely to complete high school (+11 percent) and attend college (+7 percent).

Head Start

As described in the introduction, Head Start, founded in 1965, is the largest and longest-lasting scaled-up Pre-K program in the United States. A wealth of evidence suggests that Head Start has positive, short-term effects on cognitive, social-emotional, health, and parenting outcomes (Lee et al. 1988; McKey et al., 1985; Puma et al., 2010). The evidence on the effects of Head Start can be divided into two primary groups: research from the Head Start Impact Study, a randomized control trial evaluation of Head Start that began in 2002, and other evaluation research. I begin with the latter and then turn to evidence from the Head Start Impact Study.

Currie and Thomas (1995) used a sibling-fixed effects strategy that compared student outcomes between siblings who did and did not attend Head Start. The logic behind this empirical strategy is that by using variation within families rather than across families, many home environment-related factors that may confound the relationship between Head Start and

child outcomes are accounted for. These authors find that at age six, children who attended Head Start scored significantly higher on vocabulary and reading tests than their siblings who did not. Another study by Deming (2009) also used the sibling-fixed effects model and examined a host of young adult outcomes and finds that Head Start participants gain 0.23 standard deviations on a summary index of these outcomes, with particularly strong effects for the most economically-disadvantaged children. These effects, on grade repetition, disability placement, high school graduation, and college attendance, are about 80 percent as large as gains from the smaller Perry Preschool and Carolina Abecedarian model programs.

A study by Garces et al. (2002) provides evidence specifically on later-schooling outcomes of Head Start on children who participated in the program in 1980 or earlier. Using the Panel Study of Income Dynamics and estimating effects using a sibling-fixed effects approach, the authors find that, for white students, Head Start attenders were more likely to complete high school and enroll in college. For example, white children who attended Head Start were 30 percent more likely to complete high school than their siblings who did not attend Head Start. However, this effect was not present for other racial subgroups. Garces et al. (2002) also finds that Head Start attenders were less likely to be booked or charged with a crime in school, and this effect was mostly driven by Black Head Start attenders. Specifically, the authors find that Black Head Start attenders were 12 percent less likely than their siblings who did not attend Head Start to be booked or charged with a crime by their early twenties, when the outcome data were collected from participants.

A study by Ludwig and Miller (2005) differs from the previous studies in that the authors do not use the sibling-fixed effects strategy; rather, they exploit a discontinuity in how funding was allocated to programs when Head Start was founded in 1965. Similar to Garces et al. (2002),

this study found positive effects of Head Start on high school completion, but they found positive effects for White *and* Black children. Additionally, Ludwig and Miller (2005) found that a 50-100 percent increase in Head Start funding is associated with an increase in the likelihood of attending some college of about 15 percent.

More recent evidence on the effectiveness of Head Start comes from the Head Start Impact Study (HSIS) that began in 2002. The HSIS is a nationally-representative randomized control trial that included both 3- and 4-year old cohorts. The HSIS found positive impacts of Head Start on academic outcomes at the end of the Head Start program ranging from 0.15 to 0.30 standard deviations, but there were no positive effects detected after kindergarten. In terms of social-emotional impacts, the study found no impacts immediately following Head Start or at the end of kindergarten. By the end of third grade, the effects of Head Start on social-emotional outcomes were mixed, with participants having lower levels of problem behaviors and aggression, but more emotional symptoms and poorer student-teacher relations (Puma, 2010, 2012).

The aforementioned program estimates from the HSIS report the average effects of Head Start on participant outcomes. A series of studies have used the HSIS data to study the distributional effects of Head Start. In the first study, Bitler et al. (2014) use data from the HSIS to document the distributional effects of Head Start across individual children. The authors use data from the three-year-old HSIS cohort and explore heterogeneous treatment effects using quantile treatment effects and mean treatment effects for subgroups. Generating treatment on the treated estimates (those that take up the offer of Head Start), they find that Head Start effects are generally largest at the bottom of the distributions of each achievement outcome. For example, for the Peabody Picture Vocabulary Test, they find that program impacts of over a standard

deviation at the bottom of the distribution and a quarter of a standard deviation at the top of the distribution. They also find pronounced benefits of Head Start for Hispanic students and students with limited English proficiency.

In the second study from the HSIS, Bloom and Weiland (2015) extend the work of Bitler et al. (2014) to investigate variation in program effects across program sites in addition to variation across individual children and policy-relevant subgroups. First, and consistent with Bitler et al. (2014), Bloom and Weiland (2015) find that Head Start serves a compensatory role, where program effects are largest for children at the bottom of the cognitive outcome distribution. Additionally, Bloom and Wieland (2015) find that Head Start effects are largest for dual language learning and Spanish-speaking students. Probing alternative explanations for why there were such pronounced effects for this subgroup, the authors argue it is due to their limited prior exposure to English, not systematic difference in counterfactual conditions.

Turning to Bloom and Weiland's (2015) analysis of variation in effects across Head Start program sites, they find a significant amount of variation in Head Start program effectiveness across five of the six outcomes studied. They note that their study, "verifies what has been hypothesized for decades: that this large-scale, nationally funded, locally implemented program (with 1,800 grantees and 16,000 centers at present) produces results that vary widely relative to those of competing local alternatives" (p. 30). For example, they find standard deviations of program effect estimates to range from 0.12 to 0.25 for the five significant outcome measures.

One important finding from this research is that there is significant variation around outcomes with significant, positive grand mean effects as well as outcomes with near-zero grand mean effects. So, for example, the HSIS found a null and near zero grand mean effect estimate of program participation on self-regulation. However, the standard deviation of this estimate is 0.22

and is statistically significant, which means that the average impact estimate masks significant heterogeneity. Indeed, there are many sites where Head Start has large, positive effects on self-regulation and there are many sites where Head Start has large, negative effects on self-regulation. Conversely, Head Start has a large, significant grand mean effect on the outcome of numeracy, but there is little or no observable cross-site variation in program effects for this outcome (Bloom & Weiland, 2015).

State and Local Pre-K

Evaluations of state and local Pre-K programs are the most recent addition to the body of literature on Pre-K effects because they are the newest form of program. Most of the evidence comes from a small subset of states that have established Pre-K programs, including Florida, Georgia, North Carolina, Oklahoma (Tulsa), and Tennessee. There is also recent evidence on Boston's Pre-K program. In addition to state- and local-specific evaluations, there is also some evidence on Pre-K effectiveness nationally through analysis of nationally-representative datasets, such as the ECLS-K. I begin by first detailing the state- and local-specific evaluations.

Florida

Two studies have examined the effectiveness of the Florida Pre-K program on children's later schooling outcomes. The first study by Figlio and Roth (2009) used an instrumental variables approach to estimate the effects of the program on disciplinary outcomes in kindergarten through second grade and found that attending the Pre-K program was associated with fewer disciplinary infractions in these grades. The second study, by Miller and Bassok (2017), focused on the effects of Florida's Pre-K program on grade retention in kindergarten through third grade. The authors found that by the end of kindergarten, students who participated in Pre-K were less likely to be retained in kindergarten than children who did not attend Pre-K.

However, this advantage was erased by the third grade because retention rates for children who attended Pre-K accelerated to parity in grades one through three (Miller & Bassok, 2017).

Georgia

A study by Fitzpatrick (2008) examined the effects of Georgia's Pre-K program on a range of outcomes measured in fourth grade. Using a difference-in-differences approach, Fitzpatrick finds positive effects of the program on academic achievement. Specifically, the program is associated with positive, though small, effects on math (0.03 SD) and reading (0.01 SD) National Assessment of Educational Progress (NAEP) scores. Additionally, the study revealed that students who participated in the Pre-K program were more likely to be at grade level for their age in the fourth grade.

North Carolina

A series of recently published studies have evaluated the effects of two early education initiatives in North Carolina: Smart Start and More at Four. More at Four, now known as NC Pre-K, is the state's Pre-K program that was established in 2001. Smart Start is a birth through age four program focused building community supports to ensure children enter school healthy and ready to learn. Where possible, I focus only on the NC Pre-K results, but it is challenging to disentangle the programs as these studies focus on the impacts on all students, not just those that were directly involved in each program (Dodge et al., 2017).

The first of three studies in this line of research focused on third grade outcomes (Ladd et al., 2014). Exploiting variation in the rollout of programs across counties as well as variation in the financing of programs, the researchers utilize fixed effects models to examine the effects of each program on third grade math and reading achievement. The authors found positive effects of roughly equal magnitude of NC Pre-K on math and reading achievement in third grade. The

authors translate the magnitude of the effects of the program to the equivalent of two to four months of instruction in the third grade. Another paper in this line of research by Dodge et al. (2017) extended the 2014 paper by examining math and reading achievement through fifth grade as well as including outcomes of grade retention and special education placement. In terms of achievement, the authors find stable effects for both math and reading that approximate 0.20 standard deviations. Notably, the effects did not appear to fade between grades three through five. Finally, Dodge et al. (2017) also find that NC Pre-K was associated with a lower likelihood of being placed into special education as well as being retained.

The third study in this line of research evaluating the NC Pre-K program is by Mushkin et al. (2015) and uses the same methods and data as the previous two studies. In this study, the authors estimate the impact of NC Pre-K on the likelihood that students are placed in special education in the third grade. The results suggest that for every $100 invested in the NC Pre-K program, the odds of a child being placed in special education in third grade are reduced by 3.47 percent. Note again that these North Carolina studies report the effects of the program on all students, not just those participating directly in the NC Pre-K program (or Smart Start).

Oklahoma (Tulsa)

There are two primary studies from Oklahoma that evaluate the Pre-K program in Tulsa. While the immediate-term effects of the Tulsa program have been evaluated extensively using an age-cutoff regression discontinuity design (Gormley & Gayer, 2005), these two studies use propensity score matching techniques that enable comparisons longitudinally. The first, Hill et al. (2015), examined the effects of the Tulsa Pre-K program on third grade math and reading outcomes for a single cohort. The authors found positive effects on math (0.18 SD) and null effects on reading. The second study, Phillips et al. (2016), examined the effects of the program

on eighth grade achievement and non-academic outcomes. The study revealed positive, albeit smaller, effects on math (0.13 SD) and null effects on reading. It also indicated that attending the Tulsa Pre-K program was associated with lower odds of being retained (OR=0.60) and being chronically absent (OR=0.54).

Tennessee

While the research is not yet published in a peer-reviewed publication, a recent evaluation by Lipsey, Farran, and Hofer (2016) has gained significant notoriety so I detail it here. This study reports on the third grade follow-up of a statewide evaluation of the Tennessee Pre-K program. The evaluation used a randomized control trial (RCT) approach by randomly assigning children to attend the Pre-K program in settings where demand for Pre-K exceeded supply. The design is desirable because the process of randomization helps to ensure that treatment and control group participants are equal on all measures except for the treatment— TN Pre-K. At the end of the Pre-K year, Lipsey et al. (2016) find large, positive effects of Pre-K on achievement (0.32 SD), grade preparedness (0.22 SD), interpersonal skills (0.19 SD), and work-related skills (0.20 SD). However, by the end of the third grade, all positive effects on non-academic outcomes are insignificant and the effects on achievement are negative (-0.13 SD).

Boston

Research by Weiland and Yoshikawa (2013) evaluates the effectiveness of the Pre-K program in Boston. Their impact evaluation used an age-eligibility cutoff regression discontinuity design with a sample of approximately 2000 students from the 2008-2009 academic year. Weiland and Yoshikawa (2013) examined a robust set of outcome domains, including language, literacy, mathematics, and emotional development. The authors found large, positive effects of Boston's Pre-K program on language/literacy and mathematics (0.45-0.62

SD), the two domains targeted by the intervention, and smaller but significant impacts on emotional development (0.19-0.03 SD) and executive functioning (0.20 SD). Additionally, subgroup analyses revealed more positive benefits of the program for Hispanic students. However, unlike much of the research previously detailed (e.g., Bitler et al., 2014; Gormley et al., 2005), the authors do not find differential effects for socio-economically disadvantaged children.

National Pre-K Evidence

Having discussed the state-specific evaluations, I now detail two studies that leverage the nationally-representative ECLS-K dataset to estimate the effects of Pre-K on schooling outcomes. Magnuson et al. (2007) analyzed data from the first cohort of the ECLS-K, which followed a sample of children who attended kindergarten in the 1998-99 academic year. These authors use a range of identification strategies, including instrumental variables and propensity score matching. At kindergarten entry, they find positive effects of Pre-K on reading (0.12 SD) and math (0.10 SD) assessment outcomes. They also find that Pre-K is associated with lower self-control (-0.07 SD) and higher levels of externalizing problem behaviors (0.11 SD). Looking at the same outcomes in the Spring of first grade, Magnuson et al. (2007) finds null effects on both math and reading achievement, but persistence of the negative effects on social-emotional outcomes (self-control -0.13 SD; externalizing problem behaviors 0.21 SD).

A working paper by Bassok et al. (2015) conducts a similar analysis but uses the newest cohort of the ECLS-K, which followed a sample of children who attended kindergarten in the 2010-11 academic year. This study examined academic and social-emotional outcomes in the Fall of kindergarten, Spring of kindergarten, and Spring of first grade. Reading outcomes include positive effects on reading achievement in the Fall (0.11 SD) and Spring (0.06) of kindergarten,

but null effects by the end of first grade. Math outcomes include positive effects in the Fall of kindergarten (0.08 SD), but no significant effects thereafter. Social-emotional effects include positive effects on externalizing problem behaviors that are stable at approximately 0.11 standard deviations across the three time points (positive effects suggest higher levels of problem behaviors, which is a practically negative result). The authors also examined teacher-reported self-control and found that Pre-K participation was associated with lower levels of self-control in the Spring of kindergarten (-0.10 SD) and first grade (-0.10 SD).

Conclusion

Overall, there is significant variability in the different types of Pre-K program types that children may attend. In addition to differences in program types, there is also a lot of variation across state contexts and the populations of children that these programs serve. As a result, it is difficult to draw simple, clear-cut conclusions from research that examines these diverse programs and contexts. With this caveat in mind, there are some key lessons. In general, Pre-K programs have positive, albeit short-term effects on children's cognitive and achievement outcomes. In some cases, effects on cognitive and achievement outcomes persist through elementary school and middle school, though there is significant variability in this finding across studies. During elementary and middle school, Pre-K effects generally persist on outcomes such as grade retention and special education placement. While the majority of the long-term evidence is from early model programs, these effects include lower crime incidence, better health outcomes, and higher educational attainment. Aside from the early model programs, studies of Head Start have also shown positive effects on long-term life outcomes, including lower crime incidence and higher educational attainment. Long-term evidence on state and local Pre-K programs is not yet available because these programs have not existed for enough time.

Subgroup Effects of Pre-K

Thus far, I have focused on the average treatment effects of Pre-K programs on children's schooling outcomes. There is a related line of research that has examined if Pre-K programs may be differentially effective for different subgroups of children. There are three key subgroups that researchers have examined consistently in the literature: race, income/SES, and gender. I now turn to highlight the key findings for each subgroup.

A recent review by Ladd (2017) detailed the consensus findings on the differential effects of Pre-K on children from different income/SES strata and racial and language subgroups. Ladd (2017) focused on 13 studies of Pre-K effectiveness that included subgroup analyses and used rigorous identification strategies that enable causal estimates. In terms of race/ethnicity, the subgroup that appears to benefit the most from Pre-K is Hispanic children. These subgroup differences are apparent in the short term (Gormley, 2008; Weiland & Yoshikawa, 2013), as well as later in elementary school (Muschkin et al., 2015; Ladd et al., 2014). Ladd (2017) offers two possible reasons for this positive finding for Hispanic students. First, Hispanic students are the least likely subgroup to attend Pre-K, so it is likely that the contrast between Pre-K and counterfactual experience is stark (i.e. Pre-K or home care versus Pre-K or other center-based care). Second, the positive benefits of Pre-K for Hispanic children may also be related to exposure to the English Language. Gormley (2008) found in Tulsa, for example, that English test gains were largest for Hispanic students whose parent spoke Spanish at home or were born in Mexico.

For Black students, the evidence is more mixed on the differential effects of Pre-K. One positive finding comes from the Tennessee Pre-K evaluation. Recall that nearly all outcome effects had faded or switched direction by the third grade in this study. However, for Black

students, the positive effects of Pre-K on social-emotional skills persisted though the third grade (Lipsey et al., 2016). However, in terms of cognitive and achievement effects, studies show both positive (Gormley et al., 2005) and null effects (Ladd et al., 2014; Weiland & Yoshikawa, 2013) for Black students. Thus, a clear signal about the differential effects of Pre-K for Black students is inconclusive.

In terms of income or SES, studies generally show more positive effects for lower-income/SES children. These positive effects for economically-disadvantaged children have been shown in studies of Head Start (Bitler et al., 2014; Deming, 2009), Tulsa's Pre-K program (Gormley, 2005), North Carolina's Pre-K program (Dodge et al., 2017; Ladd et al., 2014), as well as the nationally-representative Early Childhood Longitudinal Study (Bassok, 2010; Magnuson et al., 2007).

In terms of gender, there are mixed results regarding the differential effects of Pre-K on child outcomes. A recent meta-analysis by Magnuson et al. (2016) that draws on the database of studies mentioned earlier in this paper finds few significant differences between male and females in terms of Pre-K effects. On cognitive and achievement outcomes, the authors find a statistically significant but very small benefit for girls relative to boys (0.03 SD). In terms of other school outcomes, which is a composite that combines special education placement and grade retention into a single measure, there are large differences between males and females. Specifically, the difference in the effect size for the benefits on other school outcomes between males and females is 0.56 standard deviations.

Fadeout of Pre-K Treatment Effects

As the review of the literature on the effectiveness of Pre-K programs has demonstrated, longitudinal evaluations of Pre-K interventions are often characterized by significant, positive

effects immediately following treatment with a geometric decline in effects thereafter. Averaging across studies in the meta-analytic database detailed in the previous section, Bailey et al. (2017) found that effects of Pre-K on student cognitive outcomes was 0.23 standard deviations at the end of treatment, 0.10 standard deviations after one year, 0.09 standard deviations after two years, and statistically insignificant thereafter. As noted earlier, Duncan and Magnuson (2013) estimate initial program impacts of 0.35 standard deviations. Longitudinally, they estimate that program effect sizes, on average, decrease by approximately 0.03 standard deviations per year, and positive, albeit small statistically-significant effects remain for 10 years after the program ends. Despite some difference in estimates between these two studies, the general trend, for achievement and cognitive outcome measures, is a rapid fadeout in the magnitude of Pre-K treatment effects, though small, statistically significant effects persist for ten years, on average.

What explains *Pre-K fadeout*? Researchers from disciplines spanning economics and developmental psychology have offered a range of explanations. In this section, I detail three central explanations that have emerged in the literature surrounding Pre-K fadeout effects: (1) skill building, (2) foot-in-the-door, and (3) sustaining environments explanations.

The Skill Building Explanation

From economics, Cuhna and Heckman (2007) have popularized a skill building model of human capital production. There are two key components of the Cuhna and Heckman (2007) model of skill formation. The first is the concept of *self-productivity*, which means that the skills acquired at one time point augment the skills acquired at subsequent time points. For example, it may be that early academic motivation skills augment later acquisition of academic content by making the student more engaged and eager to acquire new content. The second key component of the model is the concept of *dynamic complementarity*, which means that skills acquired at one

stage increase the productivity of investment at later stages. This component implies the need for early human capital investments to be followed by later investments in order for early investments to be productive. In the words of Cuhna and Heckman (2007), "Together, dynamic complementarity and self-productivity produce multiplier effects which are the mechanisms through which skills beget skills and abilities beget abilities" (p. 35). In terms of Pre-K fadeout, this perspective suggests that early investments (e.g., Pre-K quality) will make later investments (e.g., K-12 school quality) more productive and, conversely, early investments are less likely to be productive if they are not followed by subsequent investments.

A recent study by Jackson and Johnson (2017) provides empirical support for the Cuhna and Heckman (2007) skill building model. Jackson and Johnson (2017) study the interactive and synergistic effects of the early human capital investment of Head Start and later human capital investment of school finance reforms that led to higher school funding. The authors find that both investments—increases in Head Start and K-12 school funding—independently provide positive effects on outcomes such as educational attainment and earnings. However, when the two investments are coupled, the effects are even larger. For example, a $1000 per pupil increase in Head Start funding increases education attainment by 0.096 years and increases adult wages by 1.9 percent. However, for this same increase in Head Start funding, if followed by K-12 spending at the 75^{th} percentile of the distribution, educational attainment increases by 0.22 years and adult wages increase by 5.6 percent (Jackson & Johnson, 2017).

Bailey et al. (2017) have extended the work of Cuhna and Heckman (2007) to suggest that the skills being built themselves should be considered. Bailey et al. (2017) suggests that interventions should target "Trifecta Skills", namely skills that are (1) malleable, (2) fundamental, and (3) would not develop in counterfactual conditions. Malleable skills are those

that can be altered by the provision of some treatment, such as math and literacy skills, as opposed to less malleable skills, such as conscientiousness. Bailey et al. (2017) stress that malleability is not dichotomous, with some skills susceptible to intervention and others not; rather, some skills *are more likely* to be influenced by treatments than others.

In addition to the malleability of skills, skills targeted by an intervention are likely to be sustained if the treatment targets skills that are fundamental. Fundamental skills, as defined by Bailey et al. (2017), are skills "upon which later skills are built, and that influence positive life outcomes, such as attainment or labor market success" (p. 13). Examples of fundamental skills include academic motivation, social skills, and self-regulation, for example. This element of the trifecta skills aligns with Cuhna and Heckman's (2007) notion of self-productivity. Again, fundamentality does not operate independently from malleability. For skills to be sustained well beyond a given treatment intervention, the treatment should focus on skills that are *both* malleable and fundamental.

The third factor that Bailey et al. (2017) assert is critical for the sustainment of intervention treatment effects is that the skills targeted should be skills that do not readily develop in counterfactual conditions. For example, basic academic skills develop quite rapidly in counterfactual conditions. Hill et al. (2008) found that on nationally-normed reading and mathematics assessments, children gained over a standard deviation between kindergarten and first grade alone. So, despite the fact that basic academic skills are both malleable and fundamental for later success, they do not meet the trifecta skill criteria of counterfactual conditions. For example, if a Pre-K treatment focused on early mathematics and reading skills, the effects on assessments of these outcomes may be large at the end of Pre-K, but dramatically

attenuated by first grade as children in the counterfactual condition have an opportunity to rapidly develop these skills.

To summarize, the skill building perspective as outlined by Cuhna and Heckman (2007) suggests that skill building is a dynamic and complementary process wherein early skills predict later skills and the productivity of later skills are a function of earlier skills. However, this perspective is limited in that it refers to skills in a general sense and does not provide guidance on the specific types of skills that should be targeted. Bailey et al. (2017) address this limitation and detail criteria about the skills themselves that are likely to be sustained after some intervention. In order for intervention treatment effects to be sustained, interventions should target "trifecta skills," which are (1) malleable, (2) fundamental, and (3) unlikely to develop in counterfactual conditions.

The Foot-in-the-Door Explanation

The second common explanation for the fadeout phenomenon is the Foot-in-the-Door (FITD) perspective that focuses on the precise timing of the intervention. The FITD perspective focuses on developmental timing and posits that intervention impacts are most likely to persist when applied at critical periods or junctures in development (Bailey et al., 2017). Interventions that seek to delay the onset of certain behaviors in the teenage years are common examples of this perspective, such as pregnancy prevention and drug-use prevention. The theory underpinning these interventions is that if an intervention can delay the onset (or promote) of some behavior during a time when initiation can take place, the developmental trajectory of a child can be altered in the long term through increased educational attainment and positive acclimation into the labor market, for example (Bailey et al., 2017; Dolin & Booth-Butterfield, 1995).

While many FITD interventions described in the literature take place in middle and late adolescence, examples from Pre-K interventions focus on special education placement and grade retention. Findings from the meta-analytic database described above indicate that, across all included studies, the effects of Pre-K on reducing placement in special education and grade retention are between 0.3 and 0.4 standard deviations (Bailey et al., 2017). It is possible that it is through these mechanisms that some Pre-K interventions have lasting effects, especially in terms of long-term effects such as graduation rates and postsecondary enrollment. If Pre-K makes children more likely to get onto a mainstream academic trajectory early, it is possible to see how this FITD intervention can have effects that cascade over time.

It is difficult, however, to disentangle FITD and skill building processes in Pre-K interventions because the processes are likely simultaneously operating. However, there is some empirical evidence to support the FITD perspective from Deming's (2009) analysis of Head Start data. Deming (2009) used a sibling fixed effects identification strategy that compared variation in child outcomes within siblings who did and did not attend Head Start. The findings indicate fadeout of Head Start treatment on children's cognitive assessment scores but positive, significant, and longitudinal effects of Head Start on both grade retention and placement into special education.

The FITD perspective is related to the skill building perspective in that it represents a developmental cascades perspective wherein intervention effects last over time and interact with subsequent development stages. They differ, however, in how directly they influence development. Skill building interventions are direct because they target early skills that are built upon by one another in a similar construct progression. FITD interventions are indirect because

they alter transitory skills at an opportune time that alters an individual's developmental trajectory (Bailey et al., 2017).

The Sustaining Environments Explanation

The third explanation for fadeout of Pre-K treatment effects is the sustaining environments explanation, which asserts that the quality and supportiveness of a child's environment after program treatment is critical for the sustainment of treatment effects (Bailey et al. 2017; Duncan and Magnuson, 2013; Ramey & Ramey, 2006). This perspective differs from the skill building and FITD perspectives, which both posit that targeting the right types of skills at the right time will prepare children to take advantage of any form of subsequent environmental opportunity (Bailey et al., 2017). The sustaining environments perspective, however, posits that such targeted interventions will not be successful unless they are accompanied with subsequent environmental experiences that are high-quality and enriching. A common point made in favor of this perspective is that the Carolina Abecedarian treatment children, who went on to attend well-resourced and desegregated schools in Chapel Hill, North Carolina, experienced less fadeout of treatment effects than in Perry Preschool, where children went on to attend lower-resourced and segregated schools in Ypsilanti, Michigan (Bailey et al., 2017; Duncan & Magnuson, 2013).

In addition to anecdotal evidence comparing the subsequent environmental contexts of children in Abecedarian and Perry, there is some empirical evidence to support the sustaining environments perspective. One of the first studies to highlight the sustaining environments perspective comes from Currie and Thomas (2000) who analyzed national data from the National Educational Longitudinal Survey (NELS). The authors found that cognitive scores of black students who attended Head Start faded out more rapidly than white students who attended Head Start. The authors also found that black Head Start students went on to attend lower-quality K-12

schools than white Head Start students. Finally, when stratifying by measures of school quality (school mean reading assessment performance), which enables comparisons of differential fadeout based on race while controlling for school quality, the authors found that black and white Head Start attenders performed similarly. A more recent study by Zhai et al. (2012) conducted a similar analysis and found persistence of Head Start effects only for children who subsequently attended higher-quality schools. The key implication of these two studies is suggestive evidence that Head Start fadeout occurs, at least in part, due to inferior subsequent schooling experiences of poor children.

A recent study by Swain et al. (2016) used data from the Tennessee Pre-K randomized control trial to examine the extent to which the quality of early-elementary teachers moderated the persistence of Pre-K effects. Linking student-level data from the Pre-K evaluation with records of teacher observation scores from the state's formal evaluation program, the authors found small, positive interactions between teaching quality and state Pre-K attendance on some cognitive measures. Additionally, the authors find that the interaction effects are largest for children with limited English skills and who entered Pre-K scoring in the bottom portion of the assessment performance distribution.

Additional studies have addressed the role of classroom and school-level quality as moderators of Pre-K treatment effect persistence. Magnuson et al. (2007) found that the Pre-K advantage at school entry dissipated for children who attended small classes in early elementary school but not for children who attended larger classes. In other words, the Pre-K effect faded out in small class size settings because children who did not attend Pre-K "caught up" to their Pre-K-attending peers. The same study also found that the Pre-K advantage faded more rapidly and completely when students experienced classrooms with more time spent on reading

instruction in the early grades. Similar studies by Claessens et al. (2014) and Engel et al. (2014) focused on advanced versus basic content coverage in kindergarten and found that advanced content coverage was beneficial for both mathematics and reading achievement outcomes in kindergarten for all students, regardless of prior Pre-K attendance. A paradoxical implication of these studies by Claessen's et al. (2014), Engel et al. (2014), and Magnuson et al. (2007) is that higher-quality education in the early elementary grades actually facilitates the erasure of Pre-K treatment effects because the non-treated children in the sample classes are able to "catch up".

In sum, the subsequent environments perspective stresses the importance of the quality of subsequent educational contexts for the sustainment of Pre-K effects. A few studies, though not all (Jenkins et al., 2016), provide some empirical support for this perspective. There is some evidence that the structural factors (e.g. class size) and instructional factors (e.g. content coverage, teacher quality) in early elementary schools are important moderating factors of Pre-K treatment effect persistence. If interventions that are not targeted at the "right" developmental stage or not coupled with later supports (e.g., high quality instruction) at least in part explain why the Pre-K effect fades after Pre-K, attention to possible strategies that correct these problems may help to extend the effect. Arguably, these strategies have been embedded in reform efforts that together have been called the P-3 approach (Manship et al., 2016; Takanishi, 2016).

The P-3 Approach

The P-3 approach includes myriad strategies, but often focuses on aligning standards, curricula, and assessments across grades Pre-K through third, facilitating collaboration between teachers across this continuum (Abry et al., 2015), and the provision of kindergarten transition practices (Shulting et al., 2005; Little et al., 2016). There is evidence of each of the three

aforementioned fadeout explanations apparent in the P-3 approach. For example, skill building is represented by a focus on improving the quality of *both* the Pre-K experience as well as later schooling experiences. Foot-in-the-door is represented by efforts, for example, to screen and identify children with disabilities early so that appropriate supports can be tailored. Finally, sustaining environments is represented in efforts to align instruction so that early gains are built upon in later grades. Returning to the paradoxical findings from studies that related later school quality to Pre-K persistence, these P-3 strategies are intended to help ensure steady gains for all students by promoting individualized instruction that keeps children operating in their "zone of proximal development," which is the notion that learning takes place when instruction is focused near or just beyond the child's envelope of knowledge (Vgotsky, 1997).

Summary

There are a number of theoretical explanations for the persistent problem of Pre-K treatment effect fadeout with limited empirical support for each explanation. However, the focus on Pre-K fadeout and the explanations for its existence is novel and we need to know a lot more. The fact that there is at least some empirical support for each of the three perspectives suggests that all likely play a role in the persistence of Pre-K treatment effects. A challenge for researchers is to sort out the relative impacts or importance of each perspective. An example of researchers directly attempting to disentangle these effects comes from a recent evaluation of the Building Blocks Pre-K math curriculum. In this study, classrooms were randomized to either receive or not receive the curriculum. Then, the treatment classrooms were further randomized to determine if they would receive a treatment of subsequent supports in kindergarten (Bailey et al., 2017; Clements et al., 2012). Designs such as this will help to determine the relative role and importance of elements of Pre-K treatment in the sustainment of Pre-K effects.

Dimensions of Pre-K Program Quality

The purpose of this section of the literature review is to detail what is known about components of Pre-K program quality and contextualize school-based Pre-K as a potential new component that warrants inquiry. Thus far, the review of the literature has focused on the effects of different Pre-K programs on child outcomes, without regard to how these programs may vary in important ways based on their quality. Pre-K program quality has been classified in different ways in the literature (Pianta, 2016), but these classifications often include three key categories. The first is structural quality, which includes elements such as class size and teacher qualifications, and can be seen as providing the precursors necessary for enriched teacher-child interactions. Second, process quality features are more proximal to the learning process and often include observational measures of classroom practices, such as the warmth of teacher-child interactions. Third, the temporal dimension of Pre-K program quality is concerned with dosage of Pre-K treatment, whether it be one versus two years of Pre-K or the number of hours per day that a child attends Pre-K, for example.

My consideration of school-based Pre-K and co-location in this dissertation implicates two of these quality categories: structural and temporal. School-based Pre-K is an example of structural quality in that it relates to the physical setting of Pre-K programs and sets the stage for various high-quality processes, such as coordinated and aligned instruction and data sharing across grades. School-based Pre-K and co-location is also related to the temporal dimension of Pre-K quality in that it concerns the length of time spent in a single setting type. I now turn to provide a review of the evidence on dimensions of structural, process, and temporal Pre-K quality, respectively.

Structural Dimension

Class size and teacher-child ratio is one area of structural quality that has received a lot of attention from researchers and a general consensus has emerged. In a recent review of the literature, Pianta et al. (2016) argues that class sizes above 20 are generally associated with poorer child outcomes for children. However, there is no consensus about a precise, optimal class size or teacher-child ratio. For example, studies that have experimentally varied class size ratios in Pre-K did not identify differential student outcomes, though the range of ratios that were varied were all quite small (1:5-9) (Farran, 2017).

In terms of teacher degree and certification, there is little evidence of a link with improved child outcomes in Pre-K. A study by Mashburn et al. (2008) examined three credential measures and found no link with student outcomes. The authors included measures of the Pre-K teacher having a bachelor's degree, the field of the teacher's degree being in child development, and if the teacher's assistant had a Child Development Associated (CDA) credential. In another study, researchers aggregated data from seven different Pre-K effectiveness studies to examine the relationship between Pre-K teachers having a bachelor's degree and a range of child outcomes in Pre-K (Early et al., 2007). Again, the authors found no link between teacher degree and child outcomes.

In addition to the specific structural elements of class size and teacher degree, researchers have also examined global measures— summary ratings of program quality—of structural quality. The study by Mashburn et al. (2008) focused on two global measures: The National Institute for Early Education Research's (NIEER) quality benchmarks and the Early Childhood Environmental Rating Scale (ECERS-R). The NIEER benchmarks include a set of ten structural quality elements, such as a teacher-student ratio of 1:10 or better and requiring Pre-K programs

to use comprehensive early learning and development standards. Mashburn et al. (2008) find no link between the aggregate number of NIEER benchmarks met and child outcomes in Pre-K. The second global measure, the ECERS-R, includes measures of the Pre-K classroom environment, such as physical space, different play center activities available, and furnishings for relaxation and comfort. For this measure, Mashburn et al. (2008) find that the scale is largely unrelated to child outcomes except for oral and written language outcomes.

Process Dimension

While structural quality features of Pre-K programs are generally unrelated to child outcomes, there is more consistent evidence that process measures are related to child outcomes (Farran, 2017). One common measure of Pre-K classroom processes that has been examined is the Classroom Assessment Scoring System (CLASS) (Pianta et al., 2007). The CLASS provides a measures of the quality of two global domains of teacher-child interactions: instructional support and emotional support. There are seven different scales from the CLASS that comprise these two global domains. Instructional support includes scales that capture concept development, which are strategies used by teachers to promote higher-order thinking and problem solving, and quality of feedback, which capture the verbal evaluation provided to children about their work, comments, and ideas. The emotional support scale includes scales that capture whether or not the classroom climate it positive or negative, teacher sensitivity, control level the teacher has over classroom activities, and behavior management strategies.

Evidence on the relationship between process quality, as measured by the CLASS, and student outcomes in Pre-K is available from the National Center for Early Development and Learning's (NCEDL) Multi-State Study of Pre-K and the State-Wide Early Education Programs Study (SWEEP). The two studies used common measures and the results of the study are

published by Mashburn et al. (2008). The authors find no relationship between the quality of the emotional support component of the CLASS and student outcomes but they do find that the quality of instructional interactions is positively associated with multiple measures of academic and language development but not with measures of social-emotional outcomes.

Temporal Dimension

The third and final dimension of Pre-K quality that I examine in this section is related to the time or dosage of Pre-K treatment. More specifically, this dimension has been operationalized in two different ways in the literature. The first concerns whether or not one or two years of Pre-K (including combinations of Pre-K and Head Start) is more beneficial for child outcomes. The second concerns dosage in terms of the number of hours per day or week that a child attends Pre-K. In general, the evidence supports the claim that more time in Pre-K is associated with higher student outcomes, especially for low-income children (Berhman et al., 2004; Hill et al., 2003; Marcus-Jenkins, 2016).

Previous studies that examined one versus two years of Pre-K find that the additional year is beneficial for child outcomes, but the benefit is a fraction of the benefit of the first year (Reynolds et al., 2011; Tarullo et al., 2013). A recent study looked beyond one versus two years of a single Pre-K type to investigate the benefits of Head Start followed by state Pre-K or two years or Head Start (Marcus-Jenkins et al., 2016). The authors finds that Head Start followed by state Pre-K is associated with benefits for children's pre-reading skills but there is no differential effect for pre-writing or pre-math skills. Finally, there is some evidence that two years of Pre-K may have adverse effects on children's problem behaviors while still providing benefits on academic and cognitive outcomes (Loeb et al., 2007; Magnuson et al., 2007; Vandell et al., 2010).

In addition to the years of Pre-K a child experiences, the amount of hours per day or week is also associated child outcomes. In general, researchers have operationalized Pre-K attendance into part-time and full-time. Full-time Pre-K is commonly categorized as a child attending Pre-K at least 20 hours per week. Numerous studies have used this categorization and have found that full-time Pre-K is associated with higher child outcomes than part-time Pre-K (Loeb et al., 2007; Magnuson et al., 2004; Gormley & Gayer, 2005), though others have found no relation (Pianta et al., 2005). Again, studies of duration of Pre-K during the day have also found that, while providing benefits to a host of academic and cognitive outcomes, increased dosage is linked with increases in problem behaviors (Loeb et al., 2007). Interestingly, one study found that for children who attended school-based Pre-K, the adverse effects of Pre-K on problem behaviors was *not* observed despite being observed for all Pre-K attenders (Magnuson et al., 2007).

School-Based Pre-K Programs and Effects

In this final section of the literature review, I highlight research on school-based Pre-K programs. I begin by reviewing some of the early Pre-K model programs that included either elements of school-based Pre-K or follow-on services that extended into elementary school. Next, I detail some Head Start initiatives that sought to build stronger connections between Pre-K and elementary school, including a related program, Schools of the 21st Century. I conclude the section by reviewing the one study that has examined the effectiveness of school-based Pre-K versus non-school-based Pre-K that was not a part of a comprehensive program. In other words, the study examined differences in setting type alone, not differences in other program attributes, such as Pre-K follow-on services. In some cases, the evaluation evidence on the effects of the programs highlighted in this section is low-quality or nonexistent; however, I still note these

studies to demonstrate previous efforts to build connections between Pre-K and elementary school.

The Carolina Abecedarian Project

The Carolina Abecedarian Project is, as detailed earlier in this chapter, one of the best-known Pre-K experiments in the field of early childhood education (Yoshikawa et al., 2013). In addition to the randomization for the preschool treatment, at kindergarten entry, both the treatment and control groups were randomized into the Abecedarian K-2 Educational Support Program or a K-2 Control condition. This component of the Abecedarian is of particular interest for the current review as it concerns the connection between preschool and early-elementary school. The K-2 support treatment, which was randomly assigned to half of the original treatment and half of the original control particulars, was designed to sustain children's academic development by supporting parental involvement in education and building on the gains made in the preschool treatment. The K-2 support treatment included Home-School Resource Teachers, who worked with classroom teachers to ensure that children's early school experiences were developmentally appropriate and also with parents to explain ways they could enhance schoolwork by participating in specific educational activities at home (Campbell & Ramey, 1995).

Empirical evidence on the effectiveness of the Abecedarian program is extensive and was reviewed earlier. Fewer of the Abecedarian analyses have focused on the secondary K-2 Educational Support Program. While the effects of the preschool intervention far exceeded the K-2-only condition (Reynolds, Magnuson, & Ou, 2010), there is some evidence that the combination of the preschool intervention and the K-2 intervention produced *additive* benefits than just the preschool intervention. At age 15, Ramey and colleagues (2000) found that

preschool intervention and K-2 intervention group outperformed the preschool-only group on measures of reading achievement. Finally, while statistically insignificant, the trend of benefits for reading was present when the participants were 21 years old. Given the small sample size of the initial experiment and even smaller sample size of the preschool plus K-2 intervention group, low statistical power to detect program impacts is a concern.

Chicago Child-Parent Centers

The Chicago Child-Parent Centers (CPC), also reviewed previously, included a school-based Pre-K component (Reynolds, 2000). Important in the context of this review, there is evidence that children who attended the early intervention and *continued* with the program into early-elementary school demonstrated higher academic achievement than children who only participated in the early intervention (Conrad & Eash, 1983). In addition to benefits of the "add-on" –which included curriculum alignment and professional development for teachers in the primary grades–for academic achievement outcomes, other researchers have found positive effects on school remedial services and the frequency of delinquency infractions (Reynolds et al., 2001), as well as at the age of 24, higher rates of high school completion, full-time employment, and lower rates of Medicaid usage and violet arrests (Reynolds et al., 2007).

Head Start Programs

Throughout the history of the Head Start program, founded in 1965, there have been a number of initiatives and demonstration programs meant to better connect early education to elementary school and mitigate fadeout (Beatty, 1997). The first and most widely known of these programs is Head Start Follow-Through (HSFT) (Watkins, 1995). HSFT was initially developed to simply provide the Head Start program in an elementary school setting in order to ease the

transition between preschool and elementary school. However, as Kennedy (1993) details, the program was not implemented in this way. Instead, the program evolved into a large experiment of different *planned variation* approaches in the early grades, including parent education, direct instruction, behavioral analysis, High/Scope Curriculum, and the Bank Street Model of Developmental-Interaction.

Evaluation results indicate that the direct instruction and behavioral analysis models were most consistently associated with student gains in academic achievement and socio-emotional learning in early-elementary school. The High Scope and Bank Street models were also shown to be beneficial in terms of academic achievement, but the effects were transient (Seitz et al., 1983). It is also important to note that there are serious concerns about the quality of the evaluation of these planned variations in the HSFT program (House et al., 1978). For example, there were over 2000 comparisons in the study between HSFT and non-HSFT groups and the type of program and measures collected varied dramatically across sites.

The second Head Start-related program is Project Developmental Continuity (PDC). PDC was a demonstration program that was initiated in 13 Head Start sites in 1974. The aim of the program was to develop and implement comprehensive programs that link Head Start centers with local elementary schools in order to provide continuous developmental support through the third grade (Reynolds et al., 2010). The only evaluation evidence on PDC comes from the High/Scope Educational Research Foundation, which found that (1) the program model was not implemented as intended in any of the 13 sites, (2) the programs as implemented had few effects on parents and teachers, and (3) children's social competence was not enhanced by the program (Bond & Rosario, 1982).

The third and most recent Head Start-related program is the Public-School Early Childhood Transition Demonstration Project (PSTP). The PSTP began in 1991 and involved 12,000 children from 31 sites across the country. Participants were randomly assigned to a transition demonstration group or a business-as-usual control group. The transition treatment extended Head Start-like supports (e.g. home visits) through the first four years of elementary school. The program also included developmentally-appropriate curricula and instructional practices, social support services, and health and nutrition services (Ramey et al., 2000; Reynolds et al., 2010).

Evaluation results of the PSTP show no positive effect of the intervention (Ramey et al., 2004). That said, according to Ramey et al. (2004), the null results are likely a product of poor implementation. Only approximately 20 percent of the sites implemented high-quality programs and there was a large amount of spillover between the treatment and the control group, with schools in the control group adopting PSTP practices.

Schools of the 21st Century

The final school-based preschool program is the Schools of the 21st Century model (21C), which was developed by one of the key architects of Head Start, Edward Zigler (Finn-Stevenson & Zigler, 1999). The 21C model incorporates elements of model preschool programs, but targets children from the entire socioeconomic distribution, not just disadvantaged children. The program includes full-day preschool for children aged three through five, before and after school care, home visits for parents of children aged birth through three, and other social support services, which vary by site (Heinrich et al, 2006). The first 21C schools began operating in 1988 and there are now over 1000 21C schools in operation across the United States (Desimone et al., 2004).

Despite the number of 21C schools in operation, there is limited evaluation evidence on the effectiveness of them. In a national evaluation of the program, ten schools from five districts in five states (CO, CT, KY, MA, and MO) that had been using the 21C model for more than five years were analyzed. With a total sample of 860 children, the evaluation examined child achievement and socioemotional outcomes, parent behaviors, classroom instructional practices (through observations) and focus groups with teachers and parents. While the evaluation did not have a strong comparison group (i.e., it used existing national data as comparisons), it found that the 21C schools offered high-quality instruction in comparison to national averages and that children increased their math and reading skills from preschool to second grade at a higher rate than the national norm.

Another study of this program focused on the implementation of the 21C program within elementary schools, which is particularly salient for this review. The authors draw on focus group data with 20 preschool teachers, 22 kindergarten teachers, and 53 parents from 10 schools in five states (Desimone et al., 2004). The study revealed a number of key benefits as well as challenges associated with locating Pre-K programs within elementary school buildings. Key benefits highlighted by participants in Desimone et al. (2004) included opportunities for Pre-K and elementary school teachers to collaborate and coordinate instruction, efficient transitions from Pre-K to kindergarten, and stronger parental engagement throughout elementary school. The first two benefits provide support for my theoretically-driven hypothesis that remaining in the same school following Pre-K will be beneficial due to its facilitation of alignment.

Challenges revealed by Desimone et al., (2004) include conflicts over the use of material resources in the school, such as copier machines, as well as uncertainty over the extent to which Pre-K teachers were a part of the broader school community. For example, some Pre-K teachers

were uncomfortable speaking up in faculty meetings with the whole school and also felt invalidated by the lack of small gestures, such as having a mailbox in the administration office. Other challenges included building respect for Pre-K education and its pedagogical norms, salary inequities, and protecting preschool from accountability and testing pressures (Desimone et al., 2004).

Magnuson et al. (2007)

The study that is most related to my proposed dissertation is by Magnuson et al. (2007) and focuses on the child outcomes associated with school-based Pre-K. This study used the nationally-representative Early Childhood Longitudinal Study- Kindergarten Cohort of 1998 dataset to estimate the effects of Pre-K attendance on child outcomes and to compare these effects to other forms of preschool and Head Start. The analysis used a range of econometric techniques to derive causal estimates from observational data, including regression with saturated controls, instrumental variables, and propensity score matching. While the main analysis in this study focused on the effects of Pre-K without respect to school-based Pre-K, a sub-analysis within the paper examined the differences in Pre-K effects between children who attended Pre-K in the *same school that they attended kindergarten in* versus children who attended Pre-K in a *location that differed from where they attended kindergarten*. While this contrast is informative, it does not distinguish between children who attended Pre-K in *an elementary school but switched to another school for kindergarten* and children who attended Pre-K in a *non-school-based setting*.

With this limitation in mind, the sub-analysis by Magnuson et al. (2007) found two key findings. The first is that for all Pre-K attenders, regardless of setting, Pre-K attendance was associated with positive math and reading achievement outcomes in the fall of kindergarten. The

effect sizes on academic outcomes ranged from 0.18 to 0.19 standard deviations. The analysis did reveal differences for Pre-K co-location in terms of socio-emotional outcomes. In particular, Magnuson et al. (2007) examined the outcomes of self-control and externalizing problem behaviors. In their main analysis of all Pre-K attenders, regardless of setting, the authors reported negative effects on both self-control (-0.07 SD) and externalizing problem behaviors (-0.11 SD). However, in the analysis that included Pre-K co-locaters, the authors found no significant adverse effects for co-locaters.

The key implication of the study by Magnuson et al. (2007) for this dissertation is the need to examine a holistic range of measures, including academic and socio-emotional outcomes, when estimating the effect of Pre-K on child outcomes. While other studies have also revealed simultaneously positive effects of Pre-K on academic outcomes and negative effects on socio-emotional outcomes (Yoshikawa et al., 2013), this is the first study to indicate that locating Pre-K programs in public schools may provide equally beneficial academic effects to center-based Pre-K while also providing no adverse effects on socio-emotional outcomes. However, as mentioned earlier, the study by Magnuson et al. (2007) did not probe the issue of Pre-K location fully and I build upon the findings in this dissertation.

In particular, I build on the work of Magnuson et al. (2007) in four ways. First, I use the newer version of the ECLS-K dataset that provides nationally-representative information on children who attended Pre-K in the 2009-10 academic year as opposed to the 1997-98 academic year. Second, as detailed in the Methods Chapter below, I include a richer set of outcome variables that will provide a more granular examination of sub-constructs of socio-emotional outcomes, including executive functioning, approaches to learning, self-control, problem behaviors, and interpersonal skills. Third, I construct Pre-K attendance variables that provide

clearer contrasts of the various Pre-K setting types. In particular, I examine the following two contrasts: (1) school-based versus non-school-based Pre-K, regardless of whether children co-locate and (2) co-location versus no co-location versus non-school-based Pre-K. Fourth, this study uses novel statistical methods that have been developed since Magnuson et al. (2007) was published that provide stronger support for causal inferences from observational data. Specific methods, which I will detail in subsequent sections, include propensity score weighting methods using Entropy Balancing. Fifth, I combine the ECLS-K data with qualitative interview data in order to provide a more in-depth portrait of school-based Pre-K and co-location in practice.

CHAPTER 3: CONCEPTUAL FRAMEWORK

Introduction

I conceptualize school-based Pre-K along two key dimensions. The first dimension relates to the quality of the Pre-K program itself and the second relates to the alignment of the continuum from Pre-K through elementary school. In terms of the first dimension, I theorize that location of Pre-K programs within elementary schools will increase the quality of the Pre-K experience for children and thus yield more positive child outcomes than Pre-K programs located in stand-alone centers. The second dimension of my conceptual framework considers the Pre-K experience within the continuum of Pre-K through elementary school. Here, I theorize that co-location of Pre-K programs within elementary school buildings will foster greater alignment and coordination of services that leads to more persistent Pre-K effects for children in these settings than children who do not remain in the same physical building in elementary school. In this chapter, I detail each dimension of my conceptual framework and include a review of the theoretical and empirical literatures that support each dimension. I conclude by formally detailing my research questions and hypotheses.

Program Quality

The first dimension of the conceptual framework for this study contends that children's early learning experiences matter and that the higher the quality of these experiences, the more positive child outcomes will be. Developmental theories, such as Bronfenbrenner and Morris' (2006) bioecological model, posit that the quality of children's experiences provide the groundwork for development and that early experiences lay the foundation for later

development. Further, neuroscientific evidence suggests that early childhood is a particularly "sensitive period" wherein environmental influences play a highly consequential role in the developmental trajectories of humans (Shonkoff & Phillips, 2000). As a result, it follows that the quality of Pre-K programs that serve children during this critical period of development will matter for the magnitude and persistence of program impacts. Indeed, existing empirical research provides support for the theory that high-quality Pre-K programs are associated with larger and longer lasting effects on child outcomes (Burchinal et al., 2011; Belsky et al., 2007).

We also know that not all Pre-K programs are created equal in terms of quality. Indeed, the early experiments of Pre-K, including the Perry Preschool Project and the Carolina Abecedarian Study, were very intensive and expensive programs that do not mirror the reality of Pre-K programs offered at scale today (Reynolds et al., 2010). In Perry, certified teachers with a bachelor's degree taught children in class size ratios of 6:1. The program also included weekly home visiting and cost approximately $13,000 per pupil in 2016 dollars. Today, the average per pupil expenditure on Pre-K in state-run programs is approximately $5,000 (Barnett et al., 2017). Between states, there is also variability in the types of quality elements their respective programs provide. For example, the National Institute for Early Education Research rates state Pre-K programs annually on 10 research-based quality benchmarks (e.g., student-teacher ratio) and finds that only six states meet all ten standards and 14 only meet five or fewer standards. Finally, and directly relevant to the current study, there is also variability in quality *within state programs*—something that may be predicted by the *physical location of Pre-K settings*.

There is some evidence that Pre-K programs that operate in elementary school buildings have what some might call "higher quality elements" than center-based programs located elsewhere. A recent evaluation of North Carolina's Pre-K program, which is the foci of the

qualitative portion of this study, found that Pre-K programs in public school settings had higher proportions of teachers with advanced degrees and Birth-Kindergarten (BK) licensure (Peisner-Feinberg et al., 2016). Another study, by Bellm et al. (2002), examined differences between school-based and non-school-based programs across a number of state Pre-K programs and found that school-based programs had teachers with higher credentials, higher pay and benefits, and had lower levels of staff turnover. Together, this theoretical and empirical evidence supports a hypothesis that school-based Pre-K programs may be of higher quality and will thus yield larger and longer lasting impacts on student outcomes as compared to non-school-based programs.

Alignment

In terms of alignment across the early grades, developmental theories again stress the value of coordinated, coherent progressions as critical for student learning gains. For example, the aforementioned bioecological model also includes a temporal dimension (formally, the *chronosystem*) that implicates the role of the sustainment of various contextual influences on child development over time (Bronfenbrenner & Morris, 2006). This ecological model has been applied to the transition to kindergarten by Rimm-Kaufman and Pianta (2000) and the central element of the model is the continuity of contextual supports over time. They note, "Examination of this transition period must address how contexts and relationships change over time, and how change and stability in these relationships form key aspects of children's transitions to school" (p. 491). In this regard, school-based Pre-K and co-location can be seen as facilitating stability in contextual relationships across the Pre-K to elementary school transition period through stable educational settings and school norms and practices, for example.

In addition to bioecological theory, a complementary theoretical perspective comes from the work of Vygotsky. Specifically, Vygotsky's notion of the Zone of Proximal Development contends that learning takes place when children are provided content and asked to do procedures that are just beyond their current level of knowledge. If the content or procedures are too basic or too advanced, learning is not optimized. In the context of alignment along the Pre-K to elementary school continuum, this perspective suggests that instruction should be individualized and catered so that each child is regularly being challenged at or just above their current level of knowledge. School-based Pre-K and co-location may facilitate this through greater collaboration across grades and sharing data on children's educational needs across grades. While ecological theory is useful to support consideration of the complex contextual factors that support children's development, Vygotsky is helpful in explaining *how* and *under what conditions* learning takes place within these contexts.

Empirical research on the role of alignment across the Pre-K through elementary school continuum is sparse and not well defined. A set of studies use data from the Early Childhood Longitudinal Study to probe this topic. The first, by Engel et al. (2013), compared math achievement at kindergarten entry and teacher's reported math instruction during the kindergarten year. Despite children entering kindergarten with basic math skills, teachers report spending a significant amount of time teaching these same skills (13 days per month), and this basic math instruction is negatively associated with student math achievement at the end of kindergarten. Moreover, children benefitted the most when math instruction, as reported by the teacher, was advanced. Another study, by Abry et al. (2015), examined the alignment between teacher views on the importance of academic and social school readiness skills between Pre-K and kindergarten teachers. They find that when teachers have misaligned views on school

readiness competencies (i.e. Pre-K teachers value social skills while kindergarten teachers value academic skills, for example), children are rated as having poorer approaches to learning, social skills, and lower math achievement. School-based Pre-K may serve to foster more common goals in terms of school readiness and generate greater alignment in expectations across the Pre-K and kindergarten spheres.

There is also emerging evidence from the ongoing IES-funded Early Learning Network study in North Carolina that school-based Pre-K may facilitate greater alignment and continuity in the early grades. Through in-depth interviews with state and county officials as well as Pre-K and elementary school administrators, the study has revealed that in school-based Pre-K settings, respondents report increased alignment in two key areas: (1) transition from Pre-K to kindergarten and (2) data-driven decision making (Cohen-Vogel et al,, 2018).

In terms of the former, respondents highlight the ease by which children transition to kindergarten if they attended Pre-K in the same building. For example, one respondent noted that children become familiar with the school building, learn about the staff and school norms, and have the security and confidence to begin kindergarten in stride. The second key element that school-based Pre-K facilitates, as reported by the respondents, is data sharing and coordination across grades. In many cases, respondents noted that data on children's academic and social skills is shared with the child's kindergarten teacher and, in some cases, teachers meet in "vertical meetings" to discuss the needs of individual children as they progress from Pre-K to kindergarten. However, in some cases this sharing does not take place because children who attended Pre-K in a certain elementary school switch to another building to attend elementary school. This caveat suggests that the effect of school-based Pre-K may depend, in part, on co-location—something that directly test in this dissertation.

Bringing it All Together

School-Based Pre-K and Co-location as a Facilitator of Program Quality and Alignment

To summarize, I conceptualize school-based Pre-K and co-location as a facilitator of increased Pre-K program quality and a coordinated and aligned early learning experience from Pre-K through elementary school. School-based Pre-K programs are likely to be staffed by higher credentialed teachers who are paid more. Additionally, the dimension of alignment may shape the quality of the Pre-K program itself. Co-location may prompt Pre-K programs to align curricula and assessments, for example, to best prepare children for the demands of early-elementary school. Finally, beyond the effects on the Pre-K experience, co-location may enable greater coordination and alignment that facilitates individualized instruction that optimizes early learning. It is in the context of this conceptual framework that I formally detail my research questions and associated hypotheses.

Research Questions and Hypotheses

Research Question One

Do children who attend school-based Pre-K have different academic and social-emotional outcomes in elementary school than children who attend non-school-based Pre-K?

Based on the review of the literature indicating that school-based Pre-K programs have higher quality features (e.g., higher credentialed teachers) than non-school-based Pre-K and theoretical evidence that school-based Pre-K may be better aligned with elementary school instruction, I hypothesize that children in school-based Pre-K will have higher student outcomes in elementary school. This hypothesis applies to both academic and social-emotional outcomes.

Research Question Two

Do children who co-locate—those who remain in the same elementary school where they attended Pre-K—have different academic and social-emotional outcomes than children who switch to a different elementary school after Pre-K or attend non-school-based Pre-K?

I hypothesize here that children who co-locate for kindergarten or longer will have higher and more sustained Pre-K gains than children who do not. There is very limited empirical evidence to support this hypothesis, so this hypothesis is largely supported by theoretical arguments about the role of alignment and continuity in supporting learning.

Research Question Three

For both Research Questions One and Two, do the observed relationships vary across subgroups of interest? Specifically, do the results vary by socioeconomic status, Pre-K funding type, urbanicity, school quality, and measures of P-3 vertical alignment?

I hypothesize that, for both specifications of the independent variable, the effect estimates will be largest for socioeconomically-disadvantaged children. The literature does not provide consistent evidence for differences based on the other subgroups, so I pose no hypotheses for these subgroups. I will detail in the Methodology Chapter why each of these four additional subgroup measures were included in the analysis.

Research Question Four

In what ways are Pre-K programs different when located in elementary schools versus stand-alone centers and how might co-location foster alignment and continuity?

I hypothesize that facilitators of alignment in school-based Pre-K settings will include transitions to kindergarten and data use. I expect there to also be heterogeneity in the extent to which schools leverage co-location to optimize early learning settings. In other words, the physical setting of Pre-K will open the door to possibilities with respect to quality improvement

and alignment, but much will depend on the extent to which administrators and teachers leverage this opportunity.

Research Question Five

What are the areas of alignment and misalignment between the quantitative findings for Research Questions One, Two, and Three and the qualitative findings for Research Question Four?

I hypothesize that I will find general consensus in findings from the quantitative and qualitative portions of this study. I expect positive impact estimates from the quantitative portion and I expect to find clear evidence of the ways in which school-based Pre-K can generate these impacts in the qualitative portion of the study.

CHAPTER 4: METHODOLOGY

Study Design

Why Mixed Methods?

The goal of this study was to provide a comprehensive analysis of the effects of school-based versus non-school-based Pre-K on child outcomes in elementary school *as well as to* understand the potential reasons for the observed relationships. As such, a mixture of qualitative and quantitative methods was required. Mixed methods research is often termed the "third research paradigm" in addition to quantitative and qualitative paradigms. Mixed methods is rooted in a pragmatic epistemology that eschews paradigmatic "wars" in favor of empirical research that combines methods in order to answer practical questions about "what works" and "why" (Johnson & Onwuegbuzie, 2004).

An analysis of the strengths and weaknesses of quantitative and qualitative data and methods, respectively, as applied to a particular research question is a useful tool for exploring the utility of a mixed approach (Johnson & Onwuegbuzie, 2004). According to Turner (2003), the *fundamental principle of mixed research* is that researchers should collect data using different strategies and analytic approaches so that the resulting mixture is likely to yield complementary strengths and no overlapping weaknesses.

In the present study, the key strengths of the quantitative component include strong external validity through the use of nationally-representative data, inclusion of a holistic set of academic and social-emotional outcomes, and strong internal validity through the use of a quasi-

experimental design that enables estimates of the causal impacts of Pre-K attendance on student outcomes. The key weakness of the quantitative component is the lack of information about *why* observed relationships manifest themselves. In terms of the qualitative component, key strengths include the ability to explore mechanisms by which school-based Pre-K may or may not drive observed outcomes from the quantitative component. The key weaknesses of the qualitative portion are a lack of generalizability beyond North Carolina, small sample sizes, and inability to make causal claims. Taken together, the combination of both quantitative and qualitative components minimizes weaknesses and the strengths are complementary to one another. As a result, the mixed methods approach provides a more robust and holistic methodological approach as opposed to a monomethod approach. Having detailed the justification for a mixed methods approach, I now turn to detail the specific mixed methods design that I used in my study— the convergent parallel design.

The Convergent Parallel Design

According to Creswell and Clark (2011), the convergent parallel design is the most common mixed methods approach and is used when gathering different but complementary data, at the same time and on the same topic, in order to understand the phenomena at hand. In this design, the quantitative and qualitative data and analytic methods are given equal weighting in terms of importance. The quantitative and qualitative data and analysis take place independently but occur concurrently. Once each of the two components have been analyzed independently, the two components are merged together to analyze the convergence (or divergence) of research findings. I now detail the specific design as it was applied to my dissertation.

Design Components

In Figure 4.1 I present the Research Design Procedural Diagram for the study. There are four key steps of the study design and they began with two key modes of data collection—qualitative and quantitative. In later steps of the study, the qualitative and quantitative elements were merged to form synthesized results and inferences. Step one of this study was focused on data collection for the qualitative and quantitative components. The quantitative data are nationally-representative data on children from kindergarten through fifth grade from the Early Childhood Longitudinal Study, Kindergarten Class of 2010-11. The qualitative data come from in-depth, semi-structured interviews with state, district, school leader, and teacher interviews in North Carolina.

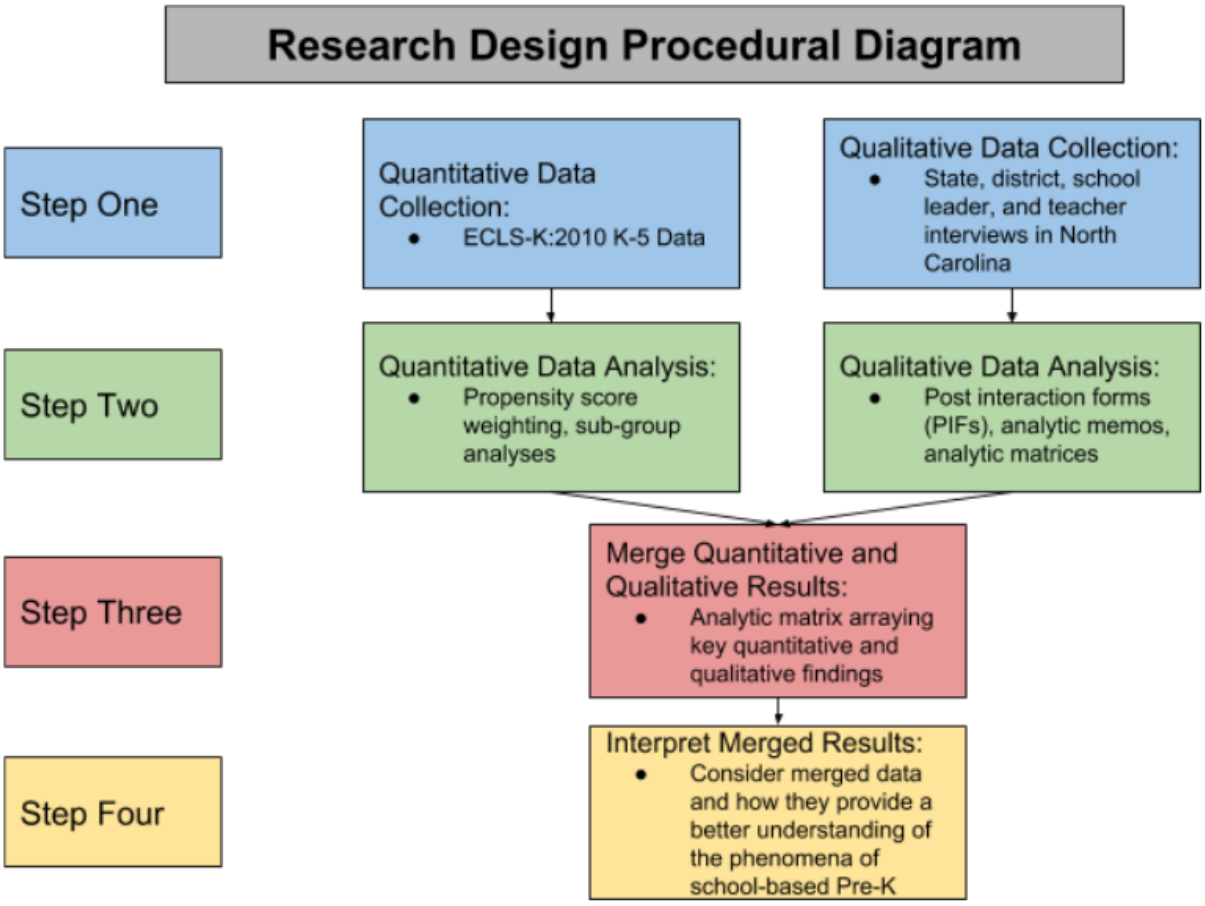

Figure 4.1. Research Design Procedural Diagram.

Step two of the study was focused on the analysis of the quantitative and qualitative components of the study, independently. The quantitative analysis used propensity score analysis to estimate the effects of school-based Pre-K on student outcomes in elementary school. The qualitative analysis probed the interview data to reveal patterns and themes in order to understand the realities of school-based versus non-school-based Pre-K settings.

Steps three and four of the study focused on the merging and analysis of the quantitative and qualitative data as one (Creswell & Clark, 2011). Step three was the merging process itself, where I used analytic matrices to array data sources in a digestible way to ease analysis. In step four, I considered the merged data to examine areas of convergence and divergence and the implications of these findings for conclusions about the phenomena of school-based Pre-K. Having detailed an overview of the study design and its component steps, I now turn to detail the analytic strategies used in each step of the study.

Quantitative Analysis

Data Source

The ECLS-K:2011 Dataset

The Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K:2011) is a nationally-representative and longitudinal study sponsored by the U.S. Department of Education's National Center for Education Statistics (Tourangeau et al., 2015). The ECLS-K:2011 provides data on child development, early learning, and school progress from a cohort of children who attended kindergarten in the 2010-11 academic year. The project gathered data from multiple sources, including children, parents, teachers, and school administrators and used a

range of data collection methods, including child assessments, surveys, and administrative data collection. Together, the dataset provides researchers with detailed, holistic information, including Pre-K experiences, about children as they progress through elementary school.

The baseline sample for the ECLS-K:2011 included approximately 18,000 kindergarteners from about 970 schools. Participants were selected for inclusion in the study using a complex survey sampling framework that took place in three phases. First, the United States was divided into primary sampling units (PSUs) and 90 of these PSUs were sampled. Second, public and private schools with kindergarten programs were sampled within each sampled PSU. Third, individual children within each sampled school were selected for participation in the study (Tourangeau et al., 2015).

Data collection took place regularly from kindergarten entry (Fall, 2010) through Spring, 2016, when most of the children were in the fifth grade. In this study, I used data from the kindergarten, first, second, and third grade data collection waves, which were all the waves available during the time of analysis. In total, there are five data collection waves that included assessments of the full sample of children that I include in my analysis. In addition to assessments at each wave of data collection, data were collected annually via surveys to parents, teachers, and school administrators.

Due to the complex nature of the ECLS-K:2011 sampling approach, weights were used to adjust estimates and standard errors (Tourangeau et al., 2015). The ECLS-K:2011 provides a number of sample weights to choose from and recommends selecting the weight that maximizes the number of sources of data included in the analysis. As such, I used sampling weight W7C17P_7T170, which covers all five child assessment waves of the full sample and includes all data sources, such as parent and teacher survey responses. In addition to using sample weights

to adjust point estimates, I also used replicate weights to adjust standard errors and correctly estimate variance of point estimates. Specifically, I used replicate weights with the jackknife method with two PSUs per stratum (JK2). This procedure estimates the point estimate of interest for the full sample as well as each of the 80 replicates. The variation of the replicate estimates around the full-sample estimate is used to estimate the variance for the full sample (Tourangeau et al., 2015).

Measures

Pre-K Setting Type

I generated two versions of the independent variable to test research questions one and two, respectively. The first specification is a dichotomous variable that designates whether a child attended school-based or non-school-based Pre-K. The second specification focuses on co-location and has three categories: (1) co-location (i.e. attended school-based Pre-K and stayed in same school for kindergarten), (2) school-based Pre-K only (i.e. moved to a different school for kindergarten, and (3) non-school-based Pre-K. In essence, the second specification subsets the school-based Pre-K attenders into two groups: co-locaters and movers.

The first specification of the independent variable was drawn from the ECLS-K:2011 parent survey that was administered in the Fall of 2010, when children entered kindergarten. The survey asked parents if their child attended center-based Pre-K on a regular basis in the year prior to entering kindergarten. I used this item to determine if a child attended Pre-K or not and all other children were removed from the dataset. An additional item asked parents about the physical location of the Pre-K program. Specifically, the item asked if the Pre-K was located in its own building, a public elementary school, a private elementary school, a college or university, or a church, synagogue or other place of worship. I used this item to sort children into either

school-based Pre-K or non-school-based Pre-K, where school-based is defined as a public or private elementary school.

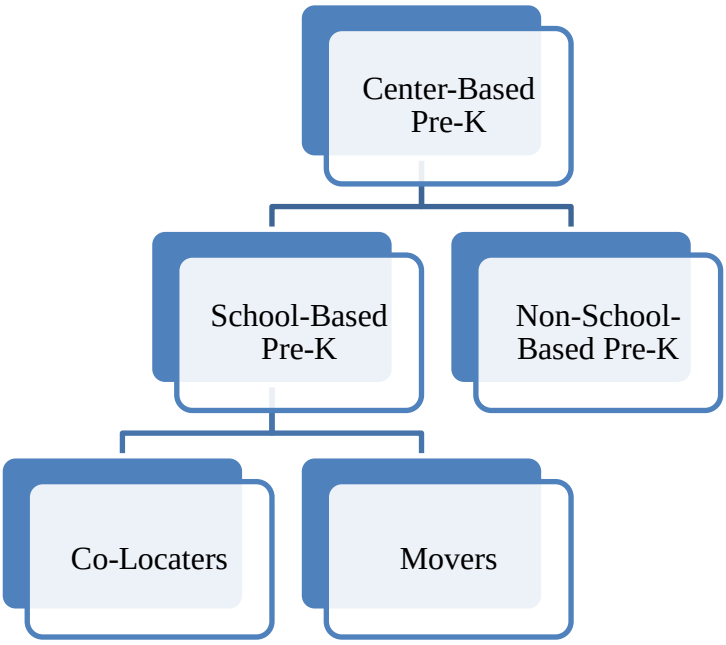

Figure 4.2. Independent Variable Structure.

The second specification of the independent variable disaggregates the school-based Pre-K attenders into co-locaters and movers. An additional item in the parent survey asked parents if the Pre-K program that their child attended was located in the school where their child now attends kindergarten. I used this item to distinguish between children who remain in the same school building for kindergarten and those that switch to a different elementary school, despite having attended a school-based Pre-K. The third category of this variable was be the same comparison from the first specification: non-school-based Pre-K attenders. In Figure 4.2 I present a hierarchical representation of Pre-K attendance in the ECLS-K sample.

Outcomes

Assessment Outcomes. The first set of outcomes that I examined come from the direct child assessments that were conducted in the Fall and Spring of kindergarten and the Spring of

first through third grades. Child assessments for the ECLS-K were conducted by trained and certified data collectors and took place during the school day. Trainings for child assessors lasted four days and certification was awarded based on successful completion of written exercises about the child assessment as well as an observation-based exercise administering the assessment to a kindergarten-aged child who was brought on-site to the training session. In each wave of data collection, the assessment battery was designed to be administered within approximately 60 minutes for each child. The mathematics and reading achievement assessments took the most time (approximately 50 minutes) and the executive function assessments took approximately 10 minutes. Child responses to assessment measures were recorded digitally using a computer-assisted interviewing program.

Mathematics. The ECLS-K:2011 mathematics assessment was designed to measure children's skills in conceptual knowledge, procedural knowledge, and problem solving. The assessment included questions on number sense, properties, and operations; measurement; geometry and spatial sense; data analysis, statistics, and probability; and patterns, algebra, and functions. The mathematics assessment used a two-stage design where all students were given a common set of 20 routing items and then scores on the routing items determined the difficulty level of the remaining assessment items (low, middle, or high). Assessors used easels with verbal prompts to administer items so children's reading ability would not affect their mathematics assessment performance. Additionally, children were offered paper and pencil to use during the assessment. For some items that involved counting, wooden cubes were provided as an aid to children.

For analysis, I used Item Response Theory (IRT)-based scores developed for the ECLS-K:2011. IRT is a method that generates domain-scores for each child that can be compared with

other children in the sample, regardless of the specific level of questions that they received. Other benefits of IRT-based scores over a raw number-right score are that (1) these scores can adjust for the possibility of a low-ability child guessing several difficult items correctly, (2) they can model the probability that a child would have been correct or incorrect on an omitted item, and (3) IRT scoring is ideal for analyses of longitudinal gains in achievement. Reliability of the mathematics IRT assessment scores, theta, was high across data collection waves, ranging from 0.95 in Fall of kindergarten to 0.91 in Spring of second grade.

Reading. The ECLS-K:2011 reading assessment included items about basic language and literacy skills, such as print familiarity, letter recognition, beginning and ending sounds, rhyming words, and word recognition, as well as vocabulary knowledge and reading comprehension. The reading comprehension items asked children to identify information stated in text (e.g., definitions and facts) and to make complex inferences within texts. All children were administered a common set of 29 routing items and then scores on these items determined the level of difficulty of the remaining items (low, middle, or high). Like the mathematics assessment scores, I used the IRT-based scale scores for reading achievement. Reliability of the reading IRT assessment scores, theta, esd high across data collection waves, ranging from 0.94 in Spring of kindergarten to 0.92 in Fall or kindergarten.

Executive Functioning. The ECLS-K:2011 included measures of two domains of executive functioning in the direct cognitive assessment: working memory and cognitive flexibility. Working memory was measured with the Numbers Reversed subtest of the *Woodcock-Johnson III Tests of Cognitive Abilities* (Woodcock et al. 2001) and cognitive flexibility was measured with the *Dimensional Change Card Sort* (DCCS) (Zelazo 2006). In the Numbers Reversed task, children were provided with a digit span from the assessor that they

were then expected to repeat in reverse order. For example, if the assessor said "3...8," then the child should respond "8...3." The task begins with five two-number sequences. If the child gets three consecutive two-number sequences incorrect, then the task ends. If not, the sequence becomes increasingly long, up until an eight-number sequence or the child gets three consecutive spans incorrect. I used the *W* score developed for the ECLS-K:2011 in accordance with developer guidelines for analysis. The *W* score is a special transformation of the Rasch ability scale and provides a common scale of equal intervals that represents a child's ability and the difficulty of the task (Tourangeau et al. 2015).

In the DCCS, children were asked to sort a series of picture cards based upon different sorting criteria. Each card featured either a red rabbit or a blue boat. In the first portion of the task, children were asked to place the cards into trays based on the *color* of the trays, regardless of the picture. In the second portion, children were asked to place cards into trays based on the *shape* on the trays, regardless its color. Based upon performance on the second portion, children may proceed to a third portion of the task where the sorting rule depends on whether or not the card has a black border around the edges. If the card has a border, they are to sort by color, and if there is no border, they are supposed to sort by shape. I used the combined score that reflects children's performance on all three portions of the task and was recommended by the developer to assess general performance (Tourangeau et al., 2015).

Teacher-Reported Outcomes. In addition to the three outcomes measures in the direct cognitive assessment (mathematics, reading, and executive functioning), I also included a set of teacher-reported outcomes of social-emotional skills. Teachers completed items from the Social Skills Rating System (SSRS) (NCS Pearson, 1990) that yielded four social skill scales: (1) self control, (2) interpersonal skills, (3) externalizing problem behaviors, (4) and internalizing

problem behaviors. The score on each of these four scales is the mean rating on the items included in the scale, where higher scores represent that the child exhibits the behavior of interest more often. Consequently, higher scores for self control and interpersonal skills are "desirable" while lower scores on the externalizing and internalizing problem behaviors scales are "desirable".

Teachers also completed items from the Approaches to Learning scale, which was developed specifically for the ECLS-K and was completed in conjunction with the SSRS items. This scale asked teachers to report how often students exhibited the following learning behaviors: keeps belongings organized, shows eagerness to learn new things, works independently, easily adapts to changes in routine, persists in completing tasks, pays attention well, and follows classroom rules. The score for this scale is the mean rating of the seven items included in the scale, where higher scores indicate that the child exhibited positive learning behaviors more often.

Covariates

As I will detail in the Analysis section, I used propensity score methods to conduct my analysis. A key to successful implementation of propensity score methods is effective use of covariate measures. Specifically, the goal is to select a vector of covariate measures that represent confounders, which are measures that are simultaneously related to both the independent variable and the outcomes, as well as measures that are highly related to the outcome measures. A key benefit of the ECLS-K:2011 is the unusually rich set of covariates that I can draw upon to implement this analysis. I included the following covariate measures: race, gender, language spoken at home, number of siblings in household, family food security, parental marital status, age of child at kindergarten entry, parent age, child birthweight in ounces,

maternal WIC receipt, maternal relationship status at child birth, urbanicity, region, and the poverty rate in the school district of the child. Additionally, I measured socioeconomic status using a measure developed especially for the ECLS-K:2011. The socioeconomic status measure is a composite that includes the occupational prestige of a child's parent(s), their income, and educational attainment. See Table 5.1 for descriptive information about the measures used in the analysis.

Analysis

Causal Inference and Observational Studies

Randomized control trials (RCTs) are frequently heralded as the "gold standard" for estimating causal effects because randomization, when implemented correctly, ensures that treatment and control groups, in expectation, are the same on all measures except for the treatment (Bloom, 2006). In addition to eliminating bias from confounding variables, RCTs are also lauded for the simplicity of analysis—researchers only need to compare the difference in means between the treatment and control groups to estimate a causal effect estimate. However, it is often the case that RCTs are not feasible, whether for economic, ethical, or other practical concerns (Cochran, 1965). As a result, researchers often work to derive causal estimates from observational data.

The key challenge when estimating causal effects from observational data is the possibility of confounding variables. For example, if children are not randomly assigned to either school-based or non-school-based Pre-K, it is possible that there are factors beyond Pre-K setting type that influence their outcomes. In the following section, I formalize this challenge in the

context of the Rubin Causal Model, commonly known as the potential outcomes framework (Holland, 1986; Rubin, 1974).

Potential Outcomes Framework. When applying Rubin's (1974) potential outcomes framework to the first research question of the present study, there are two possible treatments: school-based Pre-K (active treatment) and non-school-based Pre-K (control treatment). (Note that my second research question will have three treatment groups, but I only use the first here for an illustrative example). For each child in the sample who attended Pre-K, there are two potential outcomes: $Y_i(1)$, which is the outcome of child i if s/he attended school-based Pre-K and $Y_i(0)$, which is the outcome if the same child i attended non-school-based Pre-K. In reality, we never observe a child with both treatments—we observe one outcome and the unobserved outcome is the counterfactual. The appeal of an RCT is that, through the process of randomization, we can assume that:

$$E[Y_i(1) - Y_i(0)] = E[Y(1) - Y(0)]$$

where $E[Y_i(1) - Y_i(0)]$ is the actual, but in reality, unobservable Average Treatment Effect (ATE) and $E[Y(1) - Y(0)]$ is the estimated ATE. We are able to assume this equivalence because randomization ensures that the treatment and control groups are equal in expectation before the treatment is assigned. As a result, we can be confident, assuming randomization and SUTVA assumptions hold, that any observed differences between the two groups are due to the introduction of the treatment.

In the absence of random assignment, assignment of treatment to either school-based or non-school-based Pre-K could be non-random and influenced by confounding variables. In observational studies, researchers must work to address all of the confounding variables in order

to recover unbiased causal estimates. One such method for recovering causal estimates from observational data is propensity score methods.

Propensity Score Analysis

Background. Propensity score methods seek to address the challenge of the missing counterfactual posed in Rubin's Causal Model (RCM) by using observable measures to model selection into either treatment or control conditions. In the context of the present study, this method seeks to use measures to predict what determines whether children attend school-based or non-school-based Pre-K and in the second phase of the analysis, co-location. Using the logic of the RCM, if I can effectively model selection into treatment conditions and achieve balance on covariates, casual estimates in the absence of random assignment are attainable.

Propensity score methods were first detailed by Rosenbaum and Rubin (1983). The key element in propensity score methods is the propensity score itself, which is the probability of treatment conditional on observed covariates: $e_i = \Pr(Z_i = 1) \mid X_i$ (Austin, 2011). A key assumption when modeling treatment assignment is that of "strong ignorability" or "no unmeasured confounders," which posits that treatment assignment is independent of the potential outcomes given the observed covariates. Formally, this assumption can be written as $(Y(1), Y(0) \perp Z \mid X)$, where potential outcomes $Y(1), Y(0)$ are independent of treatment assignment, Z, conditional on the covariates, X. If this assumption is met, the propensity score can be used to achieve balance on all confounders between the treatment and control groups and then causal estimates can be estimated from the two groups.

Propensity Score Tautology and the Primacy of Balance. In reality, we never know the true propensity score, so researchers must rely upon their estimated propensity score as a substitute—a phenomena Ho et al. (2007) term the "propensity score tautology." In order to

attain the true propensity score, researchers must precisely model selection into treatment, including the use of all proper variables and functional forms—a nearly impossible task. However, since the propensity score is ultimately a tool to achieve balance, we can feel confident in the use of our propensity score estimate as long as we are able to achieve balance (Ho et al., 2007). As a result, the guiding principle that I adopted in selecting propensity score estimation techniques is to select the technique that yielded the most balance between treatment and control groups on observable characteristics.

Propensity Score Estimation. The traditional way to estimate a propensity score is to estimate the independent variable as a function of observable covariates using a logistic or probit estimation model, generate the predicted probabilities, and then use those predicted probabilities as controls, weights, or to conduct some matching procedure. These traditional approaches to propensity score analysis continue to be used today and, assuming a number of assumptions are met, can yield valid results. However, there are more novel methods that have been developed in recent years that regularly ensure high levels of covariate balance and require fewer modeling assumptions on the part of the analyst. In this study, I used one of these novel methods—Entropy Balancing.

Entropy Balancing, developed by Hainmueller et al. (2012), does not estimate a propensity score model but rather directly estimates weights by minimizing the "entropy" of weights. This method, using the Method of Moments estimator, prioritizes covariate balance as the goal to optimize on. When estimating Entropy Weights, analysts can choose to balance on mean differences in covariates as well as higher-order moment conditions, including variance and skewness. I estimate weights in this study that attain balance on all three moment conditions. Essentially, this procedure determines what weights are necessary to achieve balance on each of

these moment conditions. A potential pitfall of such an approach is the generation of extreme weights. To avoid the impact of extreme weights, I truncated the weights by recoding weights outside of the first and 99th percentiles to those percentile values. I also visually inspected the distribution of the weights and the distribution was normal.

Missing Data

It is necessary to address issues of missing data in the ECLS-K:2011 dataset in order to minimize bias, maximize use of available information, and obtain appropriate estimates of uncertainty (Allison, 2002; Little & Rubin, 2002; Rubin, 1976). While rates of missingness on key outcome variables are very low (<0.5 percent), rates of missingness on covariates are as high as 25 percent. When analyzing complete case data, the implications on the sample size can be dramatic. When the data are missing completely at random (MCAR), the estimates generated from listwise deletion will be unbiased but the standard errors can be inflated dues to drops in sample size and statistical power (Allison, 2002). When data are missing not at random (MNAR), the likeliest scenario, there is a risk for both bias and standard error inflation (Allison, 2002). I used multiple imputation methods in order to address missing data and the threats that it has for my analysis.

Multiple imputation is an iterative form of stochastic imputation that generates a vector of plausible values to replace missing data cells. In other words, rather than using a model to generate a single estimated value for a missing data cell, multiple imputation generates multiple estimated values that reflect the uncertainty around the true value (Johnson & Young, 2011; White et al., 2010). Multiple imputation has three basic steps. First, is the imputation phase where the missing data cells are filled with estimate values. This process is done for each imputation dataset, *m*. Second, analysis is conducted using each of the *m* completed datasets.

Third, the parameter estimates from the analysis on each of *m* datasets are combined for inference in a way that accounts for the uncertainty associated with estimated missing values.

To conduct multiple imputation, I used the MI Impute command set in Stata (StataCorp, 2015). My imputation model included all measures used in my analysis. I imputed using the chained equations methodology, which allows for regression for continuous variables, logistic regression for binary variables, ordered logistic regression for ordered variables, and multinomial logistic regression for nominal variables. I estimated 30 imputed datasets. I then conducted analysis over all of the 30 imputed datasets and combined the parameter estimates that appropriately reflect the uncertainty associated with imputed values. I diagnosed the success of imputation using the standard information provided in the MI command set, including Relative Increase in Variance, Fraction of Missing Information, Degrees of Freedom, Relative Efficiency, and the between and within imputation variance estimates (Ender, 2010; Rubin, 1987). The imputation process was successful in each of these diagnostic tests.

Analytic Procedures

Descriptive Analysis. After I addressed missing data via multiple imputation procedures, I began the analytic process with a descriptive analysis. I calculated basic descriptive statistics, including mean, standard deviation, minimum, and maximum, for each measure used in the analysis.

Propensity Score Estimation and Evaluation of Balance. Having run descriptive statistics on all relevant measures in the analysis, I then estimated propensity score weights using Entropy Balancing. I implemented the balancing procedure using the *ebalance* command in Stata 14. In the first specification of the estimation, I included all control variables and specified that balance should be achieved on all three moment conditions (mean, variance, and skewness). In

the second specification, I added a measure of the travel time between a child's home and the nearest school-based Pre-K in an attempt to capture exogenous prediction of children attending school-based Pre-K. To derive this measure, I used GIS software to calculate the travel time in minutes between a child's neighborhood census tract code and the nearest possible school-based Pre-K. Here, the theory is that children living closer to a school-based Pre-K program will be more likely to attend that program than a child that lives further away.

A summary of the balance results from these two procedures is presented in Appendix A. In both procedures, near perfect balance was achieved on the mean, variance, and skewness of each variable. As mentioned previously, I examined the distribution of the weights to ensure general normality of dispersion and also truncated the weights by recoding weights outside of the first and 99th percentiles to those percentile values.

Model Building and Estimation of Main Treatment Effects. Having estimated the propensity score weights using Entropy Balancing, I proceeded to estimate the final analytic models. I estimated two general models, both weighted by the product of the ECLS-K:2010 sampling weight and the estimated propensity score weight. The first model corresponds to the first specification of the independent variable and takes the following form:

$$Y_{ist} = \alpha + \beta X_{School\ Pre-K} + \varepsilon$$

Where Y_{ist} is the outcome of student *i* in subject *s* at time *t*. $X_{School\ Pre-K}$ is an indicator variable for children who attended school-based Pre-K and the reference category is students who attended non-school-based Pre-K. The second model, which corresponds to the second specification of the independent variable, takes the following general form:

$$Y_{ist} = \alpha + \beta X_{Co-locate} + \beta X_{Mover} + \varepsilon$$

where $X_{Co-locate}$ is an indicator for co-locaters, X_{Mover} is an indicator for movers and the reference category is non-school-based Pre-K attenders.

To begin, I ran five different variations of the general model in order to ensure I selected the most appropriate model for final analysis and interpretation. The first model was a standard OLS model that included the full vector of control variables. The second and third models added LEA and State fixed effects, respectively. The fourth model added the propensity score weight that was derived using only the control variables. The fifth and final model included the propensity score weight that was derived using the control variables and the GIS-calculated measure of travel time to the nearest school-based Pre-K. A summary of the results of these models using independent variable one is provided in Table 5.2. The model that I report throughout the analysis is the fifth model that includes the GIS weights and state fixed effects, along with the full vector of covariate measures.

Subgroup Analysis. In addition to estimating the main treatment effects and in order to answer research question three, focused on subgroup effects, I ran a series of additional models that included full interactions for subgroups of interest. The first model restricted the sample to include children whose families were from the bottom 25th percentile of the SES measure. The second model restricted the sample to include children who attended state-funded Pre-K (as opposed to any center-based Pre-K in the full sample). The third model restricted the sample to include children living in urban locations. The fourth model restricted the sample to include schools that met Annual Yearly Progress in the previous academic year—a proxy for school quality. Finally, the fifth model restricted the sample to schools that engaged in numerous Pre-K-kindergarten transition practices (>5, the 75th percentile)—a proxy for the extent to which an elementary school engages in vertical alignment.

Qualitative Analysis

The qualitative portion of this dissertation draws upon interview data with Pre-K stakeholders, ranging from state-level administrators to elementary school teachers, in central North Carolina. I used a cross-case, cross-sectional design within a single state policy context. While focusing on a single state policy context limits generalizability to other states, it enables me to compare across cases without the potential confounding of differences in state Pre-K program policies. As I will detail in the sampling section, I collected data and analyzed it across individuals within schools, schools within districts, and districts within North Carolina.

Sample

Figure 4.2 provides and overview of the sample. This sample is hierarchical, spanning state-level actors and individual teachers. At the state level, I interviewed 6 individuals at the North Carolina Department of Public Instruction (DPI) and the North Carolina Department of Health and Human Services (DHHS). DHHS has authority over NC Pre-K and DPI has authority over K-12 education. Since my study focuses on Pre-K and longitudinal effects through elementary school, it was necessary to interview representatives from both state agencies. Due to my work on the IES-funded Early Learning Network, I have been in contact and in some cases interviewed these respondents previously. This connection helped facilitate my access to these high-level policy actors (Marshall & Rossman, 2015).

I focused my interviews within three districts/counties in central North Carolina that vary by enrollment size (small, mid-size, and large). A summary of descriptive statistics for each of the three districts is provided in Table 4.1 below. The three districts varied by their enrollment in both K-12 public education and in terms of the number of NC Pre-K students served. The

districts also varied in terms of the percentage of minority students they serve. Pike serves the highest percentage of minority students (50%) and Olin the smallest (25%).

Table 4.1. Description of Qualitative Sample Counties.

District	Size Classification	Enrollment	NC Pre-K Students Served	% Minority
Olin	Small	7,500	100	25%
Case	Medium	12,000	200	30%
Pike	Large	24,000	600	50%

Note. Data were drawn from the Common Core of Data and the Early Childhood Integrated Data System. District names are pseudonyms. Values are rounded to ensure the confidentiality of participating districts.

I chose to vary the size of districts because of the potential role of size in enrollment patterns of students as they transition from Pre-K to kindergarten, with larger districts having more options for children to switch to different elementary schools, for example. At the district-level, I interviewed county-level actors in the Local Education Authority (LEA) as well as the county Smart Start Partnership. In North Carolina, NC Pre-K is administered out of county-level Smart Partnerships and in each partnership there is a county-level NC Pre-K program coordinator that I interviewed. Additionally, LEAs have authority over the NC Pre-K programs that are located within their school buildings. Each LEA also has a Pre-K coordinator that I interviewed.

Within each district, I sampled four different schools (two elementary schools with Pre-K programs, one elementary school without a Pre-K program, and one Pre-K center that is not associated with an elementary school building). I chose this selection of schools in each of the three districts for two principal reasons. First, it enabled me to draw comparisons across Pre-K

programs that are and are not located in elementary school buildings. Second, it enabled me to draw comparisons across elementary schools with and without Pre-K programs.

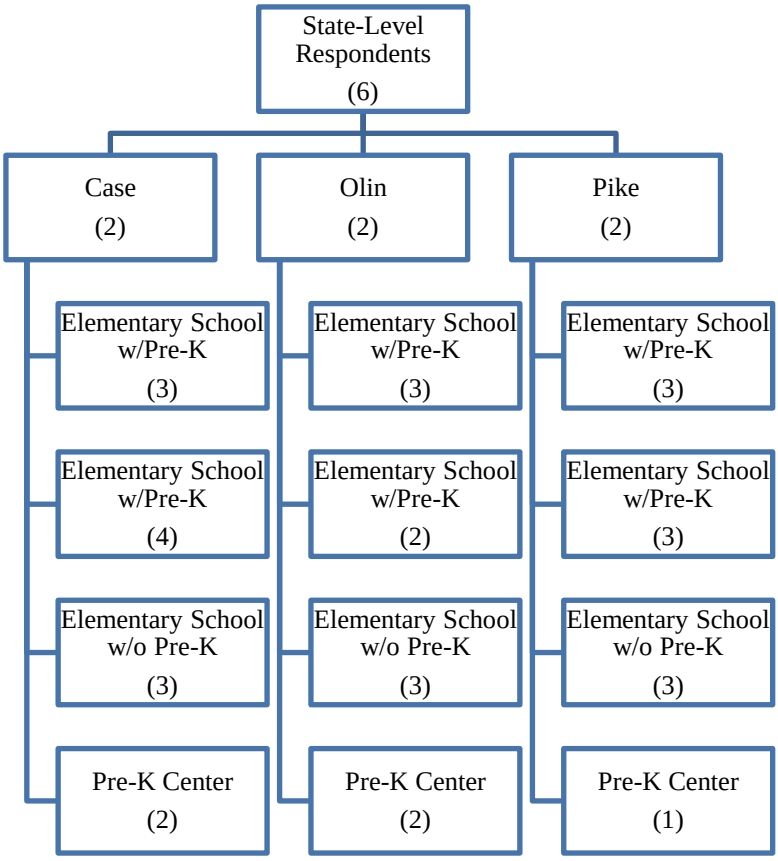

Figure 4.3 Qualitative Sample

Within non-school-based Pre-K programs, I interviewed one Pre-K teacher and the center director (Note: In Pike, only the Pre-K teacher agreed to participate). In elementary schools, I interviewed at least one Pre-K teacher and one kindergarten teacher. Since the Pre-K to kindergarten transition is the most stark transition year (Pianta et al., 2000), it is valuable to have perspectives from teachers on either side of that transition. I also interviewed the elementary school principal, where possible, in order to understand school-level processes related to school-based Pre-K and transitions. In total, I interviewed 44 individuals for the qualitative portion of the study. An overview of my qualitative sample is provided in Figure 4.3.

Data Collection Procedures

I used semi-structured interviews with study participants that lasted approximately 30 minutes each. I used interview protocols to guide the general structure and flow of each interview, but also allowed for minor deviations and follow-up questions to probe emergent ideas or discrepant responses, for example (Miles & Huberman, 1994). I developed protocols for state, district, school/center administrator, and teacher participants (see Appendix C). Interviews at the state and district/county levels focused primarily on motivations for locating Pre-K programs in elementary school buildings or stand-alone centers. Interviews with elementary school principals and center directors focused on the benefits and challenges of Pre-K setting types. Finally, interviews with teachers focused on the practical, day-to-day realities associated with either school-based or non-school based Pre-K settings, with a particular focus on connections across grades. An example of a central guiding question for all interviews was, for example, "What are some of the key benefits and challenges of (not)having a Pre-K program located within your school building?"

Directly following each interview, I completed a Post-Interaction Form (PIFs), which is a tool to organize initial findings and reflections on an interview immediately following its conclusion (Cohen-Vogel & Harrison, 2013; Miles & Huberman, 1994). The PIF, shown in Appendix B, was structured in two primary parts. In the first part, I reflected on the interview process and protocol itself. Here, I considered if any specific interview items could be improved upon, removed, or expanded in scope. The second part of the PIF was intended to spur reflection on emergent themes and findings from the single interview. This part of the PIF provided a way to document any notes about non-verbal communications, note areas for follow-up inquiry, and

brainstorm areas of emerging themes and patterns in the data. As I will detail in the next section, these forms were coded and analyzed alongside the transcribed interview text.

Data Analysis

After transcribing the audio-recorded interviews and compiling all PIFs and other supplementary artifacts, I coded the data. I used pattern coding to reveal central constructs in the data (Miles & Huberman, 1994). I used a hierarchical coding scheme that included primary, secondary, and tertiary codes, where applicable. The primary codes captured basic descriptions of the element of the interview under consideration, such as curricula or data systems. The secondary codes provided more detail on the primary codes. For example, secondary codes for data systems provided categories of data systems, such as assessment data or attendance/enrollment data. Finally, tertiary codes further clarified the content of a secondary code, where the secondary code alone is insufficient to sort the data. For example, a tertiary code for assessment data included entry/diagnostic assessments, formative assessments, and end-of-year assessments. In addition to descriptive codes, I applied codes that indicated phenomena that participants identified are challenges or facilitators in terms of providing high-quality, aligned early education, where applicable.

Once all data were coded, I analyzed the data using a variety of matrices. Essentially, the data were sorted into a large table where each participant was listed in a row and each column was a different code. I used this table to analyze the data hierarchically, from classroom to state. For example, I first examined data from participants within a single school. I then developed an analytic memo for each school that arrayed the key findings as well as areas of agreement and disagreement among participants. Once I did this for each participating school, I then analyzed these memos across schools within districts/counties. In this memo, I highlighted consistent and

inconsistent findings across schools within each district/county. Finally, I conducted the same analysis across districts within North Carolina. This analytic procedure enabled me to examine variation in findings across units within levels of analysis and better identify areas of general consensus or areas with discrepant findings.

Merged Data Analysis

Simply collecting and analyzing a combination of qualitative and quantitative data independently does not constitute a concurrent mixed-methods design (Creswell & Tashakkori, 2007). The third and final step in the analysis of this study was to consider how the previous two components, the quantitative and qualitative analyses, relate to one another and inform broader conclusions about the phenomena of school-based Pre-K. Specifically, I used an integrative joint display table that brings together data through visual means to draw out new insights that would not be available by considering the quantitative and qualitative findings in isolation (Creswell & Clark, 2006). This table is a tool to array the qualitative and quantitative data in a format that enables me to examine the two sources as one. According to Fetters et al. (2013), there are three different outcomes that joint displays generate: (1) confirmation, (2) expansion, and (3) discordance. Confirmation occurs when the findings from both the quantitative results and qualitative results confirm one another. Expansion occurs when the findings diverge and expand insights of the phenomenon by highlighting different aspects of the phenomena. Here, the findings are not in disagreement with one another, rather, they are highlighting different elements of a single phenomenon. Finally, discordance occurs when the findings are indeed in disagreement, contrast, and inconsistent. I used the joint display table to reveal how the findings from the quantitative and qualitative elements of the study revealed any of these outcomes. An

example template of a joint display table is shown in Table 4.1. In practice, my use of this table was iterative and informal, often constructed by arranging notecards in a grid-like fashion.

Table 4.2. Joint Display Table

Key Quantitative Findings	Qualitative Findings Intersecting Key Quantitative Findings				
Key Quant. Finding 1					
Key Quant. Finding 2.					
Key Quant Finding 3					
….					
Key Quant Finding N.					

CHAPTER 5: RESULTS

Quantitative Analysis Results

Descriptive Results

Descriptive statistics are provided in Table 5.1. Note that all quantitative results tables are provided at the end of this section due to their length. Based on this national sample of kindergarteners in 2010-11, 33 percent of center-based Pre-K attenders attended Pre-K in a school-based setting. In terms of the second specification of the independent variable, 20 percent of children were co-locaters, meaning they attended kindergarten in the same building where they attended Pre-K, and 13 percent of children were movers, meaning they attended Pre-K in a school building but switched to a different school building for kindergarten.

Independent Variable One: School-Based Versus Non-School-Based Pre-K

I present the results of the main estimation model for independent variable one (school-based Pre-K versus non-school-based Pre-K) in the last column of Table 5.2. The rows are separated by outcome measure and assessment wave within each outcome (Fall K. through Spring 3rd). Interestingly, I find no significant relationships between school-based Pre-K attendance on any of the outcome measures for any assessment wave. Furthermore, for each outcome, the point estimates are closely centered around zero, suggesting no effect, regardless of statistical significance.

Independent Variable Two: Co-location Versus Movers Versus Non-School-Based Pre-K

I present the results of the main estimation model for independent variable two (co-locater versus mover versus non-school-based Pre-K) in Table 5.3. In this table, each consecutive column provides the results for every assessment wave. The rows are separated by outcome measure and the two indicator variables of interest: Co-locater and Mover. The reference category is no school-based Pre-K.

In terms of co-location, I find no statistically significant difference between co-locaters and children who never attended school-based Pre-K for any outcome and for any assessment wave. In terms of movers, I find four significant differences between movers and children who never attended school-based Pre-K. In terms of math achievement in the second grade, I find that movers perform 0.167 standard deviations lower, on average, than their peers who never attended school-based Pre-K. This significant relationship remains and is larger by the end of third grade (-0.203**). In terms of reading achievement, I find that movers perform 0.157 standard deviations lower, on average, than their peers who never attended school-based Pre-K at the end of third grade. It is notable that these significant academic achievement results were only significant in second and third grade, and not immediately following the conclusion of Pre-K. Finally, I also find that in the spring of kindergarten, movers score 0.155 standard deviations higher on the internalizing problem behaviors scale, on average, than their peers who never attended school-based Pre-K. Note that positive scores on this measure suggest more internalizing problem behaviors—a negative outcome.

Despite the limited number of statistically significant relationships presented in Table 5.3, an examination of the trends in the coefficients is useful. For nearly every outcome measure, the coefficient for co-location is positive or very close to zero. Conversely, the

coefficients for the movers are almost universally negative (in the case of problem behaviors, they are positive, which is a negative outcome). Examination of the results from this perspective suggests that co-location, in general, is positively related to student outcomes while the relationship is negative for the movers.

Finally, the coefficients for movers for the two academic achievement outcomes grew in magnitude from the fall of kindergarten to the spring of third grade. As mentioned previously, the coefficients for the movers were both statistically significant in the spring of third grade—the assessment wave when the magnitude of the coefficients were largest. It was surprising to see latent significant results and magnitudes that rose over time for the movers. This finding led me to wonder if the mover variable was not only capturing students who moved to a new school from Pre-K to kindergarten, but also captured transient students who frequently switched schools. To test this, I re-estimated the models with a covariate measure that captured the number of school changes a child experienced from kindergarten through third grade. Interestingly, the results were robust to this alternative specification, suggesting that the mover variable is not a proxy for transient students, beyond the Pre-K-kindergarten transition.

Subgroup Analysis

Independent Variable One

In Table 5.4, I present the results from the subgroup analyses for the first specification of the independent variable: school-based versus non-school-based Pre-K. Each column in this table represents a different subgroup, with the first column re-presenting the main results as a reference. The first subgroup of students is the model restricted to only children whose families were in the bottom 25th percentile of the socio-economic status distribution. In this model, I find

two statistically significant relationships. First, I find that low-SES school-based Pre-K attenders perform 0.058 standard deviations higher, on average, than their peers who attended non-school based Pre-K on the numbers reversed assessment in the fall of kindergarten. However, this relationship flips sign and is not significant in subsequent assessment waves. The second significant relationship for this subgroup is for interpersonal skills in the spring of kindergarten. Here, I find that low-SES school-based Pre-K attenders perform 0.038 standard deviations lower, on average, than their peers who attended non-school-based Pre-K.

The second subgroup of interest is state-funded Pre-K attenders. For this subgroup, I find no significant relationships aside from the self-control outcome. Here, I find students in state-funded Pre-K programs that attended school-based Pre-K score 0.400 standard deviations higher, on average, than their peers who attended non-school-based Pre-K in the fall of kindergarten. The significant positive relationship persists through the end of first grade, with coefficients of 0.208 standard deviations in the spring of kindergarten and 0.234 standard deviations in the spring of first grade, respectively.

The third subgroup of interest includes students who live in urban areas. In this model specification, all statistically significant results are isolated to the two academic achievement outcomes. For math achievement, I find generally consistent negative associations across all assessment waves. The magnitude of the coefficients is consistent, ranging from -0.116 in the fall of kindergarten to -0.183 in the spring of third grade. While the coefficient is similar in magnitude, the association is not statistically significant in the spring of first grade. For reading achievement, there are consistently negative associations across all assessment waves but the only statistically significant finding is for the spring of third grade (-0.126 standard deviations).

In sum, these findings suggest a negative association between school-based Pre-K attendance and math and reading outcomes for urban students in kindergarten through third grade.

The final two subgroups, school Annual Yearly Progress (AYP) and high transition practices, were focused on understanding the potential role of elementary school quality and school practices related to vertical alignment. In the school AYP model, I find no statistically significant relationships for any outcome. In terms of the high transitions practices model, I find a single significant relationship for the internalizing problem behaviors outcome in second grade. Here, I find that school-based Pre-K attenders in high transition practice schools score 0.211 standard deviations lower than their peers who attended non-school-based Pre-K.

Independent Variable Two

In Table 5.5, I present the results from the subgroup analyses for the second specification of the independent variable: co-location versus non-school-based Pre-K and movers versus non-school-based. For the low-SES subgroup, I find large, consistent negative associations for the movers in terms of academic achievement. In math, the coefficients for the movers are statistically significant in every wave and range from -0.368 standard deviations to -0.592 standard deviations. In reading, the coefficients for the movers are also statistically significant in every wave and range from -0.308 standard deviations to -0.515 standard deviations. In addition to the consistent academic achievement findings, I also find two other statistically significant findings, one in the spring of first grade for number reversed (-0.474) and one in the spring of second grade for DCCS (-0.426). While these are isolated relationships within the two executive function measures, it is notable that nearly all coefficients for the co-locaters are positive and all coefficients for the movers are negative.

The results for the second subgroup of interest, state-funded Pre-K attenders, closely mirror the results from the main results—suggesting minimal difference for this specific type of center-based Pre-K. There are two exceptions, however, where the results deviate from the full sample. First, in terms of self-control, I find that both co-locaters and movers in state-funded Pre-K programs enter kindergarten with higher scores than their non-school-based Pre-K peers do. The coefficient for the co-locaters is 0.376 and the coefficient for the movers is 0.457. Second, in terms of externalizing problem behaviors, I two somewhat contradictory results. At kindergarten entry, the movers score 0.315 standard deviations lower, on average, than the non-school-based Pre-K attenders. By the end of kindergarten, there is no difference between the movers and the non-school-based Pre-K attenders but the co-locaters score 0.278 standard deviations lower than the non-school-based Pre-K attenders.

For the urban subgroup, I find moderately sized and consistent negative relationships for both co-locaters and movers in terms of math achievement. While the majority of the statistically significant relationships are isolated to the co-locaters, the coefficients are remarkably similar, suggesting the measure is not distinguishing between the two groups. While no statistically significant relationships emerged for reading, the coefficients follow the same, consistently negative, pattern. Beyond the achievement outcomes, there are only two other notable relationships—both for the internalizing problem behaviors outcome. I find that among students in urban areas, being a mover is associated with higher internalizing problem behaviors in first grade (0.205). Additionally, being a co-locater is associated with fewer internalizing problem behaviors in third grade (-0.145).

For the school AYP model, I only find a single statistically significant relationship. In the spring of third grade, I find that movers in schools that met AYP perform 0.164 standard

deviations worse, on average, than their peers who did not attend school-based Pre-K. Note that the finding for this outcome and wave was also significant in the main effects model. Finally, in the high transition practices model, I find a single significant relationship between co-location and internalizing problem behaviors in the spring of second grade (-0.256). The original intent of these two final subgroups—AYP and transition practices—was to understand the latent significant results for the movers in terms of the academic achievement outcomes. While the results that were significant in the main effects model are generally no longer significant in these models, that is likely a function of reduced power. The general pattern of the findings is the same, with positive effects for the co-locaters and negative effects that grow over time for the movers. That said, the positive effects for the co-locaters are almost universally larger in bother the AYP and transition practices models.

Table 5.1. Summary Statistics of Outcomes, Independent Variables, and Control Variables

Achievement Outcomes				
Math	**Mean**	**SD**	**Min.**	**Max.**
Fall K. Math	35.17	11.61	-0.76	128.68
Spring K. Math	48.81	12.50	8.18	95.03
Spring 1st Math	72.17	16.86	1.49	123.28
Spring 2nd Math	87.31	15.41	13.66	132.73
Spring 3rd Math	98.52	14.24	36.14	142.38
Reading				
Fall K. Reading	53.35	11.63	23.70	110.27
Spring K. Reading	67.05	13.67	29.97	118.35
Spring 1st Reading	90.99	16.14	36.80	134.69
Spring 2nd Reading	103.62	13.41	52.00	142.15
Spring 3rd Reading	111.25	12.41	58.69	148.07
Executive Function Outcomes				
Working Memory				
Fall K. Working Memory	437.36	30.58	341.14	581.00
Spring K. Working Memory	453.17	29.99	356.57	544.00
Spring 1st Working Memory	471.78	25.55	378.83	567.00
Spring 2nd Working Memory	481.92	22.83	402.05	581.00
Spring 3rd Working Memory	490.50	22.13	403.00	581.00
Cognitive Flexibility				
Fall K. Cognitive Flexibility	14.43	3.22	0.00	22.77
Spring K. Cognitive Flexibility	15.25	2.71	0.00	23.50
Spring 1st Cognitive Flexibility	16.17	2.27	0.00	23.46
Spring 2nd Cognitive Flexibility	6.69	1.30	1.70	10.19
Spring 3rd Cognitive Flexibility	7.19	1.07	1.75	10.87
Social-Emotional Outcomes				
Approaches to Learning				
Fall K. Approaches	2.96	0.68	0.99	5.14
Spring K. Approaches	3.10	0.70	0.75	5.68
Spring 1st Approaches	3.05	0.72	0.60	5.32
Spring 2nd Approaches	3.05	0.71	0.69	5.37
Spring 3nd Approaches	3.04	0.74	0.50	5.94
Self Control				
Fall K. Self Control	3.05	0.63	1.00	4.83
Spring K. Self Control	3.13	0.66	1.00	4.91
Spring 1st Self Control	3.15	0.65	1.00	5.21
Spring 2nd Self Control	3.14	0.65	0.93	5.04
Spring 3rd Self Control	3.18	0.66	0.86	5.43

Table 5.1. Continued.

	Mean	SD	Min.	Max.
Interpersonal Skills				
Fall K. Interpersonal	2.99	0.64	1.00	4.82
Spring K. Interpersonal	3.12	0.67	0.82	4.87
Spring 1st Interpersonal	3.09	0.66	1.00	5.34
Spring 2nd Interpersonal	3.07	0.68	0.69	5.29
Spring 3rd Interpersonal	3.08	0.69	0.69	5.20
Internalizing Problem Behaviors				
Fall K. Internalizing	1.66	0.65	-0.28	4.00
Spring K. Internalizing	1.70	0.67	-0.66	4.00
Spring 1st Internalizing	1.80	0.65	-0.11	4.28
Spring 2nd Internalizing	1.79	0.65	-0.13	4.00
Spring 3rd Internalizing	1.78	0.66	-1.09	4.27
Externalizing Problem Behaviors				
Fall K. Externalizing	1.44	0.47	-0.07	4.00
Spring K. Externalizing	1.50	0.50	0.18	4.00
Spring 1st Externalizing	1.55	0.51	-0.28	4.00
Spring 2nd Externalizing	1.59	0.51	0.05	4.00
Spring 3rd Externalizing	1.62	0.54	-0.34	4.00
Independent Variables				
School-Based Pre-K	0.33	0.33	0.00	1.00
Pre-K Co-Locater	0.20	0.19	0.00	1.00
Pre-K Mover	0.13	0.13	0.00	1.00
No School-Based Pre-K	0.67	0.67	0.00	1.00
Covariates				
White	0.47	0.50	0.00	1.00
Black	0.21	0.39	0.00	1.00
Hispanic	0.20	0.40	0.00	1.00
Asian	0.05	0.26	0.00	1.00
Other Race	0.07	0.25	0.00	1.00
Socioeconomic Status	0.03	0.81	-2.33	2.44
Male	0.52	0.50	0.00	1.00
Total Siblings in Home	1.29	1.04	0.00	9.00
Two Parent Household	0.71	0.46	0.00	1.00
Age at K. Entry	66.16	4.75	36.60	83.30
Non-English Spoken at Home	0.11	0.32	0.00	1.00
Food Security Index	0.77	2.03	-6.06	15.00
Parent Age	34.94	7.11	18.00	75.00
Birthweight	115.97	21.78	19.00	221.00
Maternal WIC Receipt	0.45	0.50	0.00	1.00
Married at Birth	0.63	0.48	0.00	1.00
District Poverty Rate	22.29	48.94	-11.30	3328.00
Urban	0.67	0.46	0.00	1.00
Northeast	0.17	0.37	0.00	1.00
Midwest	0.17	0.38	0.00	1.00

Table 5.1. Continued.

	Mean	SD	Min.	Max.
South	0.50	0.50	0.00	1.00
West	0.16	0.38	0.00	1.00
Time to School-Based Pre-K	10.51	21.56	-36.02	234.44
Hours/Week in Pre-K	33.82	8.35	20.00	70.00

Notes. N=4800, rounded to the nearest 10 in accoradance with NCES regulations.
SD= Standard Deviation, Min.= Minimum, Max.= Maximum

Table 5.2. School-Based Pre-K Main Effects

	Controls	LEA FE	State FE	PS Weight + State FE	PS Weight w/GIS + State FE
Math					
Fall K.	-0.042	-0.073	-0.062	-0.049	-0.045
	(0.046)	(0.052)	(0.046)	(0.047)	(0.047)
Spring K.	-0.091	-0.095	**-0.095 ***	-0.079	-0.072
	(0.048)	(0.083)	**(0.047)**	(0.049)	(0.049)
Spring 1st	-0.042	-0.074	-0.060	-0.040	-0.040
	(0.052)	(0.056)	(0.049)	(0.052)	(0.052)
Spring 2nd	-0.056	-0.078	-0.070	-0.050	-0.051
	(0.055)	(0.058)	(0.050)	(0.055)	(0.055)
Spring 3rd	-0.062	-0.093	-0.086	-0.059	-0.066
	(0.052)	(0.058)	(0.050)	(0.054)	(0.054)
Reading					
Fall K.	-0.025	-0.052	-0.063	-0.050	-0.044
	(0.044)	(0.051)	(0.045)	(0.046)	(0.046)
Spring K.	-0.035	-0.033	-0.052	-0.043	-0.036
	(0.047)	(0.055)	(0.047)	(0.049)	(0.049)
Spring 1st	-0.044	-0.057	-0.060	-0.051	-0.048
	(0.052)	(0.058)	(0.051)	(0.053)	(0.053)
Spring 2nd	-0.031	-0.043	-0.046	-0.031	-0.039
	(0.052)	(0.057)	(0.051)	(0.055)	(0.055)
Spring 3rd	-0.053	-0.070	-0.073	-0.068	-0.075
	(0.051)	(0.055)	(0.049)	(0.053)	(0.053)
Numbers Reversed					
Fall K.	-0.011	-0.017	-0.007	0.007	0.027
	(0.052)	(0.058)	(0.053)	(0.056)	(0.055)
Spring K.	-0.007	0.004	-0.010	0.011	0.024
	(0.057)	(0.062)	(0.057)	(0.061)	(0.062)
Spring 1st	-0.062	-0.047	-0.080	-0.084	-0.085
	(0.057)	(0.063)	(0.053)	(0.057)	(0.057)
Spring 2nd	0.063	0.086	0.053	0.089	0.085
	(0.063)	(0.068)	(0.058)	(0.061)	(0.061)
Spring 3rd	0.009	-0.020	-0.010	0.019	0.021
	(0.054)	(0.063)	(0.052)	(0.054)	(0.054)

Table 5.2. Continued.

	Controls	LEA FE	State FE	PS Weight + State FE	PS Weight w/GIS + State FE
DCCS					
Fall K.	-0.011	0.033	-0.025	-0.010	0.006
	(0.056)	(0.062)	(0.056)	(0.060)	(0.061)
Spring K.	**-0.120 ***	-0.030	-0.090	-0.046	-0.040
	(0.061)	(0.068)	(0.059)	(0.061)	(0.062)
Spring 1st	0.008	0.027	-0.008	0.010	0.019
	(0.053)	(0.061)	(0.051)	(0.054)	(0.054)
Spring 2nd	0.004	0.006	-0.019	-0.029	-0.020
	(0.063)	(0.067)	(0.060)	(0.063)	(0.063)
Spring 3rd	0.031	-0.070	-0.032	-0.010	-0.012
	(0.056)	(0.064)	(0.056)	(0.058)	(0.058)
Approaches to Learning					
Fall K.	-0.052	-0.064	-0.054	-0.029	-0.033
	(0.052)	(0.059)	(0.052)	(0.055)	(0.055)
Spring K.	0.002	0.005	0.019	0.051	0.046
	(0.054)	(0.061)	(0.054)	(0.057)	(0.058)
Spring 1st	-0.006	-0.014	-0.013	0.012	0.008
	(0.053)	(0.060)	(0.053)	(0.055)	(0.056)
Spring 2nd	-0.022	-0.061	-0.036	-0.019	-0.019
	(0.054)	(0.057)	(0.054)	(0.055)	(0.055)
Spring 3rd	0.005	-0.023	-0.008	0.011	0.006
	(0.054)	(0.060)	(0.054)	(0.057)	(0.058)
Self-Control					
Fall K.	0.001	0.017	-0.001	0.041	0.041
	(0.056)	(0.064)	(0.057)	(0.059)	(0.060)
Spring K.	-0.019	-0.012	-0.008	0.024	0.019
	(0.054)	(0.057)	(0.054)	(0.055)	(0.055)
Spring 1st	0.005	-0.003	-0.013	0.017	0.014
	(0.054)	(0.062)	(0.056)	(0.057)	(0.058)
Spring 2nd	-0.022	-0.071	-0.045	-0.032	-0.040
	(0.056)	(0.058)	(0.056)	(0.058)	(0.057)
Spring 3rd	-0.032	-0.027	-0.026	-0.030	-0.025
	(0.054)	(0.058)	(0.055)	(0.056)	(0.058)

Table 5.2. Continued.

	Controls	LEA FE	State FE	PS Weight + State FE	PS Weight w/GIS + State FE
Interpersonal Skills					
Fall K.	-0.044	-0.079	-0.047	-0.036	-0.034
	(0.054)	(0.058)	(0.054)	(0.055)	(0.056)
Spring K.	-0.044	-0.044	-0.040	-0.006	-0.003
	(0.052)	(0.059)	(0.053)	(0.065)	(0.056)
Spring 1st	-0.026	0.003	-0.027	0.018	0.024
	(0.057)	(0.063)	(0.058)	(0.058)	(0.059)
Spring 2nd	-0.000	-0.065	-0.013	-0.010	-0.012
	(0.054)	(0.058)	(0.054)	(0.055)	(0.055)
Spring 3rd	-0.029	-0.039	-0.018	-0.019	-0.015
	(0.054)	(0.060)	(0.056)	(0.057)	(0.058)
Externalizing Problem Behaviors					
Fall K.	0.032	-0.001	0.010	-0.025	-0.025
	(0.058)	(0.066)	(0.061)	(0.062)	(0.062)
Spring K.	-0.004	0.009	-0.024	-0.049	-0.040
	(0.054)	(0.062)	(0.056)	(0.056)	(0.056)
Spring 1st	-0.014	-0.027	-0.006	-0.027	-0.029
	(0.053)	(0.059)	(0.056)	(0.056)	(0.057)
Spring 2nd	-0.044	-0.015	-0.032	-0.056	-0.053
	(0.058)	(0.061)	(0.059)	(0.059)	(0.058)
Spring 3rd	-0.006	0.019	0.002	-0.011	-0.011
	(0.054)	(0.058)	(0.054)	(0.053)	(0.054)
Internalizing Problem Behaviors					
Fall K.	0.060	-0.004	0.013	0.017	0.016
	(0.059)	(0.066)	(0.060)	(0.061)	(0.061)
Spring K.	0.095	0.047	0.037	0.009	0.020
	(0.061)	(0.065)	(0.061)	(0.062)	(0.061)
Spring 1st	0.002	-0.030	0.006	-0.029	-0.036
	(0.058)	(0.062)	(0.060)	(0.059)	(0.060)
Spring 2nd	-0.012	0.019	0.001	-0.009	0.001
	(0.059)	(0.064)	(0.059)	(0.063)	(0.062)
Spring 3rd	-0.048	-0.068	-0.064	-0.076	-0.072
	(0.052)	(0.060)	(0.055)	(0.055)	(0.056)

Notes: LEA FE includes controls and LEA fixed effects. State FE includes controls and State fixed effects. PS=Propensity Score. The final model includes the PS Weight derived with controls and GIS-calculated travel distance to nearest school-based Pre-K plus State fixed effects.

Table 5.3. Co-Location and Movers Main Results

	Fall K.	Spring K.	Spring 1st	Spring 2nd	Spring 3rd
Math					
Co-locater	0.004	-0.034	0.022	0.025	0.029
	(0.056)	(0.058)	(0.061)	(0.064)	(0.062)
Mover	-0.120	-0.034	-0.138	**-0.167 ***	**-0.213 ****
	(0.066)	(0.058)	(0.078)	**(0.083)**	**(0.080)**
Reading					
Co-locater	-0.022	-0.004	-0.005	-0.001	-0.022
	(0.054)	(0.057)	(0.063)	(0.065)	(0.062)
Mover	-0.080	-0.086	-0.114	-0.099	**-0.157 ***
	(0.067)	(0.076)	(0.079)	(0.081)	**(0.077)**
Numbers Reversed					
Co-locater	0.014	0.063	-0.034	0.099	0.066
	(0.063)	(0.072)	(0.065)	(0.072)	(0.062)
Mover	0.047	-0.036	-0.163	0.062	-0.049
	(0.080)	(0.083)	(0.087)	(0.092)	(0.088)
DCCS					
Co-locater	0.058	0.021	0.063	0.040	0.060
	(0.068)	(0.071)	(0.063)	(0.075)	(0.064)
Mover	-0.075	-0.134	-0.049	-0.111	-0.125
	(0.088)	(0.102)	(0.072)	(0.104)	(0.093)
Approaches					
Co-locater	0.015	0.098	0.101	0.023	0.039
	(0.066)	(0.068)	(0.067)	(0.065)	(0.069)
Mover	-0.108	-0.034	-0.137	-0.083	-0.046
	(0.075)	(0.076)	(0.078)	(0.076)	(0.079)
Self-Control					
Co-locater	0.050	0.033	0.070	-0.027	-0.007
	(0.070)	(0.065)	(0.068)	(0.067)	(0.068)
Mover	0.026	-0.003	-0.074	-0.061	-0.052
	(0.082)	(0.077)	(0.078)	(0.077)	(0.074)
Interpersonal Skills					
Co-locater	0.010	0.033	0.080	0.019	0.042
	(0.064)	(0.064)	(0.068)	(0.064)	(0.069)
Mover	-0.103	-0.059	-0.064	-0.059	-0.103
	(0.079)	(0.079)	(0.080)	(0.075)	(0.076)
Externalizing Problem Behaviors					
Co-locater	-0.018	-0.066	-0.102	-0.075	0.007
	(0.071)	(0.064)	(0.066)	(0.071)	(0.067)
Mover	-0.035	0.000	0.082	-0.020	-0.038
	(0.077)	(0.073)	(0.078)	(0.077)	(0.068)
Internalizing Problem Behaviors					
Co-locater	-0.062	-0.067	-0.118	-0.029	-0.095
	(0.071)	(0.069)	(0.069)	(0.071)	(0.066)
Mover	0.089	**0.155 ***	0.091	0.049	-0.036
	(0.088)	**(0.087)**	(0.086)	(0.083)	(0.072)

Table 5.4. School-Based Pre-K Subgroup Effects

	Main Effects	Low-SES	State-Funded Pre-K	Urban	School AYP	High Transition Practices
	(*n* =4800)	(*n* =1210)	(*n* =1150)	(*n* =3350)	(*n* =2500)	(*n* =2070)
Math						
Fall K.	-0.045	-0.120	-0.073	**-0.116 ***	-0.027	-0.024
	(0.047)	(0.100)	(0.092)	**(0.056)**	(0.067)	(0.078)
Spring K.	-0.072	-0.226	-0.098	**-0.163 ****	-0.063	-0.003
	(0.049)	(0.104)	(0.112)	**(0.057)**	(0.068)	(0.073)
Spring 1st	-0.040	-0.123	-0.042	-0.109	-0.042	0.008
	(0.052)	(0.107)	(0.127)	(0.058)	(0.068)	(0.082)
Spring 2nd	-0.051	-0.139	-0.062	**-0.147 ***	0.002	0.024
	(0.055)	(0.134)	(0.128)	**(0.062)**	(0.071)	(0.086)
Spring 3rd	-0.066	-0.127	-0.006	**-0.183 ****	-0.029	0.002
	(0.054)	(0.126)	(0.128)	**(0.061)**	(0.070)	(0.084)
Reading						
Fall K.	-0.044	-0.143	-0.068	-0.102	-0.046	0.029
	(0.046)	(0.087)	(0.094)	(0.060)	(0.064)	(0.069)
Spring K.	-0.036	-0.103	-0.030	-0.089	-0.033	0.109
	(0.049)	(0.099)	(0.101)	(0.063)	(0.066)	(0.072)
Spring 1st	-0.048	-0.163	-0.103	-0.107	-0.023	0.040
	(0.053)	(0.112)	(0.114)	(0.065)	(0.069)	(0.085)
Spring 2nd	-0.039	-0.287	-0.102	-0.113	-0.040	0.027
	(0.055)	(0.129)	(0.116)	(0.065)	(0.076)	(0.085)
Spring 3rd	-0.075	-0.226	-0.129	**-0.126 ***	-0.119	-0.005
	(0.053)	(0.123)	(0.107)	**(0.063)**	(0.074)	(0.079)
Numbers Reversed						
Fall K.	0.027	**0.058 ***	0.034	0.019	0.076	-0.041
	(0.055)	**(0.110)**	(0.117)	(0.064)	(0.075)	(0.077)
Spring K.	0.024	-0.029	-0.019	-0.015	-0.024	0.011
	(0.062)	(0.147)	(0.124)	(0.069)	(0.087)	(0.088)
Spring 1st	-0.085	-0.285	-0.148	-0.086	-0.122	-0.032
	(0.057)	(0.126)	(0.116)	(0.067)	(0.076)	(0.092)
Spring 2nd	0.085	0.003	0.057	0.049	0.115	0.144
	(0.061)	(0.131)	(0.132)	(0.074)	(0.078)	(0.103)
Spring 3rd	0.021	0.050	-0.028	0.007	0.041	0.091
	(0.054)	(0.115)	(0.106)	(0.069)	(0.073)	(0.088)
DCCS						
Fall K.	0.006	0.049	-0.003	0.019	-0.010	-0.094
	(0.061)	(0.132)	(0.123)	(0.073)	(0.086)	(0.094)
Spring K.	-0.040	0.065	-0.156	-0.002	-0.010	-0.114
	(0.062)	(0.119)	(0.270)	(0.074)	(0.094)	(0.096)
Spring 1st	0.019	0.008	0.076	0.028	0.004	0.123
	(0.054)	(0.099)	(0.117)	(0.068)	(0.079)	(0.088)
Spring 2nd	-0.020	-0.174	-0.076	-0.036	-0.081	0.006
	(0.063)	(0.135)	(0.127)	(0.077)	(0.088)	(0.097)
Spring 3rd	-0.012	-0.122	0.071	-0.079	-0.009	0.041
	(0.058)	(0.130)	(0.130)	(0.067)	(0.080)	(0.099)

Table 5.4. Continued.

	Main Effects	Low-SES	State-Funded Pre-K	Urban	School AYP	High Transition Practices
Approaches to Learning						
Fall K.	-0.033	-0.103	0.115	-0.041	-0.077	0.025
	(0.055)	(0.126)	(0.119)	(0.066)	(0.076)	(0.086)
Spring K.	0.046	-0.118	0.187	0.042	0.009	0.042
	(0.058)	(0.128)	(0.136)	(0.068)	(0.082)	(0.090)
Spring 1st	0.008	-0.047	0.115	-0.065	-0.021	0.062
	(0.056)	(0.120)	(0.124)	(0.066)	(0.080)	(0.094)
Spring 2nd	-0.019	-0.065	0.123	-0.064	0.013	0.079
	(0.055)	(0.108)	(0.130)	(0.065)	(0.075)	(0.094)
Spring 3rd	0.006	-0.020	-0.053	0.021	-0.040	0.040
	(0.058)	(0.119)	(0.127)	(0.066)	(0.081)	(0.083)
Self-Control						
Fall K.	0.041	0.092	**0.400 ****	-0.025	0.004	0.131
	(0.060)	(0.124)	**(0.124)**	(0.072)	(0.083)	(0.085)
Spring K.	0.019	-0.018	**0.208 ***	0.010	0.036	-0.013
	(0.055)	(0.116)	**(0.110)**	(0.070)	(0.075)	(0.080)
Spring 1st	0.014	0.145	**0.234 ***	-0.025	-0.022	0.048
	(0.058)	(0.126)	**(0.118)**	(0.067)	(0.082)	(0.091)
Spring 2nd	-0.040	-0.119	0.075	-0.067	0.008	-0.054
	(0.057)	(0.109)	(0.131)	(0.068)	(0.083)	(0.088)
Spring 3rd	-0.025	0.022	-0.085	0.003	-0.142	-0.031
	(0.058)	(0.109)	(0.117)	(0.069)	(0.081)	(0.077)
Interpersonal Skills						
Fall K.	-0.034	-0.064	0.200	-0.046	-0.039	0.086
	(0.056)	(0.120)	(0.113)	(0.068)	(0.077)	(0.084)
Spring K.	-0.003	**-0.038 ***	0.173	-0.042	0.007	0.018
	(0.056)	**(0.124)**	(0.111)	(0.067)	(0.081)	(0.086)
Spring 1st	0.024	0.174	0.154	-0.036	0.029	0.128
	(0.059)	(0.125)	(0.129)	(0.067)	(0.084)	(0.092)
Spring 2nd	-0.012	-0.037	0.078	0.009	-0.009	0.034
	(0.055)	(0.106)	(0.127)	(0.066)	(0.080)	(0.086)
Spring 3rd	-0.015	-0.044	0.087	0.010	-0.071	-0.031
	(0.058)	(0.112)	(0.114)	(0.067)	(0.080)	(0.086)
Externalizing Problem Behaviors						
Fall K.	-0.025	-0.125	-0.189	0.039	0.026	-0.002
	(0.062)	(0.128)	(0.125)	(0.077)	(0.082)	(0.093)
Spring K.	-0.040	-0.035	-0.253	-0.047	0.002	0.041
	(0.056)	(0.105)	(0.137)	(0.068)	(0.076)	(0.084)
Spring 1st	-0.029	-0.184	-0.117	0.055	-0.001	-0.031
	(0.057)	(0.117)	(0.107)	(0.067)	(0.079)	(0.094)
Spring 2nd	-0.053	-0.039	-0.235	0.008	-0.076	0.012
	(0.058)	(0.110)	(0.122)	(0.070)	(0.083)	(0.095)
Spring 3rd	-0.011	-0.003	0.064	0.005	-0.018	-0.031
	(0.054)	(0.106)	(0.122)	(0.067)	(0.076)	(0.080)

Table 5.4. Continued.

	Main Effects	Low-SES	State-Funded Pre-K	Urban	School AYP	High Transition Practices
Internalizing Problem Behaviors						
Fall K.	0.016	0.100	0.075	0.033	0.075	0.000
	(0.061)	(0.136)	(0.114)	(0.076)	(0.084)	(0.088)
Spring K.	0.020	0.015	-0.058	0.057	0.055	0.054
	(0.061)	(0.120)	(0.148)	(0.077)	(0.085)	(0.089)
Spring 1st	-0.036	-0.186	-0.157	0.074	0.087	-0.102
	(0.060)	(0.125)	(0.131)	(0.068)	(0.078)	(0.103)
Spring 2nd	0.001	0.026	-0.291	0.041	0.060	**-0.211 ***
	(0.062)	(0.127)	(0.152)	(0.071)	(0.092)	**(0.093)**
Spring 3rd	-0.072	-0.045	-0.083	-0.056	-0.041	-0.118
	(0.056)	(0.126)	(0.131)	(0.062)	(0.076)	(0.095)

Notes: All sample sizes are rounded to the nearest ten for NCES restricted-use data.
SES=Socio-economic Status. AYP= Annual Yearly Progress

Table 5.5. Colocation and Movers Subgroup Effects

	Main Effects	Low-SES	State-Funded Pre-K	Urban	School AYP	High Transition Practices
Math	(*n* =4800)	(*n* =1210)	(*n* =1150)	(*n* =3350)	(*n* =2500)	(*n* =2070)
Fall K.						
Colocate	0.004	0.073	-0.032	-0.117	0.053	0.079
	(0.056)	-0.113	(0.100)	(0.069)	(0.088)	(0.095)
Mover	-0.120	**-0.368**	-0.169	-0.114	-0.113	-0.187
	(0.066)	**(0.127) ****	(0.125)	(0.076)	(0.076)	(0.116)
Spring K.						
Colocate	-0.034	0.000	-0.070	**-0.190 ****	-0.021	0.092
	(0.058)	(0.122)	(0.118)	**(0.067)**	(0.090)	(0.088)
Mover	-0.034	**-0.515 *****	-0.162	-0.127	-0.109	-0.155
	(0.058)	**(0.140)**	(0.160)	(0.081)	(0.082)	(0.128)
Spring 1st						
Colocate	0.022	0.132	-0.005	**-0.134 ***	0.052	0.134
	(0.061)	(0.121)	(0.135)	**(0.067)**	(0.091)	(0.095)
Mover	-0.138	**-0.449 ****	-0.125	-0.075	-0.143	-0.192
	(0.078)	**(0.146)**	(0.170)	(0.081)	(0.085)	(0.143)
Spring 2nd						
Colocate	0.025	0.179	-0.029	**-0.146 ***	0.083	0.160
	(0.064)	(0.146)	(0.138)	**(0.073)**	(0.099)	(0.103)
Mover	**-0.167 ***	**-0.548 ****	-0.139	-0.147	-0.085	-0.194
	(0.083)	**(0.180)**	(0.172)	(0.089)	(0.084)	(0.148)
Spring 3rd						
Colocate	0.029	0.236	0.060	**-0.176 ***	0.097	0.142
	(0.062)	(0.143)	(0.137)	**(0.071)**	(0.093)	(0.091)
Mover	**-0.213 ****	**-0.592 *****	-0.156	**-0.193 ***	**-0.164 ***	-0.220
	(0.080)	**(0.161)**	(0.169)	**(0.086)**	**(0.081)**	(0.153)
Reading						
Fall K.						
Colocate	-0.022	-0.013	-0.014	-0.136	0.045	0.056
	(0.054)	(0.109)	(0.101)	(0.072)	(0.084)	(0.080)
Mover	-0.080	**-0.308 ****	-0.194	-0.055	-0.145	-0.013
	(0.067)	**(0.110)**	(0.130)	(0.085)	(0.075)	(0.113)
Spring K.						
Colocate	-0.004	0.060	0.013	-0.125	0.078	0.124
	(0.057)	(0.120)	(0.109)	(0.075)	(0.088)	(0.082)
Mover	-0.086	**-0.312 ***	-0.130	-0.040	-0.152	0.084
	(0.076)	**(0.138)**	(0.160)	(0.090)	(0.087)	(0.124)
Spring 1st						
Colocate	-0.005	-0.023	-0.081	-0.147	0.024	0.136
	(0.063)	(0.130)	(0.122)	(0.079)	(0.094)	(0.103)
Mover	-0.114	**-0.342 ***	-0.155	-0.053	-0.073	-0.112
	(0.079)	**(0.156)**	(0.163)	(0.085)	(0.091)	(0.133)
Spring 2nd						
Colocate	-0.001	-0.109	-0.106	-0.126	-0.016	0.156
	(0.065)	(0.146)	(0.126)	(0.080)	(0.103)	(0.099)
Mover	-0.099	**-0.515 ****	-0.093	-0.094	-0.066	-0.180
	(0.081)	**(0.171)**	(0.158)	(0.088)	(0.091)	(0.137)
Spring 3rd						
Colocate	-0.022	-0.037	-0.103	-0.100	-0.062	0.050
	(0.062)	(0.140)	(0.116)	(0.075)	(0.100)	(0.093)
Mover	**-0.157 ***	**-0.468 ****	-0.188	-0.160	-0.181	-0.094
	(0.077)	**(0.165)**	(0.155)	(0.088)	(0.086)	(0.132)

Table 5.5. Continued.

	Main Effects	Low-SES	State-Funded Pre-K	Urban	School AYP	High Transition Practices
Numbers Reversed						
Fall K.						
Colocate	0.014	0.111	0.007	0.001	0.053	-0.065
	(0.063)	(0.128)	(0.124)	(0.071)	(0.098)	(0.089)
Mover	0.047	-0.009	0.098	0.044	0.100	-0.004
	(0.080)	(0.151)	(0.165)	(0.094)	(0.096)	(0.115)
Spring K.						
Colocate	0.063	0.100	-0.067	-0.054	0.046	0.006
	(0.072)	(0.172)	(0.134)	(0.080)	(0.113)	(0.104)
Mover	-0.036	-0.194	0.094	0.038	-0.099	0.020
	(0.083)	(0.175)	(0.171)	(0.091)	(0.103)	(0.122)
Spring 1st						
Colocate	-0.034	-0.137	-0.159	-0.071	-0.134	0.067
	(0.065)	(0.139)	(0.119)	(0.075)	(0.102)	(0.108)
Mover	-0.163	**-0.474 ****	-0.121	-0.101	-0.109	-0.189
	(0.087)	**(0.173)**	(0.165)	(0.100)	(0.099)	(0.154)
Spring 2nd						
Colocate	0.099	0.134	0.050	0.051	0.033	0.194
	(0.072)	(0.150)	(0.138)	(0.088)	(0.111)	(0.121)
Mover	0.062	-0.164	0.073	0.047	0.203	0.065
	(0.092)	(0.182)	(0.172)	(0.102)	(0.100)	(0.177)
Spring 3rd						
Colocate	0.066	0.243	0.021	0.016	0.044	0.106
	(0.062)	(0.134)	(0.113)	(0.082)	(0.090)	(0.096)
Mover	-0.049	-0.198	-0.142	-0.004	0.037	0.067
	(0.088)	(0.166)	(0.164)	(0.105)	(0.093)	(0.164)
DCCS						
Fall K.						
Colocate	0.058	0.154	0.045	0.070	0.003	-0.064
	(0.068)	(0.147)	(0.129)	(0.080)	(0.106)	(0.109)
Mover	-0.075	-0.086	-0.116	-0.051	-0.023	-0.143
	(0.088)	(0.169)	(0.160)	(0.102)	(0.112)	(0.130)
Spring K.						
Colocate	0.021	0.283	-0.063	0.022	0.118	-0.156
	(0.071)	(0.151)	(0.133)	(0.088)	(0.116)	(0.120)
Mover	-0.134	-0.214	-0.373	-0.033	-0.149	-0.046
	(0.102)	(0.166)	(0.207)	(0.112)	(0.132)	(0.138)
Spring 1st						
Colocate	0.063	0.110	0.036	0.041	0.014	0.156
	(0.063)	(0.111)	(0.127)	(0.076)	(0.105)	(0.106)
Mover	-0.049	-0.123	0.170	0.011	-0.007	0.073
	(0.072)	(0.126)	(0.135)	(0.095)	(0.088)	(0.105)
Spring 2nd						
Colocate	0.040	0.022	-0.009	0.001	-0.047	0.102
	(0.075)	(0.167)	(0.144)	(0.094)	(0.115)	(0.113)
Mover	-0.111	**-0.426 ***	-0.231	-0.085	-0.117	-0.147
	(0.104)	**(0.200)**	(0.223)	(0.105)	(0.130)	(0.168)
Spring 3rd						
Colocate	0.060	0.040	0.146	-0.029	0.035	0.142
	(0.064)	(0.151)	(0.134)	(0.080)	(0.094)	(0.118)
Mover	-0.125	-0.331	-0.104	-0.146	-0.057	-0.121
	(0.093)	(0.179)	(0.181)	(0.097)	(0.103)	(0.182)

Table 5.5. Continued.

	Main Effects	Low-SES	State-Funded Pre-K	Urban	School AYP	High Transition Practices
Approaches to Learning						
Fall K.						
Colocate	0.015	0.016	0.091	-0.037	-0.062	0.055
	(0.066)	(0.145)	(0.127)	(0.080)	(0.099)	(0.107)
Mover	-0.108	-0.255	0.169	-0.047	-0.094	-0.023
	(0.075)	(0.160)	(0.152)	(0.087)	(0.098)	(0.109)
Spring K.						
Colocate	0.098	0.012	0.191	0.066	0.034	0.087
	(0.068)	(0.158)	(0.147)	(0.084)	(0.106)	(0.110)
Mover	-0.034	-0.285	0.179	0.010	-0.018	-0.029
	(0.076)	(0.155)	(0.171)	(0.092)	(0.098)	(0.114)
Spring 1st						
Colocate	0.101	0.088	0.193	-0.036	0.105	0.128
	(0.067)	(0.149)	(0.138)	(0.078)	(0.101)	(0.108)
Mover	-0.137	-0.220	-0.065	-0.106	-0.156	-0.042
	(0.078)	(0.146)	(0.177)	(0.091)	(0.100)	(0.125)
Spring 2nd						
Colocate	0.023	-0.018	0.128	-0.098	0.062	0.120
	(0.065)	(0.127)	(0.146)	(0.077)	(0.095)	(0.109)
Mover	-0.083	-0.125	0.111	-0.019	-0.040	0.014
	(0.076)	(0.144)	(0.158)	(0.089)	(0.092)	(0.130)
Spring 3rd						
Colocate	0.039	0.105	0.018	0.008	0.021	0.062
	(0.069)	(0.151)	(0.142)	(0.081)	(0.108)	(0.102)
Mover	-0.046	-0.180	-0.218	0.040	-0.104	0.004
	(0.079)	(0.142)	(0.168)	(0.089)	(0.098)	(0.114)
Self-Control						
Fall K.						
Colocate	0.050	0.093	**0.376 ****	-0.065	0.046	0.094
	(0.070)	(0.149)	**(0.132)**	(0.085)	(0.105)	(0.102)
Mover	0.026	0.090	**0.457 ****	0.030	-0.042	0.190
	(0.082)	(0.160)	**(0.157)**	(0.098)	(0.108)	(0.112)
Spring K.						
Colocate	0.033	-0.004	0.195	0.006	0.025	-0.005
	(0.065)	(0.136)	(0.125)	(0.086)	(0.096)	(0.097)
Mover	-0.003	-0.035	0.237	0.015	0.048	-0.026
	(0.077)	(0.149)	(0.159)	(0.095)	(0.096)	(0.103)
Spring 1st						
Colocate	0.070	0.181	**0.296 ***	0.006	0.037	0.075
	(0.068)	(0.147)	**(0.130)**	(0.081)	(0.103)	(0.103)
Mover	-0.074	0.099	0.091	-0.068	-0.086	0.004
	(0.078)	(0.150)	(0.162)	(0.093)	(0.098)	(0.128)
Spring 2nd						
Colocate	-0.027	-0.164	0.060	-0.067	-0.013	-0.133
	(0.067)	(0.130)	(0.143)	(0.079)	(0.106)	(0.105)
Mover	-0.061	-0.062	0.111	-0.065	0.031	0.073
	(0.077)	(0.136)	(0.154)	(0.091)	(0.099)	(0.118)
Spring 3rd						
Colocate	-0.007	0.042	-0.057	0.015	-0.123	0.006
	(0.068)	(0.145)	(0.128)	(0.084)	(0.103)	(0.097)
Mover	-0.052	-0.004	-0.148	-0.013	-0.162	-0.090
	(0.074)	(0.118)	(0.144)	(0.094)	(0.094)	(0.095)

Table 5.5. Continued.

	Main Effects	Low-SES	State-Funded Pre-K	Urban	School AYP	High Transition Practices
Interpersonal Skills						
Fall K.						
Colocate	0.010	0.012	0.161	-0.062	-0.015	0.169
	(0.064)	(0.138)	(0.123)	(0.077)	(0.094)	(0.101)
Mover	-0.103	-0.162	0.292	-0.024	-0.064	-0.046
	(0.079)	(0.154)	(0.140)	(0.093)	(0.104)	(0.115)
Spring K.						
Colocate	0.033	0.027	0.177	-0.050	-0.013	0.053
	(0.064)	(0.144)	(0.122)	(0.077)	(0.099)	(0.105)
Mover	-0.059	-0.121	0.165	-0.032	0.028	-0.039
	(0.079)	(0.154)	(0.142)	(0.095)	(0.102)	(0.108)
Spring 1st						
Colocate	0.080	0.208	0.259	-0.012	0.083	0.141
	(0.068)	(0.152)	(0.142)	(0.079)	(0.107)	(0.106)
Mover	-0.064	0.131	-0.089	-0.069	-0.028	0.107
	(0.080)	(0.141)	(0.172)	(0.095)	(0.099)	(0.122)
Spring 2nd						
Colocate	0.019	-0.068	0.085	0.018	-0.015	0.013
	(0.064)	(0.131)	(0.140)	(0.076)	(0.105)	(0.103)
Mover	-0.059	0.003	0.063	-0.005	-0.002	0.068
	(0.075)	(0.135)	(0.145)	(0.089)	(0.095)	(0.116)
Spring 3rd						
Colocate	0.042	0.054	0.172	0.029	0.029	0.033
	(0.069)	(0.147)	(0.126)	(0.080)	(0.103)	(0.109)
Mover	-0.103	-0.171	-0.109	-0.016	-0.178	-0.132
	(0.076)	(0.136)	(0.137)	(0.090)	(0.095)	(0.111)
Externalizing Problem Behaviors						
Fall K.						
Colocate	-0.018	-0.111	-0.134	0.062	0.052	0.028
	(0.071)	(0.149)	(0.131)	(0.091)	(0.104)	(0.114)
Mover	-0.035	-0.142	**-0.315 ***	0.008	-0.001	-0.051
	(0.077)	(0.150)	**(0.148)**	(0.098)	(0.095)	(0.109)
Spring K.						
Colocate	-0.066	0.049	**-0.278 ***	-0.057	-0.025	0.061
	(0.064)	(0.128)	**(0.143)**	(0.080)	(0.095)	(0.105)
Mover	0.000	-0.143	-0.196	-0.033	0.032	0.009
	(0.073)	(0.126)	(0.182)	(0.089)	(0.092)	(0.107)
Spring 1st						
Colocate	-0.102	-0.220	-0.183	0.012	-0.046	-0.058
	(0.066)	(0.142)	(0.115)	(0.076)	(0.099)	(0.110)
Mover	0.082	-0.137	0.036	0.112	0.047	0.010
	(0.078)	(0.129)	(0.148)	(0.095)	(0.097)	(0.119)
Spring 2nd						
Colocate	-0.075	-0.045	-0.232	0.038	-0.133	0.050
	(0.071)	(0.132)	(0.139)	(0.090)	(0.102)	(0.111)
Mover	-0.020	-0.031	-0.240	-0.033	-0.016	-0.048
	(0.077)	(0.131)	(0.141)	(0.094)	(0.101)	(0.132)
Spring 3rd						
Colocate	0.007	0.026	0.121	0.048	-0.006	0.042
	(0.067)	(0.129)	(0.140)	(0.087)	(0.099)	(0.102)
Mover	-0.038	-0.040	-0.069	-0.055	-0.032	-0.145
	(0.068)	(0.128)	(0.129)	(0.082)	(0.090)	(0.095)

Table 5.5. Continued.

	Main Effects	Low-SES	State-Funded Pre-K	Urban	School AYP	High Transition Practices
Internalizing Problem Behaviors						
Fall K.						
Colocate	-0.062	0.053	0.063	-0.003	0.035	-0.101
	(0.071)	(0.162)	(0.123)	(0.095)	(0.107)	(0.096)
Mover	0.089	0.161	0.102	0.082	0.118	0.162
	(0.088)	(0.175)	(0.147)	(0.100)	(0.114)	(0.127)
Spring K.						
Colocate	-0.067	-0.105	-0.147	-0.012	-0.054	0.015
	(0.069)	(0.138)	(0.157)	(0.087)	(0.103)	(0.103)
Mover	**0.155 ***	0.170	0.149	0.150	0.173	0.116
	(0.087)	(0.151)	(0.199)	(0.110)	(0.107)	(0.125)
Spring 1st						
Colocate	-0.118	-0.216	-0.234	-0.023	0.020	-0.159
	(0.069)	(0.148)	(0.146)	(0.077)	(0.099)	(0.115)
Mover	0.091	-0.147	0.020	**0.205 ***	0.159	-0.010
	(0.086)	(0.163)	(0.169)	**(0.096)**	(0.103)	(0.138)
Spring 2nd						
Colocate	-0.029	0.027	**-0.350 ***	0.028	0.027	**-0.256 ***
	(0.071)	(0.150)	**(0.165)**	(0.081)	(0.117)	**(0.107)**
Mover	0.049	0.024	-0.156	0.060	0.095	-0.139
	(0.083)	(0.156)	(0.174)	(0.095)	(0.109)	(0.121)
Spring 3rd						
Colocate	-0.095	-0.134	-0.123	**-0.145 ***	-0.021	-0.127
	(0.066)	(0.143)	(0.145)	**(0.074)**	(0.094)	(0.111)
Mover	-0.036	0.070	0.010	0.065	-0.063	-0.104
	(0.072)	(0.157)	(0.143)	(0.082)	(0.092)	(0.118)

Qualitative Analysis Results

I present the findings from the qualitative portion of the study in a manner consistent with the interview protocol and analytic coding structure. First, I focus on general preferences for school-based or non-school-based Pre-K and the key benefits and challenges of each setting. Next, I focus on the extent to which school-based Pre-K facilitates collaboration and vertical alignment. Finally, I conclude with an overview of my findings related to the role of leadership at the principal and teacher levels. Within each of these topic areas, I also discuss how participant reports varied between teachers within schools, between schools within districts, and between districts within the state, where notable.

Benefits and Challenges of School-Based Pre-K

In Table 5.6, I provide a summary of the key benefits and challenges of school-based Pre-K, as identified by the interview participants. While the list in this table is not exhaustive, I chose to highlight the benefits and challenges most frequently raised by participants and that came from a diversity of participants.

By far, the most frequently cited benefit of school-based Pre-K is what I term the "normalization of school" for school-based Pre-K attenders. Because these students spend their Pre-K year inside of an elementary school building, they become familiar with the school building and the routines and practices in elementary school. Participants gave specific examples of children learning the layout of the school campus and learning how to go through the line at lunch, for example. As one principal noted, for the kids in Pre-K in the school, "they simply know what kindergarten is like because it's down the hall and they have been exposed to the environment for a year." This benefit of the normalization of school is closely related to the second key benefit of school-based Pre-K: Easier transitions into kindergarten.

Participants often framed their discussion about the normalization of school in terms of how it facilitates easier transitions into kindergarten, particularly for co-locaters. Here, participants cited how children who attended school-based Pre-K often had higher levels of "school behaviors", including the ability to walk in a line, for example. Furthermore, some participants reported that the transition was particularly easy for co-locaters. Beyond acquisition of concrete "school behaviors" that could be applied to any kindergarten context, co-locaters benefit from learning about the staff and specific school culture. One respondent, for example, noted how she sees the children who attended Pre-K in their school building enter kindergarten

with much more confidence because they already feel like they are a part of the school and have established relationships with staff members throughout the school.

In addition to the "normalization" of school and its links to facilitating easier transitions into kindergarten, respondents also cited a number of practical resource benefits to locating Pre-K in elementary school buildings. I learned from state- and county-level respondents that Pre-K teachers in district schools are generally paid more than Pre-K teachers in center-based settings. Further, and likely relatedly, the teachers in school-based settings have higher credentials (degree levels) and there is less staff turnover. From the school-level respondents, I learned that Pre-K programs located in school buildings benefit from the broader array of resources that elementary school buildings offer. For example, one Pre-K teacher discussed how the Pre-K students were able to use the school library, art room, and computer lab—resources that are not frequently available in smaller, stand-alone centers.

The final key benefit of school-based Pre-K that emerged from my interviews is the opportunities it provides for collaboration and vertical alignment. While I will detail my findings as they relate to collaboration and alignment in a subsequent section, respondents frequently noted how having Pre-K and kindergarten teachers in the same building set the stage for collaboration around curriculum and instruction as well as simply learning about each other's students and the needs that they have. In one school in Case county, a kindergarten teacher talked about how she was "just down the hall" from the Pre-K classroom and how she would often go eat lunch in the Pre-K classroom to "get to know the kids that would be in my class next year".

Table 5.6. Key Benefits and Challenges of School-Based Pre-K

Key Benefits of School-Based Pre-K	Key Challenges of School-Based Pre-K
"Normalization" of school.	Teacher isolation.
Easier transition into kindergarten.	Risk of developmentally inappropriate classroom activities.
Additional resources.	Physical space not appropriate for Pre-K.
Opportunities for collaboration/alignment.	Mixing NC Pre-K and K-12 requirements and regulations.

In addition to benefits of school-based Pre-K, respondents also highlighted a number of key challenges associated with locating Pre-K in elementary school buildings. The first key challenge is teacher isolation. In most of the elementary schools with Pre-K programs that I visited, there was only one Pre-K classroom. As a result, it is difficult for teachers to have same-grade peer networks to support them in their practice. Conversely, in center-based settings, there are often multiple teachers who work closely with one another on a daily basis. An interesting change related to this issue was taking place in Case county. Due to the lack of Pre-K teacher professional community *and* practical space constraints in the district elementary schools, the district plans to open a singular "Pre-K Center" where all children will attend Pre-K.

Another common challenge raised by respondents is the risk of developmentally inappropriate classroom practices in Pre-K because of location in elementary school buildings. These concerns were almost universally raised by Pre-K teachers located in elementary school buildings. These teachers often characterized the differences between Pre-K and kindergarten to be vast—with Pre-K focused on social and emotional development in a play-based manner and kindergarten focused on academic development in a more structured, didactic manner. These Pre-K teachers in elementary school settings often felt pressures from higher grade levels, most

often kindergarten teachers, to give kids a jump start on early academic skills. These Pre-K teachers would often push back. For example, one Pre-K teacher reported having to tell kindergarten teachers, "This is what studies have shown. This is what is developmentally appropriate for four-year olds" and added that she would "just put it back on the research and what my standards are."

It is important to note, however, that not all reports of this "academic-developmental" debate were conflictual and negative. In one school, for example, the Pre-K teacher did feel that she was sometimes pushed to engage in developmentally inappropriate classroom activities, but in these instances, she would work with the kindergarten teacher to work through the issue. For example, this respondent described a specific situation where the kindergarten teacher wanted Pre-K to focus on letter identification and suggested some methods for doing so. These methods, as determined by the Pre-K teacher were developmentally inappropriate. However, the two teachers worked together to develop ways to increase coverage of letter identification in Pre-K using appropriate methods. This same Pre-K teacher also has worked to inform her principal about developmentally appropriate classroom practice. Since the elementary school principals observe and evaluate Pre-K teachers in their buildings, she felt it was necessary to provide the principal with the early learning standards for Pre-K and help her realize that effective instruction in Pre-K may look very different from effective instruction in higher grades. She did this because she feels she has to advocate for Pre-K and "look after herself"—a common theme that I will detail in later sections.

Another key challenge with school-based Pre-K is that the physical building is sometimes not suitable for Pre-K children. For example, Pre-K students often have to eat in their classroom because the tables in the cafeteria are not appropriately sized for four-year-olds. Another

example is the playground equipment. One school reported having to install a separate playground that is only used by the Pre-K students because it meets the NC Pre-K regulations. These concerns were not universal, however. One school in Case county was only a couple of years old and was actually built with Pre-K in mind. The Pre-K teacher and principal in this school both reported how the construction included appropriately sized play equipment, and in the Pre-K classroom, appropriately sized counters and bathroom fixtures.

Concerns over the appropriateness of elementary school buildings for Pre-K are closely related to the final key concern that participants identified—mixing NC Pre-K and K-12 requirements and regulations. Despite some NC Pre-K programs being located in elementary school buildings, NC Pre-K classrooms are separate entities from a regulatory perspective. Participants highlighted a range of instances where these divisions manifested themselves, in addition to the space concerns previously detailed. One example includes coordinating the Pre-K and K-5 school schedules. Not only does Pre-K start and end at different times in the calendar year, the daily schedule in terms of start and end times differs between the two. These differences were cited as an obstacle to making the Pre-K program feel like they are a full part of the broader elementary school and led to instances where the Pre-K program was left out of school activities, such as assemblies.

Collaboration and Vertical Alignment

A major focus of my interviews was the extent to which school-based Pre-K enables collaboration and vertical alignment. My analysis of the coded data revealed two primary types of collaboration. First, there was formal collaboration in the form of kindergarten transition practices and data sharing. Second, there was informal collaboration in the form of social networks, often facilitated by proximity and teacher/principal initiative. A final key finding, as it

relates to collaboration, is that there is significant variability in the extent of collaboration *between* schools. I detail my findings for each of these key results below.

Formal Collaboration: Transition Practices and Data Sharing

Nearly all school- and center-level respondents reported using some form of kindergarten transition practices. The most common practices included orientation nights and having children visit kindergarten prior to the start of the school year. In the school-based settings, the transition practices often included the Pre-K program and started during the Pre-K year. For example, one school had a "Moving-up Day" wherein Pre-K students would spend the day in kindergarten, learning about the various ways their routines would change for the next academic year. None of the center-based Pre-K respondents reported any types of transition practices where Pre-K children visited or experienced kindergarten in some way prior to the start of kindergarten.

Another formal collaboration practice was the sharing of data between Pre-K and kindergarten. In nearly all school-based Pre-K programs and one center-based Pre-K program, respondents reported developing a portfolio that could be shared with kindergarten teachers. These portfolios often included examples of student work, teacher notes, and results from formative assessments systems. There was, however, significant variability in the extent to which respondents viewed the utility of the portfolios and how they were shared. Some Pre-K teachers were hesitant that kindergarten teachers would look at the data and some kindergarten teachers even acknowledged that they would rather form their own assessments of incoming students than take the word of Pre-K teachers. In other cases, particularly in schools with collaborative cultures, Pre-K and kindergarten teachers would meet to discuss the data in the portfolios and discuss the needs of individual children and plan for kindergarten classroom placements.

A final area that I probed in interviews related to formal collaboration was the extent to which there were additional transition practices or data sharing for children who co-locate. My questions focused on if children who attended Pre-K in a school and stayed in the same school for kindergarten were provided any type of additional support or contact during their transition. I did not find significant evidence of additional supports for this subgroup of students. Schools engaged in the same types of practices regardless of whether or not a child would be returning to the same school for kindergarten. The only instance where there was any difference was when Pre-K and kindergarten teachers discussed classroom placements for co-locating children.

Informal Collaboration: Social Networks

In addition to formal transition practices and sharing of data, a significant amount of collaboration that took place in schools was informal in nature. This collaboration generally took place in the form of social networks among Pre-K and kindergarten teachers. I found few instances of relationships with principals and teachers in higher grades. One of the key facilitators of these relationships was proximity to one another in the same hallway. As mentioned previously, one kindergarten teacher in Case county reported frequently visiting the Pre-K classroom to have lunch with the students and get to know many of the students she would have next year. This respondent talked about how establishing these relationships with the students helps them adjust to kindergarten.

Respondents also talked about how their informal collaborations focused on the content of instruction. In one school in Case county, Pre-K and kindergarten teachers talked about how they coordinate instruction so that content is aligned well between Pre-K and kindergarten. A specific example of this is coordination over letter instruction in Pre-K and kindergarten. A kindergarten teacher noted,

> In Pre-K they used to learn only capital letters and then when they got to me they only wanted to write in capital letters. It was a problem. We [kindergarten teachers] talked to the Pre-K teacher and asked her to start introducing the lower case letters once the kids learned the lower case letters. We definitely noticed the change. The students are much more open to doing lower case letters now.

This kindergarten teacher further described how this form of communication is always friendly in nature and "goes both ways". In one case, this kindergarten teacher asked the Pre-K teacher to work on a certain skill but the Pre-K teacher explained that such content was not "part of her standards" and was inappropriate for Pre-K.

Even in some school contexts where the Pre-K teacher does not feel like a full part of the elementary school, they may still have close connections with the kindergarten team. For example, one Pre-K teacher in Pike county reported that she would chat with the kindergarten teachers on a daily basis since they share a hallway in order to "discuss what is going on in the classroom." This same respondent reported feeling like Pre-K was not included as a part of the elementary school and feels very little support from the administration.

Variability in Collaboration between Schools

I found that the extent of formal and informal collaboration varied significantly and the variability was generally *between* schools. My hierarchical analysis of the data revealed consistent findings when analyzing data between participants within schools, but the findings differed when looking between schools. Further, the school-by-school findings were not clustered together by county—suggesting a limited role of county/district effects on the extent of collaboration and vertical alignment. In order to illustrate the extent of school-by-school

variation, I provide a brief narrative description of two separate elementary schools in Case county.

The first school was a highly collaborative school. In this school, the Pre-K teacher felt like the Pre-K classroom was a full part of the school. She collaborates regularly with the kindergarten teacher to coordinate instruction and ensure the transition from Pre-K to kindergarten is as seamless as possible. In fact, the Pre-K and kindergarten classrooms use a common early literacy curriculum. The principal at this school is also very involved in fostering vertical alignment. She includes the Pre-K classroom in as many school-wide events as possible and includes the Pre-K teacher in staff meetings and "vertical" PLC meetings. When asked if the Pre-K classroom was just another grade at the school, the principal replied, "Yes, 100 percent."

The second school was not a collaborative school. In this school, the Pre-K program is essentially renting space in the elementary school building, with extremely limited contact and interaction with the broader elementary school. When speaking with a kindergarten teacher in this school about the extent to which she collaborates with the Pre-K teacher, she replied, "I never see the Pre-K kids, they stay at their end of the hall". The Pre-K teacher had a similar response that indicated a minimal amount of interaction. The principal also acknowledged that there was limited collaboration and that they "needed to do better."

These two schools clearly represent the wide variability in the extent to which elementary schools with Pre-K programs have collaborative, vertically aligned cultures. An important and somewhat surprising finding was that this variability existed between schools, with very limited differences between counties. At the county-level, I found very few instances of efforts to facilitate collaboration. In fact, one county-level respondent acknowledged herself how it varies dramatically from school to school. She could easily list off schools that were "doing it

[collaboration and vertical alignment] well" and those that were not. Due to the lack of many mandated policies and procedures related to vertical alignment and collaboration, many respondents attributed the variability to individual initiative and leadership, by both teachers and principals.

Leadership

The final set of key findings from the qualitative portion of this dissertation focus on leadership. In this section, I organize my findings into two key forms of leadership: principal leadership and teacher leadership.

Principal Leadership

Given the aforementioned variability in collaboration and vertical alignment between schools, the role of school principals as leaders of schools emerged as an important foci of my interviews. As mentioned previously, the extent to which principals were engaged or not engaged with the Pre-K programs in their buildings varied extensively. When probing for possible explanations for this variability, two key findings emerged.

One of the key factors related to principal support of Pre-K and fostering a collaborative culture was their personal views related to the value of Pre-K and early education, generally. I found that principals who actively engaged with the Pre-K program cited the importance of early childhood education and the benefits that Pre-K provides in terms of preparing students for kindergarten. These principals would frame having Pre-K in their schools as an opportunity to be taken advantage of and something that will help students become more successful throughout elementary school.

The second key factor related to principal support of Pre-K was the extent to which principals viewed the administrative tasks associated with Pre-K as burdensome. In the school contexts with low levels of collaboration, principals often cited the bureaucratic responsibilities associated with Pre-K programs. Elementary school principals are required to observe Pre-K classrooms (in addition to the county-level Pre-K administrator), ensure that childcare regulations are met, and cover Pre-K teacher absences with substitutes. As one principal in Pike county put it, "We have to do all of these things and we don't get anything for it. They don't count in our enrollment counts." This principal felt like there were additional responsibilities associated with having the Pre-K program in her building with limited benefits or compensation.

A final finding related to principal leadership is related to the academic-developmental debate. Recall that one of the key challenges identified by respondents associated with school-based Pre-K is the potential for developmentally inappropriate influences to permeate into Pre-K. Many respondents, ranging from state-level officials to Pre-K teachers identified elementary school principals as key to this issue. In many cases, respondents characterized elementary school principals as lacking appropriate knowledge of developmentally appropriate practice and play-based learning, which is the norm in early education settings.

For example, one Pre-K teacher in an elementary school in Olin county discussed how she has a conflictual relationship with her principal over the proper nature of instruction in Pre-K. As she puts it, the principal "wants to throw our curriculum out" since the principal does not see it as adequately preparing students for the rigors of kindergarten. This Pre-K teacher has over 20 years of experience and reported feeling like the principal is extremely dismissive of her experiences and judgement as an early education professional. When I asked her if she would prefer to teach in a center-based setting, she replied "yes" because she is always getting pressure

to focus on “what is coming” in higher grades, despite her strong convictions that such instruction is not appropriate for Pre-K.

Teacher Leadership

In addition to principal leadership, teacher leadership, particularly on the part of Pre-K teachers in elementary schools, emerged as a prominent theme. More specifically, I found that Pre-K teachers in highly collaborative school settings would actively advocate for the inclusion of the Pre-K program in the elementary school. As mentioned previously, variability in the extent to which schools had collaborative, vertically aligned cultures existed between schools. My findings indicate that a combination of principal and teacher leadership are a key explanatory factor. On the part of teachers, this came in the form of teachers pushing elementary school principals to ensure their inclusion in school activities and educating principals on best practices in early education.

Pre-K teachers that I spoke to often cited how they had to take it upon themselves to ensure that their Pre-K class was consistently included in school activities. This may include school assemblies, field trips, and other activities that included multiple grades of the school at once. One Pike county Pre-K teacher noted that Pre-K’s presence in the school has improved because she keeps pushing on the administration and reminding them that, “We are in this school. We are a part of it.” Beyond advocating for inclusion is specific activities, Pre-K teachers also reported pushing for symbolic forms of inclusion. For example, one Pre-K teacher in Olin county talked about how school emails and other forms of communication would often reference “K-5” and not “PK-5”. Small messages like this reinforced the notion that Pre-K was a separate entity in the school and was a point of frustration for Pre-K teachers. This teacher raised her

concerns with the administration and she reported that there have been subsequent efforts for the school to be more inclusive in its messaging.

Merged Analysis Results

As a final analytic step in this concurrent mixed-methods study, I examined the ways in which the findings from the quantitative and qualitative analysis related to one another. In many ways, I was surprised by how little convergence existed between the two modes of analysis. While the quantitative analysis provided some hints at benefits of co-location, the findings were not clearly defined. Conversely, the qualitative analysis provided a clearer, positive portrait of the promises of school-based Pre-K, with many respondents noting their preference for locating Pre-K in schools for all of the aforementioned potential benefits to children.

A specific example of this discordance is with the Approaches to Learning findings from the quantitative analysis and the "Normalizing School" findings from the qualitative analysis. The Approaches to Learning measure captures the extent to which children exhibit school behaviors, such as following instructions, adjusting to routines, and holding a pencil appropriately, for example. Given how the most commonly cited benefit of school-based Pre-K identified in my interviews was how it provided children with these skills through the process of "Normalizing School," it was unexpected that there were no significant relationships found quantitatively.

Another area of discordance is related to the transition to kindergarten. A consistent theme from the qualitative analysis is that school-based Pre-K facilitates easier transitions to kindergarten for children. Kindergarten teacher respondents, for example, frequently cited how they could see significant differences between children who attended Pre-K in a school and those that did not, in terms of their readiness for kindergarten. This evidence suggests that I would

observe benefits for school-based Pre-K attenders at kindergarten entry, on a range of outcomes. However, as my quantitative analysis revealed, I found no significant relationships between school-based Pre-K attendance.

It is important to note that part of why I find few areas of concordance between the quantitative and qualitative analysis is that the two approaches probed different elements of the phenomena of school-based Pre-K. For example, I did not ask interview participants about their views on school-based Pre-K for the specific subgroups of students analyzed in the quantitative analysis. My focus in the quantitative analysis was to relate school-based Pre-K and co-location to student outcomes while my focus in the qualitative analysis was to understand the practical on-the-ground realities of locating Pre-K programs within elementary school buildings. Another potential reason for the discordance, which I detail in the Discussion, is that the quantitative data come from a national sample while the qualitative data come only from North Carolina—potentially masking the nuance of state-by-state variability in Pre-K programming as it relates to school-based Pre-K.

CHAPTER 6: DISCUSSION

Pre-K programs have been located in school-based settings for decades, however, very few studies have directly examined the role that the physical location of programs has on student outcomes (Magnuson et al., 2007). This dissertation provides the most comprehensive analysis of the relative effectiveness of school-based versus non-school-based Pre-K and the practical realities of locating Pre-K in schools to date. In this Discussion, I begin with a brief review of my key findings and relate those findings to existing literature. Next, I discuss the implications of my findings for policy and practice. Finally, I discuss the limitations of my study and the ways that future research can address them.

Discussion of Findings and Existing Literature

Quantitative Findings

From the quantitative analysis, the primary findings are threefold. The first key finding is that I find no differences between school-based and non-school-based Pre-K attenders in terms of their outcomes in grades K-3. However, when examining the second specification of the independent variable, which divides school-based Pre-K attendees into co-locaters and movers, I find suggestive evidence that co-location is beneficial and moving is harmful, especially for math and reading achievement. My original hypothesis was that there would be benefits for school-based Pre-K attenders on all outcomes and that these benefits would be even larger for the subset of students who co-located. While the results are not as expected for the school-based Pre-K attendance variables, the results provide some evidence for a co-location benefit.

This co-location finding is roughly consistent with research from Magnuson and colleagues (2007), which found no outcome differences between Pre-K attenders in school-based or center-based settings, but did find benefits for co-locaters in terms of social-emotional learning. As detailed in the literature review, this study is not directly comparable to the present study as the contrast for the co-locaters in the Magnuson et al. (2007) study included school-based Pre-K attenders that did not co-locate. Nonetheless, there is some concordance between my findings for co-location and previous research on the topic.

The second key finding from the quantitative analysis is that I find consistent negative associations for children who live in urban areas in terms of academic achievement, in particular. It is possible that this finding relates to differences in the childcare markets in large, urban areas. For example, center-based programs (the counterfactual) may be more prevalent and high quality, dampening any potential benefit that school-based settings may offer. Relatedly, if the process of moving between Pre-K and kindergarten is detrimental for student outcomes, dynamic markets in urban settings with more moving may be driving these findings. That said, these two explanations are hypothetical and future research should more fully probe the reasons why these associations were consistently negative for students living in urban areas.

The third key finding from the quantitative analysis is that I find limited evidence that measures of school quality and vertical alignment play an important role in explaining the persistent negative achievement findings for movers in the second specification of the independent variable. Recall that the negative achievement findings for the movers in math and reading achievement became more negative over time and were largest at the end of third grade. These findings were robust to alternative specifications of the model that included controls for the total number of school switches a child had after entering kindergarten, with the assumption

that children who are movers are really just transient students. Consistent with the sustaining environments hypothesis detailed in the literature review (Bailey et al., 2017), I hypothesized that subsequent school quality or the extent to which schools engage in vertical alignment may moderate this observed relationship. However, I do not find support for this with the AYP and transition practices models. As I discuss in the limitations section, these measures are relatively crude proxies of school quality and vertical alignment, so future research should seek to better measure these constructs.

Qualitative Findings

The qualitative portion of this dissertation provides important insights into the practical, on-the-ground realities of locating Pre-K programs in schools. The central finding from this analysis is that locating Pre-K programs in elementary schools is in and of itself not enough to ensure high-quality, aligned instruction. Rather, schools vary in the extent to which they have collaborative, vertically aligned cultures. I also find that one of the key explanatory factors in this variation between schools is the role of principal and teacher leadership. From a policy perspective, it is disappointing that location alone is not more predictive, given its clear-cut implications for policy design (i.e., locate Pre-K programs in elementary schools). That said, it is not surprising that schools vary dramatically in their cultures, leadership, and organization—factors that have been shown to be critical for improvement (Bryk et al., 2010).

The most direct existing research that has examined the practical realities of school-based Pre-K is from Desimone et al. (2004), which focused on the implementation of the Schools of the 21st Century Pre-K model in elementary schools. A number of my findings are corroborated by the Desimone et al. (2004) study, including benefits of opportunities for coordinating instruction and easier Pre-K to kindergarten transitions as well as challenges of feelings of isolation and

developmentally-inappropriate influences from higher grades. That said, I did not find any evidence of challenges of conflicts over material resources in the school (e.g., copier machine) and inequitable pay scales between Pre-K and elementary school teachers, as Desimone et al. (2004) did. In terms of the latter, this is likely because North Carolina Pre-K teachers who teach in public schools are paid on the district salary schedule. It is important to note that the Desimone et al. (2004) study took place in schools where Pre-K programs were *just integrated* into the schools. In many of the schools I conducted interviews in, the Pre-K programs have been a part of the schools for years.

Like my study, interview respondents in the Desimone et al. (2004) study also reported that school-based Pre-K was beneficial for "normalizing school," wherein students gain advantages in learning school behaviors (e.g. walking in a line, following instructions). In my analysis, this was the single-most cited benefit of school-based Pre-K. The skills that respondents listed that characterize this domain are largely focused on obedience. While routines and skills related to an orderly educational environment are important, it does raise the question of if these skills come at the expense of fostering inquiry and social interaction among peers.

Another area of alignment between the qualitative analysis and existing research is related to the finding that school-based Pre-K may facilitate easier transitions between Pre-K and kindergarten and vertical data sharing. Work from the IES Early Learning Network, taking place in North Carolina, has reported similar findings (Cohen-Vogel et al., 2018). Specifically, this work has highlighted the importance of school-based Pre-K in making children familiar with elementary school procedures and developing schooling-related behaviors, which make the transition to kindergarten easier. The Early Learning Network research has also found evidence of vertical data sharing through the use of portfolios and "vertical meetings" between Pre-K and

kindergarten teachers to discuss the needs of individual children and classroom placements. Note that my study and this research took place during similar time spans, in similar locales in North Carolina, and probed similar research questions.

Finally, respondents in this study consistently cited benefits of school-based Pre-K in terms of human capital. Specifically, respondents noted that teachers in school-based settings are paid more than those in center-based settings and that the staff turnover rates in school-based settings are lower than in center-based settings. These findings are corroborated by evidence from both North Carolina (Peisner-Feinberg et al., 2016) and nationally (Bellm et al., 2002). In sum, I find much more concordance between my qualitative analysis and the existing literature than for the quantitative analysis.

Implications for Policy and Practice

While the results of this study are preliminary, there are a number of implications of the findings for policy and practice. Many of these implications stem from the qualitative section of the study, which illuminated practical, on-the-ground issues associated with school-based Pre-K that changes in policy and practice may ameliorate.

First, there should be a focus on delivering professional development to administrators and teachers in school-based Pre-K settings that focuses on creating vertically aligned and developmentally appropriate P-3 education settings. As my findings revealed, there was significant variability in the extent to which schools had these collaborative and aligned environments. Professional development supports and other formal requirements (e.g., requiring Pre-K teacher involvement in kindergarten professional learning communities) may overcome this variability and ensure that more schools are taking advantage of the opportunities afforded by having Pre-K programs located in elementary school buildings. Relatedly, this professional

development should ensure that all actors involved in the collaboration are informed about what developmentally appropriate practices look like for each grade level, ensuring that Pre-K programs are not pushed to use developmentally inappropriate classroom practices—a common concern raised by school-based Pre-K teachers in my interviews.

Second, school-based Pre-K settings should staff more than a single Pre-K classroom, where possible. My interviews revealed that Pre-K teachers often felt isolated as a result of the lack of same-grade teacher peers. Even in the collaborative school environments, collaboration was often between Pre-K and kindergarten teachers because there was only one Pre-K teacher. If policymakers choose to increase Pre-K programming in school-based settings, having multiple Pre-K classrooms may naturally occur. If increasing the number of Pre-K classrooms is impractical given space or other constraints, efforts should be made to provide opportunities for Pre-K teachers within school districts to develop rich networks where they can discuss issues of practice—even if they are located in separate school buildings.

Third, policymakers should work to break down the logistical divisions that exist between Pre-K and K-12 educational systems. There are two specific examples of challenges that were raised in the interviews that should be addressed. First, Pre-K programs located in elementary school buildings should operate on the same schedule as the elementary school so that Pre-K programs are more fully integrated in the daily routines and programmatic offerings of the school. Second, data systems should link seamlessly from Pre-K to kindergarten. Data collected from Pre-K teachers (e.g., developmental screening data) may be a valuable resource to kindergarten teachers and systems that link kindergarten to higher grades should also link backwards to Pre-K. More generally, policymakers should examine the various rules and regulations that govern Pre-K and K-12 education and evaluate how they can streamline and

coordinate them to ease the burden on elementary school leaders in settings that house Pre-K programs.

Limitations and Directions for Future Research

There are important limitations of my dissertation that future research should address. One key limitation is the concordance between the quantitative and qualitative data sources. The quantitative data comes from a nationally representative sample of children who were kindergarteners in 2010-11. The qualitative data comes from interviews with a convenience sample of respondents in central North Carolina. Ideally, I would have either quantitative data from North Carolina or qualitative data on a national scale. It is possible that the North Carolina context is unique and my qualitative findings do not align well to the quantitative findings for this reason. Unfortunately, administrative data from North Carolina that links Pre-K to elementary school outcomes were not available at the time of this study and conducting qualitative analysis on a national scale was impractical. As states continue to develop their longitudinal data systems that link early education and K-12 data sources, future research should use mixed methods approaches that have greater concordance between the quantitative and qualitative data sources, in terms of context.

A limitation of the quantitative analysis is that the findings remain correlational, not causal, in nature. Despite using sophisticated estimation strategies aimed at removing omitted variable bias, in the absence of random assignment, this threat to internal validity remains. While my findings are suggestive of what causal relationships may be, it would be inappropriate to term them as causal effects. That said, it is unlikely that children will be randomly assigned to school-based or non-school-based Pre-K, so future researchers should use a range of quasi-experimental

designs to generate a more robust body of evidence that could point to the causal effects of school-based Pre-K.

Another limitation of this work is the lack of effective measures of P-3 vertical alignment and collaboration in the ECLS-K. As my qualitative analysis revealed, location of Pre-K programs in elementary schools is likely in and of itself not sufficient to ensure high-quality programming. Rather, schools vary in the extent to which they take advantage of Pre-K programs and develop collaborative, vertically aligned cultures. While I used a measure of schools' use of kindergarten transition practices as a rough proxy for vertical alignment, better measurement of this phenomenon would be helpful in testing if it moderates the relationship between school-based Pre-K and student outcomes.

A final limitation of this research is that I was the sole investigator. While this is necessary in the completion of a dissertation, it is best practice, particularly for the qualitative analysis, to work in collaborative teams. It is possible that I have interpreted thematic findings or characterized the tone of participant responses in a way that another individual may not have. Ideally, I would have worked with other individuals on this study, continually meeting to discuss emerging findings and areas of potential disagreement.

Appendix A: Balance Statistics

Appendix A: Balance Statistics

	Treatment: Pre-Weight			Control: Pre-Weight		
	Mean	**Variance**	**Skewness**	**Mean**	**Variance**	**Skewness**
White	0.42	0.24	0.32	0.50	0.25	-0.01
Black	0.19	0.16	1.55	0.18	0.15	1.66
Hispanic	0.26	0.19	1.07	0.16	0.14	1.83
Asian	0.06	0.05	3.80	0.08	0.08	3.00
Male	0.54	0.25	-0.14	0.52	0.25	-0.07
SES	-0.13	0.60	0.42	0.16	0.66	0.16
Siblings	1.43	1.22	1.20	1.24	1.01	1.13
Two Parent	0.69	0.21	-0.83	0.71	0.20	-0.94
Age K. Enter	66.03	25.22	-1.18	65.96	21.30	-0.59
Northeast	0.13	0.12	2.17	0.17	0.14	1.73
Midwest	0.56	0.25	-0.24	0.44	0.25	0.22
South	0.14	0.12	2.08	0.20	0.16	1.53
Non-English	0.16	0.13	1.89	0.10	0.09	2.76
Food Secure	0.93	4.93	2.59	0.70	3.71	2.73
Parent Age	33.69	51.20	0.93	34.66	49.93	0.78
Birth Weight	114.70	498.70	-0.65	116.30	462.20	-0.57
WIC	0.53	0.25	-0.12	0.41	0.24	0.38
Dist. Poverty	22.99	128.80	0.34	19.38	108.50	0.51
Urban	0.64	0.23	-0.58	0.73	0.20	-1.01
Married @ Birth	0.58	0.24	-0.33	0.65	0.23	-0.62
	Treatment: Post-Weight			**Control: Post-Weight**		
White	0.42	0.24	0.32	0.42	0.24	0.32
Black	0.19	0.16	1.55	0.19	0.16	1.55
Hispanic	0.26	0.19	1.07	0.26	0.19	1.07
Asian	0.06	0.05	3.80	0.06	0.05	3.80
Male	0.54	0.25	-0.14	0.54	0.25	-0.14
SES	-0.13	0.60	0.42	-0.13	0.60	0.42
Siblings	1.43	1.22	1.20	1.43	1.22	1.20
Two Parent	0.69	0.21	-0.83	0.69	0.21	-0.83
Age K. Enter	66.03	25.22	-1.18	66.02	25.21	-1.18
Northeast	0.13	0.12	2.17	0.13	0.12	2.16
Midwest	0.56	0.25	-0.24	0.56	0.25	-0.24
South	0.14	0.12	2.08	0.14	0.12	2.08
Non-English	0.16	0.13	1.89	0.16	0.13	1.89
Food Secure	0.93	4.93	2.59	0.92	4.93	2.59
Parent Age	33.69	51.20	0.93	33.69	51.19	0.93
Birth Weight	114.70	498.70	-0.65	114.70	498.60	-0.64
WIC	0.53	0.25	-0.12	0.53	0.25	-0.12
Dist. Poverty	22.99	128.80	0.34	22.98	128.80	0.34
Urban	0.64	0.23	-0.58	0.64	0.23	-0.58
Married @ Birth	0.58	0.24	-0.33	0.58	0.24	-0.33

Appendix B: Post Interaction Form (PIF)

To be completed electronically after each interview.

Participant ID:
Date:

1. **What were the most notable take-aways from this interaction?**

2. **What information did the participant provide about differences in school-based versus non-school-based Pre-K settings?**

3. **What information did the participant provide about co-location?**

4. **Did this interaction reveal any new items that should be added to the protocol?**

5. **Did this interaction reveal the need for modification or elimination of any items from the protocol?**

6. **Was there any information provided in the interview that I need to follow-up on?**

Appendix C: Interview Guide

[Create a warm and friendly environment; give interviewee consent form.]

Thank you for taking the time to meet with me today. My name is Michael Little and I am with the University of North Carolina at Chapel Hill. Our meeting today is part of a study about different types of Pre-K settings and the transition of children from Pre-K into elementary school.

I asked to meet with you today because I value your insight and expertise.

I have several categories of questions. I understand your time is valuable, and have planned 30 minutes for this discussion *once we get started*.

This interview will be confidential. I will audio-tape this discussion to accurately capture what you say. The digital audio recording will go straight to UNC and will not be shared with anyone. Your name will not be shared in any documentation or reports that come out of this discussion. At any time, you are welcome to ask me to turn off the recorder.

Do I have your permission to proceed? If so, please read and sign the consent form in front of you.

*If yes:

PRE-K TEACHER QUESTIONS

First, I would like you ask you some questions about your experiences as a Pre-K teacher.

1. How many years have you been a Pre-K teacher at [NAME OF SCHOOL/CENTER?]

2. How many years have you been a Pre-K teacher in total?

3. Have you ever taught in a Pre-K program located in a [CENTER/ELEMENTARY SCHOOL]. (Note: Say the setting that the teacher is NOT currently located in.)

Now I would like to shift to discuss Pre-K at [NAME OF SCHOOL/CENTER].

4. How many Pre-K classrooms are in [NAME OF SCHOOL/CENTER]?

5. How many students are in each Pre-K classroom?

As you may know, NC Pre-K classrooms can be located in either stand-alone centers or inside of elementary schools. My next set of questions focuses on the potential differences in children's Pre-K experiences based on location.

6. Do you think there are any differences in the Pre-K experience between center-based and school-based Pre-K programs?

   Probe: What are the key benefits of each?

   Probe: What are the key challenges of each?

7. Do you think there are any differences in the transition from Pre-K to elementary school between center-based and school-based Pre-K programs?

   Probe: What are the key benefits of each?

   Probe: What are the key challenges of each?

I'd like to now turn back to discussing your experiences and opinions regarding [NAME OF SCHOOL/CENTER].

8. To what extent to you think that Pre-K in [NAME of SCHOOL/CENTER] prepares children to succeed in kindergarten?

9. To what extent do you see alignment between Pre-K at [NAME of SCHOOL/CENTER] and elementary school?

10. Do you coordinate with kindergarten teachers in order to align instruction and facilitate the transition to kindergarten?

    If yes: In what ways?

    If no: Are there specific reasons why you do not?

11. Are there any specific transition activities that [NAME OF SCHOOL/CENTER] provides to help ease the transition from Pre-K to kindergarten?

12. What data on children, if any, are shared between Pre-K and kindergarten?

    Probe: Does data sharing vary based on where the child goes to kindergarten?

That concludes my prepared questions for this interview. Is there anything else you would like to share with me about Pre-K at [NAME OF SCHOOL/CENTER]?

I sincerely appreciate you taking the time to speak with me today. Your insights are extremely valuable to me and will help inform our study.

K-3 TEACHER QUESTIONS

First, I would like you ask you some questions about your experiences as an elementary school teacher.

1. How many years have you been a teacher at [NAME OF SCHOOL?]

2. How many years have you been a teacher in total?

3. Have you ever taught in an elementary school with/without a Pre-K program? (Note: Say "without" if school has Pre-K, "with" if school does not have a Pre-K.)

*If elementary school does not have Pre-K classrooms:

Now I would like to talk about the transition of children into elementary school.

4. Does [SCHOOL NAME] offer any transition activities to help ease children's transition into kindergarten?

5. What data, if any, do you receive about children when they enter kindergarten?

   Probe: Do you receive any additional information for children who attended NC Pre-K?

*If elementary school has Pre-K classrooms:

As you may know, NC Pre-K classrooms can be located in either stand-alone centers or inside of elementary schools. My next set of questions focuses on the potential differences in children's Pre-K experiences based on location.

6. Do you think there are any differences in the Pre-K experience between center-based and school-based Pre-K programs?

   Probe: What are the key benefits of each?

   Probe: What are the key challenges of each?

7. Do you think there are any differences in the transition from Pre-K to elementary school between center-based and school-based Pre-K programs?

   Probe: What are the key benefits of each?

   Probe: What are the key challenges of each?

I'd like to now turn back to discussing your experiences and opinions regarding [NAME OF SCHOOL].

8. To what extent to you think that Pre-K in [NAME of SCHOOL] prepares children to succeed in elementary school?

9. To what extent do you see alignment between Pre-K at [NAME of SCHOOL] and elementary school?

10. Do you coordinate with Pre-K teachers in order to align instruction and facilitate the transition to kindergarten?

    If yes: In what ways?

    If no: Are there specific reasons why you do not?

11. What proportion of children who attended Pre-K in [NAME OF SCHOOL] stay in the school for kindergarten?

12. Are there any specific transition activities that [NAME OF SCHOOL] provides to help ease the transition from Pre-K to kindergarten?

13. What data on children, if any, are shared between Pre-K and kindergarten?

    Probe: Are data only shared if children remain in [NAME OF SCHOOL]?

That concludes my prepared questions for this interview. Is there anything else you would like to share with me?

I sincerely appreciate you taking the time to speak with me today. Your insights are extremely valuable to me and will help inform our study.

SCHOOL/CENTER ADMINISTRATOR QUESTIONS

First, I would like you ask you some questions about your experiences as an administrator.

1. How many years have you been an administrator at [NAME OF SCHOOL/CENTER?]

2. How many years have you been an administrator in total?

Now I would like to talk about Pre-K and the transition of children into elementary school.

*If center administrator:

3. To what extent to you think that Pre-K in [NAME of CENTER] prepares children to succeed in elementary school?

4. To what extent do you see alignment between Pre-K at [NAME of CENTER] and elementary school?

5. Do you coordinate with individuals in elementary schools in order align instruction and facilitate the transition to kindergarten?

    If yes: In what ways?

    If no: Are there specific reasons why you do not?

*If elementary school administrator:

6. To what extent to you think that Pre-K in [NAME of SCHOOL] prepares children to succeed in elementary school?

7. To what extent do you see alignment between Pre-K at [NAME of SCHOOL] and elementary school?

8. Do you [or your K-3 teachers] coordinate with Pre-K teachers in your school in order align instruction and facilitate the transition to kindergarten?

    If yes: In what ways?

    If no: Are there specific reasons why you do not?

9. To what extent do you feel responsible for the Pre-K classrooms located in your school?

10. To what extent are the Pre-K classrooms located in your school integrated into the broader school?

As you may know, NC Pre-K classrooms can be located in either stand-alone centers or inside of elementary schools. My next set of questions focuses on the potential differences in children's Pre-K experiences based on location.

11. Do you think there are any differences in the Pre-K experience between center-based and school-based Pre-K programs?

    Probe: What are the key benefits of each?

    Probe: What are the key challenges of each?

12. Do you think there are any differences in the transition from Pre-K to elementary school between center-based and school-based Pre-K programs?

    Probe: What are the key benefits of each?

    Probe: What are the key challenges of each?

That concludes my prepared questions for this interview. Is there anything else you would like to share with me?

I sincerely appreciate you taking the time to speak with me today. Your insights are extremely valuable to me and will help inform our study.

COUNTY/STATE OFFICIALS

First, I would like you ask you some questions about your experiences as a [COUNTY/STATE] official.

1. Can you briefly describe your role at [COUNTY/STATE] office?

2. How many years have you been with [COUNTY/STATE] office?

As you may know, NC Pre-K classrooms can be located in either stand-alone centers or inside of elementary schools.

3. If you are familiar with the process, can you explain how decisions are made regarding where to locate Pre-K classrooms?

4. Do you have a preference for locating Pre-K programs in elementary schools or stand-alone centers?

   Probe: Why?

My next set of questions focuses on the potential differences in children's Pre-K experiences based on location.

5. Do you think there are any differences in the Pre-K experience between center-based and school-based Pre-K programs?

   Probe: What are the key benefits of each?

   Probe: What are the key challenges of each?

6. Do you think there are any differences in the transition from Pre-K to elementary school between center-based and school-based Pre-K programs?

   Probe: What are the key benefits of each?

   Probe: What are the key challenges of each?

Now I would like to talk about Pre-K and the transition of children into elementary school.

7. To what extent do you see alignment between NC Pre-K and elementary school?

   Probe: Do you see differences in alignment between school-based and center-based Pre-K programs?

8. Does your office work to coordinate alignment between Pre-K and kindergarten and facilitate the transition to kindergarten?

    If yes: In what ways?

    If no: Are there specific reasons why you do not?

That concludes my prepared questions for this interview. Is there anything else you would like to share with me?

I sincerely appreciate you taking the time to speak with me today. Your insights are extremely valuable to me and will help inform our study.

REFERENCES

Abry, T., Latham, S., Bassok, D., & LoCasale-Crouch, J. (2015). Preschool and kindergarten teachers' beliefs about early school competencies: Misalignment matters for kindergarten adjustment. Early Childhood Research Quarterly, 31, 78-88.

Allison, P. D. (2002). Missing data: Quantitative applications in the social sciences. *British Journal of Mathematical and Statistical Psychology*, *55*(1), 193-196.

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, *46*(3), 399-424.

Bailey, D., Duncan, G. J., Odgers, C., & Yu, W. (2017). Persistence and fadeout in the impacts of child and adolescent interventions, Journal of Research on Educational Effectiveness, 10(2), 7-39.

Barnett, W. S. (2011). Effectiveness of early educational intervention. Science, 333(6045), 975-978.

Barnett, S., Friedman-Krauss, A., Weisenfeld, G., Horowitz, M., Kasmin, M., & Squires, J. (2016). The State of Preschool 2016, The National Institute for Early Education Research.

Bassok, D., Gibbs, C. R., & Latham, S. (2015). Do the Benefits of Early Childhood Interventions Systematically Fade? Exploring Variation in the Association Between Preschool Participation and Early School Outcomes. Working Paper Series No, 36, EdPolicy Works. http://curry. virginia. edu/uploads/resourceLibrary/36_Preschool_Fade_Out. Pdf.

Bassok, D. (2010). Do Black and Hispanic children benefit more from preschool? Understanding differences in preschool effects across racial groups. Child Development, 81(6), 1828-1845.

Behrman, J. R., Cheng, Y., & Todd, P. E. (2004). Evaluating preschool programs when length of exposure to the program varies: A nonparametric approach. *Review of Economics and Statistics*, *86*(1), 108-132.

Bloom, H. S. (2006). *The Core Analytics of Randomized Experiments for Social Research.* MDRC Working Papers on Research Methodology.

Campbell, F. A., & Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. Child Development, 65, 684–698.

Campbell, F. A., & Ramey, C. T. (1995). Cognitive and school outcomes for high risk African American students at middle adolescence: Positive effects of early intervention. American Educational Research Journal, 32, 743–772.

Campbell, F. A., Ramey, C. T., Pungello, E. P., Sparling, J. J., & Miller- Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecedarian Project. Applied Developmental Science, 6, 42– 57.

Campbell, F. A., Pungello, E., Burchinal, M. R., Kainz, K., Pan, Y., Wasik, B. H., Barbarin, O., Sparling, J., Ramey, C. T. (2012). Adult outcomes as a function of an early childhood educational program: An Abecedarian Project follow-up. Developmental Psychology, 48(4), 1033-1043.

Campbell, F. A., & Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: a follow-up study of children from low-income families. Child Development, 65(2), 684-698.

Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. Psychological Bulletin, 54(4), 297.

Campbell, D. T., & Stanley, J. C. (2015). Experimental and quasi-experimental designs for research. Ravenio Books.

Cohen-Vogel, L., & Harrison, C. (2013). Leading with data: Evidence from the national center on scaling up effective schools. *Leadership and Policy in Schools*, *12*(2), 122-145.

Cochran, W. G. (1965). *Sampling Techniques: 2d Ed*. J. Wiley.

Creswell, J. W., & Clark, V. L. P. (2011). *Designing and Conducting Mixed Methods Research.* New York, NY: J. Wiley & Sons.

Creswell, J. W., & Tashakkori, A. (2007). Developing publishable mixed methods manuscripts. *Journal of Mixed Methods Research*, 1(2), 107-111.

Crocker, R., Thomas, D., & Currie, J. (2002). Longer-term effects of Head Start. The American Economic Review, 92(4), 999-1012.

Cunha, F., & Heckman, J. J. (2007). The technology of skill formation. The American Economic Review, 97 (2), 31–47.

Currie, J., & Thomas, D. (1995). Medical care for children: Public insurance, private insurance, and racial differences in utilization. Journal of Human Resources, 135-162.

Currie, J., & Thomas, D. (2000). School quality and the longer-term effects of Head Start. Journal of Human Resources, 35(4), 755-774.

Deaton, A., Cartwright, N. (2016). Understanding and Misunderstanding Randomized Controlled Trials. NBER Working Paper Series: 22595

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from Head Start. American Economic Journal: Applied Economics, 1(3), 111-134.

Dodge, K. A., Bai, Y., Ladd, H. F., & Muschkin, C. G. (2017). Impact of North Carolina's Early Childhood Programs and Policies on Educational Outcomes in Elementary School. Child development, 88(3), 996-1014.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. Developmental Psychology, 43(6), 1428–1446.

Duncan, G. J., & Magnuson, K. (2013). Investing in preschool programs. The Journal of Economic Perspectives, 27(2), 109-132.

Duncan, G. J., & Magnuson, K. (2011). The nature and impact of early achievement skills, attention skills, and behavior problems. Whither opportunity, 47-70.

Dolin, D. J., & Booth-Butterfield, S. (1995). Foot-in-the-door and cancer prevention. *Health Communication*, *7*(1), 55-66.

Enders, C.K. (2010). *Applied Missing Data Analysis.* New York, NY: The Guilford Press.

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs—principles and practices. *Health Services Research*, *48*(62), 2134-2156.

Figlio, D., & Roth, J. (2007). The behavioral consequences of pre-kindergarten participation for disadvantaged youth. In *The problems of disadvantaged youth: An economic perspective* (pp. 15-42). University of Chicago Press.

Fitzpatrick, M. D. (2008). Starting school at four: The effect of universal pre-kindergarten on children's academic achievement. The Journal of Economic Analysis & Policy, 8(1).

Fixsen, D. L., Naoom, S. F., Blase, K. A., & Friedman, R. M. (2005). Implementation research: a synthesis of the literature.

Gibbs, C., Ludwig, J., & Miller, D. L. (2011). Does Head Start do any lasting good? (No. w17452). National Bureau of Economic Research.

Gibbs, C., Ludwig, J., & Miller, D. L. (2013). Head Start origins and impacts. Legacies of the War on Poverty, 39-65.

Gormley, W. T., & Gayer, T. (2005). Promoting school readiness in Oklahoma: An evaluation of Tulsa's pre-k program. Journal of Human Resources, 40(3), 533-558.

Gormley, W. T. (2008). The effects of Oklahoma's pre-k program on Hispanic children. Social Science Quarterly, 89(4), 916-936.

Griefer, N. (2017). WeighIt: Weighting for Covariate Balance in Observational Studies. R package version 0.3.0.

Griefer, N. (2017). Cobolt: Covariate Balance Tables and Plots. R package version 3.1.0.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, *20*(1), 25-46.

Hill, J. L., Brooks-Gunn, J., & Waldfogel, J. (2003). Sustained effects of high participation in an early intervention for low-birth-weight premature infants. *Developmental Psychology*, 39 , 730–744

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, *15*(3), 199-236.

Imai, K., & Ratkovic, M. (2014). Covariate balancing propensity score. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *76*(1), 243-263.

Johnson, R. C., & Jackson, C. K. (2017). Reducing inequality through dynamic complementarity: Evidence from Head Start and public school spending. Working Paper 23489: National Bureau of Economic Research.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.

Kline, P., & Walters, C. R. (2016). Evaluating public programs with close substitutes: The case of Head Start. The Quarterly Journal of Economics, 131(4), 1795-1848.

Ladd, H. F., Muschkin, C. G., & Dodge, K. A. (2014). From birth to school: Early childhood initiatives and third-grade outcomes in North Carolina. Journal of Policy Analysis and Management, 33(1), 162-187.

Lee, V. E., Brooks-Gunn, J., & Schnur, E. (1988). Does Head Start work? A 1-year follow-up comparison of disadvantaged children attending Head Start, no preschool, and other preschool programs. Developmental psychology, 24(2), 210.

Lipsey, M. W., Farran, D. C., & Hofer, K. (2016). Effects of a state prekindergarten program on children's achievement and behavior through third grade. Peabody Research Institute: Working Paper.

Lipsky, M. (2010). Street-level bureaucracy.: Dilemmas of the individual in public service. New York, NY: Russell Sage Foundation.

Little, M.H., Cohen-Vogel, L., and Curran, F.C. (2016). Facilitating the transition to kindergarten: What ECLS-K data say about school practices then and now. AERA Open, 2(3), 1-18.

Little, R. & Rubin, D. (2002). *Statistical Analysis with Missing Data, 2nd edition.* New York: John Wiley.

Loeb, S., Bridges, M., Bassok, D., Fuller, B., & Rumberger, R. W. (2007). How much is too much? The influence of preschool centers on children's social and cognitive development. *Economics of Education Review*, 26 , 52–66.

Jenkins, J. M., Farkas, G., Duncan, G. J., Burchinal, M., & Vandell, D. L. (2016). Head Start at ages 3 and 4 versus Head Start followed by state pre-k: Which is more effective? *Educational Evaluation and Policy Analysis*, *38*(1), 88-112.

Magnuson, K. A., Ruhm, C., & Waldfogel, J. (2007). The persistence of preschool effects: Do subsequent classroom experiences matter? Early Childhood Research Quarterly, 22(1), 18-38.

Magnuson, K. A., Ruhm, C., & Waldfogel, J. (2007). Does prekindergarten improve school preparation and performance?. Economics of Education Review, 26(1), 33-51.

Magnuson, K. A., Kelchen, R., Duncan, G. J., Schindler, H. S., Shager, H., & Yoshikawa, H. (2016). Do the effects of early childhood education programs differ by gender? A meta-analysis. Early Childhood Research Quarterly, 36, 521-536.

Magnuson, K., & Shager, H. (2010). Early education: Progress and promise for children from low-income families. Children and Youth Services Review, 32(9), 1186-1198.

Marshall, C., & Rossman, G. B. (2014). *Designing qualitative research*. Sage.

McCaffrey, D. F., Ridgeway, G., & Morral, A. R. (2004). Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological Methods*, *9*(4), 403.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Sage.

Muschkin, C. G., Ladd, H. F., & Dodge, K. A. (2015). Impact of North Carolina's early childhood initiatives on special education placements in third grade. Educational Evaluation and Policy Analysis, 37(4), 478-500.

O'brien, M., Weaver, J., Burchinal, M., Clarke-Stewart, A., Vandell, D. (2013). Women's work and child care: Perspectives and prospects In Gershoff, E., Mistry, R., Crosby, D. (Ed.), Societal Contexts of Child Development: Pathways of Influence and Implications for Practice and Policy. (pp. 37-40). New York, NY: Oxford University Press

Odden, A. (Ed.). (1991). Education policy implementation. Albany: State University of New York Press.

Office of Child Care. (2017). Office of Child Care fact sheet. Retrieved from https://www.acf.hhs.gov/occ/fact-sheet-occ

Office of Head Start. (n.d.). Head Start programs. Retrieved from https://www.acf.hhs.gov/ohs/about/head-start

Ogawa, R. T. (2009). Improvement or reinvention: Two policy approaches to school reform. In G. Sykes, B. Schneider, & D. N. Plank (Eds.), Handbook of education policy research (pp. 534-540). New York, NY: Routledge.

Penuel, W. R. (2006). Implementation and effects of one-to-one computing initiatives: A research synthesis. Journal of Research on Technology in Education, 38(3), 329-348.

Phillips, D. A., & Shonkoff, J. P. (Eds.). (2000). From neurons to neighborhoods: The science of early childhood development. Washington, D.C.: National Academies Press.

Phillips, D., Gormley, W., & Anderson, S. (2016). The effects of Tulsa's CAP Head Start program on middle-school academic outcomes and progress. Developmental psychology, 52(8), 1247.

Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., Mashburn, A., Downer, J. (2010). Head Start Impact Study: Final Report. U.S. Department of Health and Human Services, Administration for Children and Families, Washington, DC.

Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., Mashburn, A., Downer, J. (2012). Third grade follow-up to the Head Start Impact Study: Final Report. U.S. Department of Health and Human Services, Administration for Children and Families, Washington, DC.

Pungello, E. P., Kainz, K., Burchinal, M., Wasik, B. H., Sparling, J. J., Ramey, C. T., & Campbell, F. A. (2010). Early educational intervention, early cumulative risk, and the early home environment as predictors of young adult outcomes within a high-risk sample. Child Development, 81(1), 410-426.

Reynolds, A. J., Temple, J. A., Ou, S.-R., Arteaga, I. A., & White, B. A. B. (2011). School-based early childhood education and age-28 well-being: Effects by timing, dosage, and subgroups. Science , 333 , 360–364.

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

Rubin, D. (1976). Inference and missing data. *Biometrika*, 63(3), 581-592.

Rubin, D. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York, NY: J. Wiley & Sons.

Schweinhart, L. J., Barnes, H. V., & Weikhart, D. P. (2005). Significant benefits: The High/Scope Perry preschool study through age 27. Child Welfare: Major Themes in Health and Social Welfare, 9-29.

Schweinhart, L. J., Berrueta-Clement, J. R., Barnett, W. S., Epstein, A. S., & Weikart, D. P. (1985). Effects of the Perry Preschool program on youths through age 19: A summary. Topics in Early Childhood Special Education, 5(2), 26-35.

Schulting, A. B., Malone, P. S., & Dodge, K. A. (2005). The effect of school-based kindergarten transition policies and practices on child academic outcomes. Developmental Psychology, 41(6), 860.

Shadish, W. R., Cook, T.D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Independence, KY: Wadsworth Cengage learning.

Shadish, W. R. (2010). Campbell and Rubin: A primer and comparison of their approaches to causal inference in field settings. Psychological Methods, 15, 3–17.

Shager, H. M., Schindler, H. S., Magnuson, K. A., Duncan, G. J., Yoshikawa, H., & Hart, C. M. (2013). Can research design explain variation in Head Start research results? A meta-analysis of cognitive and achievement outcomes. Educational Evaluation and Policy Analysis, 35(1), 76-95.

Shaw, L. (2014). How Boston's Preschools Went From Mediocre to Outstanding. The Seatle Times. Retrieved from https://www.seattletimes.com/education-lab/how-bostons-preschools-went-from-mediocre-to-outstanding/

StataCorp. (2015). *Stata 14 Base Reference Manual.* College Station, TX: Stata Press.

Tourangeau, K., Nord, C., Lê, T., Sorongon, A. G., Hagedorn, M. C., Daly, P., & Najarian, M. (2015). Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K: 2011). User's Manual for the ECLS-K: 2011 Kindergarten Data File and Electronic Codebook, Public Version. NCES 2015-074. *National Center for Education Statistics*.

Vandell, D. L., Belsky, J., Burchinal, M., Steinberg, L., Vandergrift, N., & NICHD Early Child Care Research Network. (2010). Do effects of early child care extend to age 15 years? Results from the NICHD study of early child care and youth devel opment. *Child Development*, 81,, 737–756

Vygotsky, L. S. (1997). The collected works of LS Vygotsky: Problems of the theory and history of psychology (Vol. 3). Springer Science & Business Media.

Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. Child Development, 84(6), 2112-2130.

White, I., Royston, P., & Wood, A. (2010). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, 30(4), 377-399.

Young, R. & Johnson, D. (2011). Imputing the Missing Y's: Implications for Survey Producers and Survey Users. Proceedings of the AAPOR Conference Abstracts, pp. 6242–6248.

Zubizarreta, J. R. (2015). Stable weights that balance covariates for estimation with incomplete outcome data. *Journal of the American Statistical Association*, *110*(511), 910-922.