HEAD START: ASSESSING COMMON EXPLANATIONS FOR THE APPARENT DISAPPEARANCE OF INITIAL POSITIVE EFFECTS

by

Pete M. Bernardy
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Public Policy

Committee:

David J. Armor, Chair

Todd A. Olmstead

C. Stephen White

Judy A. Temple, External Reader

James P. Pfiffner, Program Director

Edward Rhodes, Dean

Date: 11/6/12

Fall Semester 2012
George Mason University
Fairfax, VA

Head Start: Assessing Common Explanations for the Apparent Disappearance of Initial Positive Effects

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

By

Pete M. Bernardy
Master of Public Policy
University of Minnesota, 2004
Bachelor of Arts
University of Saint Thomas, 2002

Dissertation Director: David Armor, Professor Emeritus
School of Public Policy

Fall Semester 2012
George Mason University
Fairfax, VA

ACKNOWLEDGEMENTS

I am indebted to my parents for promoting, from my earliest years, the habits of learning and perseverance that I repeatedly drew upon in completing this dissertation. My mother, Connie, continued to read my favorite children's books to me well after she had them memorized, made sure I knew all of my spelling and vocabulary words before going to bed the night before a test, and has provided comforting emotional support during my most vulnerable moments. My father, Jerry, never missed an opportunity to ask me if my homework was done, routinely demonstrated how to rigorously evaluate potential courses of action, and always put my learning opportunities before his personal interests.

While my parents were my first teachers, I would not have completed this dissertation without the selfless instruction and inspiration of the many other educators that have assisted me along the way. Most recently and directly, my dissertation committee generously shared their time and expertise. In particular, Dr. David Armor provided reliably thoughtful and prompt feedback, Dr. Todd Olmstead tactfully identified countless inconsistencies in my analysis that needed to be corrected, and Dr. White facilitated my learning of key child development and education concepts. Of course, I alone am responsible for any flaws contained in this final product.

My colleagues at the Department of Health and Human Services regularly renew my commitment to public service and have played an important role in my development. My various supervisors provided the flexibility needed to work full-time while pursuing an advanced degree. Any views expressed herein do not necessarily represent the views of the Department of Health and Human Services or the United States Government.

I would also like to extend appreciation to my family and friends, who have remained supportive despite the selfish choices I have made in the interest of conducting this research. My brother, Dave, continues to inspire me through his passion for life and commitment to those closest to him. My dear friends, Sean, Bryce, Michael, Scott, Mark, and Devin bring me great joy through their humor and friendship. My running pals, in particular the *Thursday Morning Breakfast Crew*, reliably provide a healthy and camaraderie-filled escape from work and school.

Through the process of completing this dissertation I have accrued my greatest debt to my wife and dearest friend, Robyn, who has graciously accommodated the demands of my various pursuits.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

HEAD START: ASSESSING COMMON EXPLANATIONS FOR THE APPARENT DISAPPEARANCE OF INITIAL POSITIVE EFFECTS

Pete M. Bernardy, Ph.D

George Mason University, 2012

Dissertation Director: Dr. David J. Armor

Experimental design evaluations have consistently found children given access to early childhood services through the federal Head Start program experience better academic and social outcomes relative to comparable peers by the end of their participation, but this early advantage is not sustained through the early elementary grades. However, two studies of the long-term impact of Head Start have found the program to produce improved rates of high school completion. Given these seemingly contradictory findings, this research uses data from a recent nationally representative random assignment study of this program to examine whether there is evidence of enduring effects of Head Start participation: (1) when controlling for within-child variation; (2) for learning skills not previously analyzed in published reports; (3) for children with higher quality Head Start experiences; (4) for children with higher quality early elementary school experiences; and (5) compared to a counterfactual of no preschool participation. No evidence of initial positive effects enduring into kindergarten or first grade is found.

CHAPTER ONE: INTRODUCTION

Policy Context

The federal Head Start program has provided education, health, nutrition, and parent involvement services to low-income preschool-aged children and their families since 1965. Grantees provide these services using various program models, including partial and full day programs carried out in centers, schools, family child care homes, and through visits to children's homes. In addition to the more than 800,000 three- and four-year-olds receiving services through Head Start each year, more than 100,000 pregnant women, infants, and toddlers receive services through Early Head Start.[1]

Evaluations of Head Start have consistently found that children granted access to the program demonstrate improved cognitive and social/emotional abilities by the end of their participation, above and beyond improvement that would be expected absent program participation (McKey et al., 1985). Recently, a nationally representative random assignment Head Start Impact Study found access to Head Start services to be associated with positive cognitive and social/emotional outcomes by the end of the program year (Department of Health and Human Services, 2005). Consistent with previous research on Head Start, this study found that these initial positive effects were not sustained through

[1] The term "Head Start" is often used to refer to the Head Start and Early Head Start programs collectively. In this paper, the term "Head Start" is used to refer only to the Head Start program that targets three- and four-year olds and their families.

first grade (Department of Health and Human Services, 2010a; McKey et al., 1985).

Some policy analysts cite the relative absence of detectable program effects in first grade as evidence that the more than $6 billion spent on the program annually should be diverted to uses believed to be more effective. This recommendation has been raised throughout the program's history, including in the wake of a 1969 report from a national impact study that found children attending Head Start did not demonstrate better outcomes in early elementary school than a matched control group, with the exception of children that attended full year programs, who performed marginally better on select measures of cognitive development (Ohio University and Westinghouse Learning Corporation, 1969).

Other analysts rebut this critique of the program's effectiveness by critiquing the methods used by studies that have not found positive effects and citing studies that have found positive long term effects associated with Head Start participation; by suggesting the program has been improved since the period covered by the most recent evaluation and ongoing improvements will make it even more effective; by suggesting that improving the nation's early elementary education system would enable initial program effects to endure; and by noting that a sizable proportion of children that make up the comparison group for evaluations of Head Start attend publicly funded preschool and other early childhood programs.

It is important for policymakers to understand whether there is empirical support for these explanations of the evidence that initial positive effects associated with Head Start are not sustained through the early elementary years. Such an understanding is

relevant for discussions regarding the overall level of funding for the program, as well as for identifying ways to improve the program and to effectively target its resources.

Research Questions

Toward this end, this paper presents secondary analyses of data from the Head Start Impact Study (Department of Health and Human Services, 2010a), which collected information on 4,667 children that applied to a nationally representative sample of Head Start programs in 2002, to answer the following questions:

- Does controlling for within-child variation facilitate the detection of previously unidentified effects associated with being given access to Head Start?
- Is there evidence of enduring effects associated with being given access to Head Start for measures of child attention, persistence, and confidence not addressed in the official report of the Impact Study?
- Is there evidence that the quality of care and education provided in Head Start classrooms explains children's outcomes through first grade?
- Is there evidence that early elementary school experiences influence whether initial effects of Head Start endure?
- Is there evidence of enduring effects associated with being given access to Head Start compared to a counterfactual of no preschool participation?

Collectively, these questions focus on whether there is evidence to support the most common explanations for the apparent disappearance of initial positive effects associated with Head Start. The following chapter presents past research findings related

to the prospects for quality early childhood education yielding enduring effects, and the subsequent chapters present the research design and findings of the present study.

CHAPTER TWO: BACKROUND LITERATURE

A young child's immediate environment plays a critical role in her early development (Bronfenbrenner, 1979). Within this context, a child's primary caregivers and educators are in a unique position to facilitate interactions that foster cognitive and social/emotional development and development in other key domains (Vygotskii, 1994). The mere consistent presence of a trusted adult facilitates exploration that is important for early learning (Shonkoff & Phillips, 2000).

This chapter introduces key findings from research on early childhood development; the effectiveness of early childhood interventions (with an emphasis on the Head Start program); and the role of early elementary experiences in child development. Within the context of the five research questions addressed in this study, the findings presented in this chapter are oriented toward assessing the prospects for quality early childhood education yielding positive enduring effects.

Early Childhood Development and its Implications

Cognitive and social/emotional development during early childhood sets the stage for future learning and relationship building. Cognitive development is facilitated by encouragement of exploration, promotion of a rich verbal environment, opportunity to experiment with numerical concepts, and exposure to tasks that are challenging but achievable with assistance (Shonkoff & Phillips, 2000). Social/emotional development is facilitated by setting clear expectations for conduct, enforcing these expectations firmly, and providing emotional support (Shonkoff & Phillips, 2000; White, 1993).

These two key aspects of early childhood development influence each other (Berk, 2009; Denham, 1998; National Research Council, Committee on Early Childhood Pedagogy, 2000). For instance, much of a child's cognitive development occurs through play and other interactions with peers. A socially adept child that is able to regulate his own emotions and interpret the emotions of his peers can be expected to spend a greater proportion of his childhood playing with his peers. However, these social/emotional skills themselves depend on cognitive capacities (Berk, 2009) that influence the way a child expresses, understands, and regulates his emotions (Denham, 1998).

While humans remain open to change throughout their lifecycle, their potential for fundamental change declines steadily with age (Shonkoff & Phillips, 2000). In economic terms, this results in progressively lower rates of return for investments in human capital over the course of the lifecycle (Heckman, 2006). The importance of early cognitive development can be seen in the strong correlation between intelligence and achievement scores measured in early elementary school and at older ages (Armor, 2003; Berk, 2009). While early brain development has enduring significance, the brain retains plasticity throughout life (Shonkoff & Phillips, 2000). In the social/emotional domain, studies have shown that temperamental qualities appear early in infancy and show signs of stability at four months (Shonkoff & Phillips, 2000). After 38 years of research involving visits to the homes of thousands of families, Burton White (1993) wrote, "My feeling is that once a child reaches two years of age, his primary social orientation has been established and from then on, it becomes increasingly difficult to alter it significantly." Managing one's emotions and establishing relationships with other children are key developmental tasks of early childhood, and the success of a given child at this task is a predictor of subsequent competence or deviance in adolescence (Shonkoff & Phillips, 2000).

A meta-analysis of data from 62 distinct samples found variation in children's social development and behavior in preschool or kindergarten to account for 10 percent of the variance in these domains a year or two later (La Paro & Pianta, 2000). By comparison, variation in children's cognitive/academic performance in preschool or kindergarten accounted for 25 percent of the variance in these domains a year or two later. Informed by findings such as these, it has been suggested that social skills and motivation hold greater promise for serving as the mechanism through which early childhood education can promote improved long-term performance in school and the workplace (Heckman, 2000). One interpretation of stability in cognitive and social/emotional competencies over time is that early development influences subsequent development. An alternative interpretation is that children are generally exposed to similar environmental factors, like parenting behaviors, from birth through early adulthood (Entwisle, Alexander, & Olson, 2005). For instance, children that experience poverty for longer durations, in greater extremes, and during preschool and early school years have been found to be at increased risk for adverse outcomes (Brooks-Gunn & Duncan, 1997). Children living in single-parent homes have also been found to realize less favorable educational, occupational, and behavioral outcomes, which can be accounted for by changes in economic resources and parental resources (McLanahan & Sandefur, 1994).

Effectiveness of Early Childhood Interventions

Meta-analyses of early childhood education studies carried out in the United States have consistently found positive cognitive and social/emotional outcomes for children participating in the intervention relative to a comparison group of children with similar characteristics (Camilli, Vargas, Ryan, & Barnett, 2010; Gorey, 2001; Karoly, Kilburn, & Cannon, 2005; Nelson, Westhues, & MacLeod, 2003). The most

comprehensive analysis drew data from 123 studies of center-based programs that provided educational services directly to children for at least 10 hours per week for at least two months and found mean effect sizes of 0.23 in the cognitive domain and 0.16 in the social/emotional domain (Camilli et al., 2010).

Other meta-analyses have only included studies that use a specific research design (e.g., experimental design), but allowed the inclusion of studies of programs that target the improvement of parent-child interactions rather than provide direct educational services to children. One such analysis, which drew from 34 studies, found a mean effect size of 0.30 for school-age cognitive outcomes and a mean effect size of 0.33 for school-age and adult social/emotional outcomes (Nelson et al., 2003). This analysis found larger cognitive effects for studies that provided educational services directly to the children (rather than parent-centered approaches) and were longer in duration in terms of hours per day and period of enrollment. This analysis did not find that early childhood programs that begin working with children at a younger age produce better outcomes. A separate meta-analysis of data from 20 studies that met similar inclusion criteria found a mean effect size of 0.28 for school-age cognitive outcomes (Karoly et al., 2005).

A meta-analysis of data from 60 programs that used home visiting as their primary service mechanism found end-of-treatment effect sizes of 0.18 for cognitive child outcomes, 0.10 for social/emotional outcomes, and 0.11 to 0.24 for parenting outcomes (Sweet & Appelbaum, 2004). Home visiting programs tend to focus on training parents to help their children, rather than on providing educational services directly to children, which may account for the slightly lower effect sizes. This study did not examine whether these effects increased or decreased following completion of the intervention, for those studies for which such data were available.

The Perry Preschool Study carried out in Michigan randomly assigned 123 at-risk three- and four-year-olds to a treatment or control group in the early 1960s and collected data at numerous subsequent ages. Those children assigned to the treatment group attended a preschool (2.5 hours a day for one or two preschool years) led by teachers (one for each five children) certified to teach in early childhood that implemented a curriculum that integrated child and teacher control of activities. Results immediately following preschool included intelligence gains, which were not identifiable in second grade or beyond (Schweinhart & Weikart, 1998). Nevertheless, the treatment group demonstrated better academic and behavioral outcomes through high school and into adulthood, in comparison to the control group (Belfield, Nores, Barnett, & Schweinhart, 2006; Heckman, Moon, Pinto, Savelyev, & Yavitz, 2010; Schweinhart & Weikart, 1998).

The Chicago Longitudinal Study is an ongoing quasi-experimental study of 1,539 low-income children born in 1979-1980 who participated in Child-Parent Center full-day kindergarten and received enriched school-aged services through third grade. The intervention group consists of 989 children that attended Child-Parent Center preschool classrooms of 17 children, led by teachers with bachelor degrees and early childhood certification who were assisted by teacher aides. Approximately 15 percent of the comparison group attended Head Start preschool and the rest were in home care. The association between participation in this program and positive long-term academic, behavioral, and employment outcomes has been documented extensively (Ou & Reynolds, 2010; Reynolds & Ou, 2011; Reynolds, Temple, & Ou, 2010; Reynolds, Temple, Ou, Arteaga, & White, 2011; Reynolds, Temple, White, Ou, & Robertson,

2011). Notably, data from this study has been used to examine which early childhood and youth outcomes mediate longer-term outcomes such as high school completion and incarceration. For instance, the intervention group was nearly nine percentage points more likely to graduate from high school than the comparison group, of which 46 percent was accounted for by shorter term cognitive and non-cognitive effects (Reynolds et al., 2010). Similarly, the intervention group was five percentage points less likely to be incarcerated as adults than the comparison group, of which fifty-nine percent was accounted for by shorter term cognitive and non-cognitive effects.

The Abecedarian Early Childhood Intervention study randomly assigned 112 at-risk children born in North Carolina between 1972 and 1977 to a treatment or control group when they were between six and 12 weeks of age. The treatment group attended high quality full-day center-based care, up to 10 hours a day, through kindergarten. The control group, which was not offered such care through the study but did receive iron-fortified formula for the first fifteen months to reduce the likelihood that group differences were a result of differences in early nutrition, participated in a variety of child care arrangements that included center-based care. This intervention produced positive cognitive and academic benefits (Campbell, Pungello, Miller-Johnson, Burchinal, & Ramey, 2001), but did not produce social/emotional benefits (Barnett & Masse, 2007). Adult outcomes for the treatment group relative to the control group were similar to, or exceeded, those found in the Perry Preschool study (Barnett & Masse, 2007).

Early Head Start provides home-based and/or center-based services to pregnant women and at-risk children up to three years of age, and their families. The Early Head Start Research and Evaluation Study followed 3,001 children that were randomly assigned in 1996-1998 to participate in Early Head Start programming or precluded from enrolling in such programming. Upon program completion at age three, modest favorable cognitive and social/emotional impacts were identified for the enrolled group, along with modest parenting and family self-sufficiency outcomes (Department of Health and Human Services, 2010b). Two years after program completion, prior to the children's entry in kindergarten, social/emotional and parenting impacts endured, but cognitive impacts were no longer detectable for the overall sample (Department of Health and Human Services, 2010b; Zill, 2010). Seven years after program completion, when the children were in fifth grade, there were no longer impacts for the overall sample (Department of Health and Human Services, 2010b).

Ideally, examination of the extensive research that has been conducted on such a varied set of approaches to early child education would yield insights into what types of interventions work best, for which children, under what circumstances, and in which developmental domains. Unfortunately, even those studies that have applied meta-analytic methods comprehensively in an attempt to answer these questions have found few answers (Camilli et al., 2010).

Immediate Effects of Head Start Participation

The federal Head Start program has provided education, health, nutrition, and parent involvement services to low-income preschool-aged children and their families since 1965. These services are provided by grantees that use various program models, which include partial and full day programs carried out in centers, schools, family child care homes, and/or through visits to children's homes. In addition to the more than

800,000 three- and four-year-olds receiving services through Head Start each year, more than 100,000 pregnant women, infants, and toddlers receive services through Early Head Start.

Studies of Head Start have consistently found that participating children demonstrate improved cognitive and social/emotional outcomes by the time they complete the program, above and beyond the improvement that would be expected in the absence of participation. A meta-analysis of more than 210 research reports on local Head Start programs conducted from the late 1960s through the early 1980s found consistent positive cognitive and social/emotional gains by the spring of the Head Start year (McKey et al., 1985). Among the studies with treatment and control groups that were included in this analysis, positive cognitive effects were found for both intelligence (0.59 effect size) and achievement (0.54 effect size). Slightly smaller effects were found among the pre/post studies included in this meta-analysis. Among all studies reporting social/emotional effects, statistically significant effects were found for all three sub-domains, which were self esteem (0.17 effect size), achievement motivation (0.22 effect size), and social behavior (0.35 effect size).

During the 25 years since this meta-analysis was published, two studies have assessed the impact of Head Start on cognitive and social/emotional outcomes by forming treatment and control groups using random assignment. The first such study placed 87 four-year-olds in seven high quality Head Start classrooms and another 86 comparable children on a wait list (Abbott-Shim, Lambert, & McCarty, 2003). This study found that in the spring of the preschool year, Head Start participants demonstrated greater improvement than the control group on the two cognitive measures for which they had been found to be equivalent the prior fall (i.e., receptive vocabulary and phonemic awareness). This study did not find statistically significant differences between these two

groups on children's social skills or problem behaviors, as reported by their parents. Given the small sample size and the imprecise data reported by the children's parents, it is possible that there were undetected differences in social/emotional outcomes between these two groups.

The other experimental design study of Head Start, which is the most comprehensive and rigorous evaluation of the program to date, found the provision of access to Head Start services in 2002-2003 to be associated with small to moderate positive cognitive and social/emotional outcomes for children as well as positive parenting outcomes, upon the completion of the program (Department of Health and Human Services, 2005). The size of these effects were smaller than those found in the meta-analysis described above, which may be attributable to the fact that a greater percentage of the control group participants attended preschool as a result of state and local investments in preschool programming that have been made since the studies examined in the meta-analysis were conducted. This experimental design study of Head Start, which is known as the Head Start Impact Study, continued to track these outcomes into elementary school, as discussed in the following subsection.

Early Elementary Effects of Head Start Participation

The meta-analysis introduced in the previous section (McKey et al., 1985) also assessed the impact of Head Start beyond the completion of the program. This study found intelligence effects fell from 0.59 upon completion of the program to 0.09 a year after completion and were no longer detectable two years after completion and beyond. Achievement effects fell from 0.54 upon completion of the program to 0.20 a year after completion and to 0.13 two years after completion, before being undetectable three years after completion and beyond. Positive effects initially seen for self-esteem, achievement

motivation, and social behavior demonstrated a similar trend during the three years following program completion.

Consistent with the findings of this meta-analysis, the nationally-representative Head Start Impact Study found that the positive effects that were found upon program completion were not sustained through first grade (Department of Health and Human Services, 2010a). The fact that initial effects did not endure as long among the participants in this study as among those included in the meta-analysis may be attributable to the fact that a greater proportion of the control group participants in the more recent study attended preschool.

While the Head Start Impact Study is the only evaluation of the Head Start program to use random assignment to assess the effect of the program beyond the period of participation, quality longitudinal observational studies have been conducted to assess outcomes in elementary school. A nationally representative observational study controlled for observable and unobservable characteristics by comparing the outcomes of sibling pairs that included one child that attended Head Start and another that did not (Currie & Thomas, 1995). This study found that White children participating in Head Start were 16 percentage points less likely to repeat a grade by age 10 than their siblings that did not participate. Black children participating in Head Start were no less likely than their siblings to repeat a grade by age 10. This study also found Black and White children who participated in Head Start demonstrated similar initial gains in receptive vocabulary compared to non-Head Start participants and that these effects were sustained longer among White children than among Black children. At age 10, the effect for White children remained five percentile points while there was no remaining effect for Black children.

In a subsequent study, the same researchers found that Black Head Start participants attended schools with lower overall student performance than those attended by Black non-Head Start participants, but found no evidence of this among White children (Currie & Thomas, 2000). These findings provided added nuance to the findings of a previous study which found that on average Head Start participants attended lower performing schools than their non-Head Start counterparts (Lee & Loeb, 1995). Currie and Thomas suggest this differential in school quality may explain the shorter duration of cognitive advantage for Black Head Start participants (relative to Black non-participants) compared to the advantage for White participants over non-participants. However, this theory seems to be contradicted by the findings of a subsequent study that found preschool achievement effects to diminish more quickly in smaller elementary classrooms and elementary classrooms with higher levels of academic instruction (Magnuson, Ruhm, & Waldfogel, 2007).

An observational study of 696 children eligible for Head Start in Trenton, New Jersey and Portland, Oregon explored the effect of participating in Head Start, controlling for family background, demographic factors, and baseline cognitive and social skills (Lee, Brooks-Gunn, Schnur, & Liaw, 1990). This study found an effect size of 0.4 for social competence at the beginning of first grade, as reported by the child's teacher using a 30-item instrument capturing child work habits, interpersonal skills, frustration, and help seeking. This study did not find statistically significant effect sizes for verbal achievement and perceptual reasoning, the two cognitive skills it measured, despite having found such effects upon program completion.

The effect sizes of 0.15 to 0.30 found in the meta-analyses presented above are similar in magnitude to the effects found in the large experiments for Head Start and Early Head Start upon completion of the intervention. However, the meta-analyses

considered outcomes ranging from the conclusion of the intervention all the way through adulthood. Given this consideration, it is not clear why the meta-analyses found effect sizes that were comparable to (or larger than) the post-intervention effects of these large experiments, which were not sustained. Further, it is not clear why the two meta-analyses that only included studies that collected outcome data from elementary school or beyond found larger effect sizes than the third meta-analysis (which used data from studies that only collected post-intervention outcomes). One potential explanation is that those studies that followed participants for a longer period of time tended to be studying higher quality interventions, as measured by structural features such as higher ratios of teachers to children and higher proportions of teachers with credentials. The studies of the Perry Preschool, Chicago Child-Parent, and Abecedarian programs would each fall into this category. Given Head Start centers are not funded at a level that would facilitate matching the quality of care provided in programs like the three just mentioned, this hypothesis would also account for the discrepancy between the two meta-analyses that found effects into elementary school and beyond and the findings from the Head Start experiments. An alternative explanation would be that the benefits of early childhood education are not easily detected in early elementary school, but reemerge later in life (Peisner-Feinberg et al., 2001). While there is not a theoretical explanation for this explanation, evidence suggestive of it has been identified in at least three studies (Broberg, Wessels, Lamb, & Hwang, 1997; Magnuson et al., 2007; Schweinhart & Weikart, 1998).

Longer-term effects of Head Start Participation

The only two studies that have assessed the long-term impact of Head Start have found the program to be associated with higher rates of high school completion. A nationally representative observational study controlled for observable and

unobservable characteristics by comparing the outcomes of sibling pairs that included one child that attended Head Start and another that did not (Garces, Thomas, & Currie, 2002). This study found that White Head Start participants, relative to their siblings that did not participate in Head Start, were 20 percentage points more likely to complete high school, 28 percentage points more likely to attend college, and no more or less likely to be booked or charged for a crime. Black Head Start participants, relative to their siblings, were 12 percentage points less likely to be booked for a crime, and were no more or less likely to complete high school or attend college.

Jens Ludwig and Douglas Miller (2007) utilized regression discontinuity analysis to capitalize on the way Head Start was launched in order to estimate the program's effects on high school completion. In 1965 the federal government provided grant-writing assistance to the 300 poorest counties in the United States, which produced a discontinuity in county-level Head Start funding per four-year-old that persisted through the late 1970s. Through examination of outcome data for children in cohorts that would have been preschool age during this period in the counties on both sides of the county poverty cutoff, this study found a 50-100 percent increase in Head Start funding to be associated with a four percentage point increase in completion of high school or an equivalent degree. To put this effect in context, it is worth noting that the difference in Head Start enrollment rates between the two groups of counties was between 12,000-30,000 per 100,000 four-year olds. This study did not find differences in eighth grade with respect to reading or math scores, absences, grades, or time spent on homework.

Long-Term Effects despite Unidentifiable Effects in Early Elementary School?

As indicated above, studies of the long-term impact of Head Start show the program to be associated with improved high school completion rates (Garces et al.,

2002; Ludwig & Miller, 2007). What accounts for this apparent effect on high school completion, given the disappearance of measured gains for cognitive and social/emotional outcomes by early elementary school found in other studies? One potential explanation for longer term benefits of Head Start despite the apparent discontinuation of initial advantages is that undetected effects endure throughout the school years and yield large high school completion effects. This explanation is supported by findings elsewhere that small cognitive and noncognitive differences can lead to sizable differences in school completion (Heckman, Stixrud, & Urzua, 2006).

In addition to the possibility of small effects not registering as statistically significant, undetected effects could be the result of competencies that have not been measured (or have been measured imprecisely) or not been analyzed as part of evaluations conducted to date. Attention skills are an example of a competency that has shown promise to predict subsequent achievement outcomes (Barriga et al., 2002; Duncan et al., 2007; Howse, Lange, Farran, & Boyles, 2003), but has not been emphasized in evaluations of Head Start. Studies of other early childhood education programs have reported evidence of positive long-term effects despite undetected effects in the early elementary years (Broberg et al., 1997) and one study found evidence of the effects of preschool on reading achievement scores increasing from first grade to third grade (Magnuson et al., 2007).

The Role of Early Elementary Experiences

It is possible that children's experiences in early elementary school account for the apparent disappearance of the initial advantage in academic and behavioral outcomes

experienced by Head Start participants in the experimental design studies discussed previously. Research has consistently found great variation in the characteristics of early elementary classrooms (Bryant, Clifford, & Peisner, 1991; Meyer, Wardrop, Hastings, & Linn, 1993; National Institute of Child Health and Human Development & Network, 2002; Pianta, Paro, Payne, Cox, & Bradley, 2002). Characteristics of early elementary settings are often grouped into structural features, which are enduring characteristics of the educational environment, and process features reflecting the experiences of a child in that environment. Structural measures of quality are not necessarily consistent or strong predictors of process quality (Bryant et al., 1991; National Institute of Child Health and Human Development & Network, 2002).

Structural features such as smaller class sizes and longer school days are commonly assumed to yield improved child outcomes. A well-known large-scale experimental design study conducted in Tennessee in the 1980s found that children participating in smaller classes from kindergarten through third grade demonstrated improved academic and behavioral outcomes into seventh grade (Finn & Achilles, 1999) and a quasi-experimental study in Wisconsin also found smaller classes to be associated with improved academic and behavioral outcomes (Molnar et al., 1999). There is evidence that the relationship between class size and student outcomes is mediated by social and instructional processes (Allhusen et al., 2004; Betts & Shkolnik, 1999). However, analyses of broader observational data have failed to find evidence of systematic advantages to smaller class sizes (Hanushek, 1999; Milesi & Gamoran, 2006).

With respect to the impact of extending school days or academic years, a review of research over the last quarter of a century suggests that such extensions yield positive child outcomes (Patall, Cooper, & Allen, 2010). By contrast, simple measures of teacher experience and completion of advanced degrees do not necessarily translate into improved child outcomes either (Rivkin, Hanushek, & Kain, 2005).

Unlike structural features, process features are related to the actual experiences of children in the learning environment. Certain relationship dynamics between children and their early elementary teachers, including closeness and the absence of conflict and dependence, have been found to explain positive child outcomes (Birch & Ladd, 1997; Hamre & Pianta, 2001). There is evidence that this occurs through higher levels of child engagement as well as greater teacher attention to the child (O'Connor & McCartney, 2007; Pianta et al., 2002; Pianta & Stuhlman, 2004).

Unfortunately, as Robert Pianta and Bridget Hamre (2009), leaders in the movement to improve researchers' ability to examine such features, observe: "…little to no population-level data exist pertaining to exposure of children and adolescents to particular classroom practices that are either known to relate to academic success or failure…or even hypothetically expected to relate to outcomes."

Conclusion

This literature review presents evidence that quality early childhood care and education contributes to modestly improved cognitive and social/emotional outcomes during early childhood. Long-term studies have found examples of these initial effects translating into improved adolescent and adult outcomes. For the Head Start program in

particular, two long-term studies found participation to be associated with higher high school completion rates. However, shorter term studies have found that initial positive effects associated with program participation were not sustained through the early elementary grades. The following chapter introduces a research design for the present study that is informed by this literature.

CHAPTER THREE: DATA & METHODS

This chapter presents the conceptual model, sample, measures, and methods used in the analyses presented in subsequent chapters. The data described in this chapter and analyzed in subsequent chapters was collected through the Head Start Impact Study and is available to researchers that sign a data use agreement protecting the confidentiality of the study's participants. The conceptual model, which is informed by the literature discussed in the preceding chapter, provides a framework for the measures and methods presented later in the chapter, which are used to address the following research questions:

- Does controlling for within-child variation facilitate the detection of previously unidentified effects associated with being given access to Head Start?
- Is there evidence of enduring effects associated with being given access to Head Start for measures of child attention, persistence, and confidence not addressed in the official report of the Impact Study?
- Is there evidence that the quality of care and education provided in Head Start classrooms explains children's outcomes through first grade?
- Is there evidence that early elementary school experiences influence whether initial effects of Head Start endure?
- Is there evidence of enduring effects associated with being given access to Head Start compared to a counterfactual of no preschool participation?

Conceptual Model

Figure 1 depicts a conceptual model through which early childhood education may influence adult outcomes (Ou, 2005; Reynolds & Ou, 2011). Consistent with the research presented in the preceding chapter, this model shows children's baseline cognitive and social/emotional development predicting their first grade status in these same areas of development (Berk, 2009; Denham, 1998; National Research Council, Committee on Early Childhood Pedagogy, 2000; Shonkoff & Phillips, 2000) and their learning skills (Barriga et al., 2002; Duncan et al., 2007; Howse et al., 2003). In the model, Head Start participation influences the quality of a child's preschool education (Department of Health and Human Services, 2010a), which in turn influences developmental outcomes (Camilli et al., 2010; Currie & Thomas, 1995; Department of Health and Human Services, 2010a; Gorey, 2001; Karoly et al., 2005; McKey et al., 1985; Nelson et al., 2003), controlling for developmental status entering preschool and other risk factors and covariates. Elementary school quality also influences first grade outcomes (Betts & Shkolnik, 1999; Birch & Ladd, 1997; Finn & Achilles, 1999; Hamre & Pianta, 2001; Magnuson et al., 2007; Molnar et al., 1999; O'Connor & McCartney, 2007; Pianta et al., 2002; Pianta & Stuhlman, 2004). In turn, improved first grade outcomes influence adult outcomes (Armor, 2003; Berk, 2009; Heckman et al., 2006; La Paro & Pianta, 2000; Shonkoff & Phillips, 2000), which are included in this model for illustrative purposes despite being beyond the scope of this study.

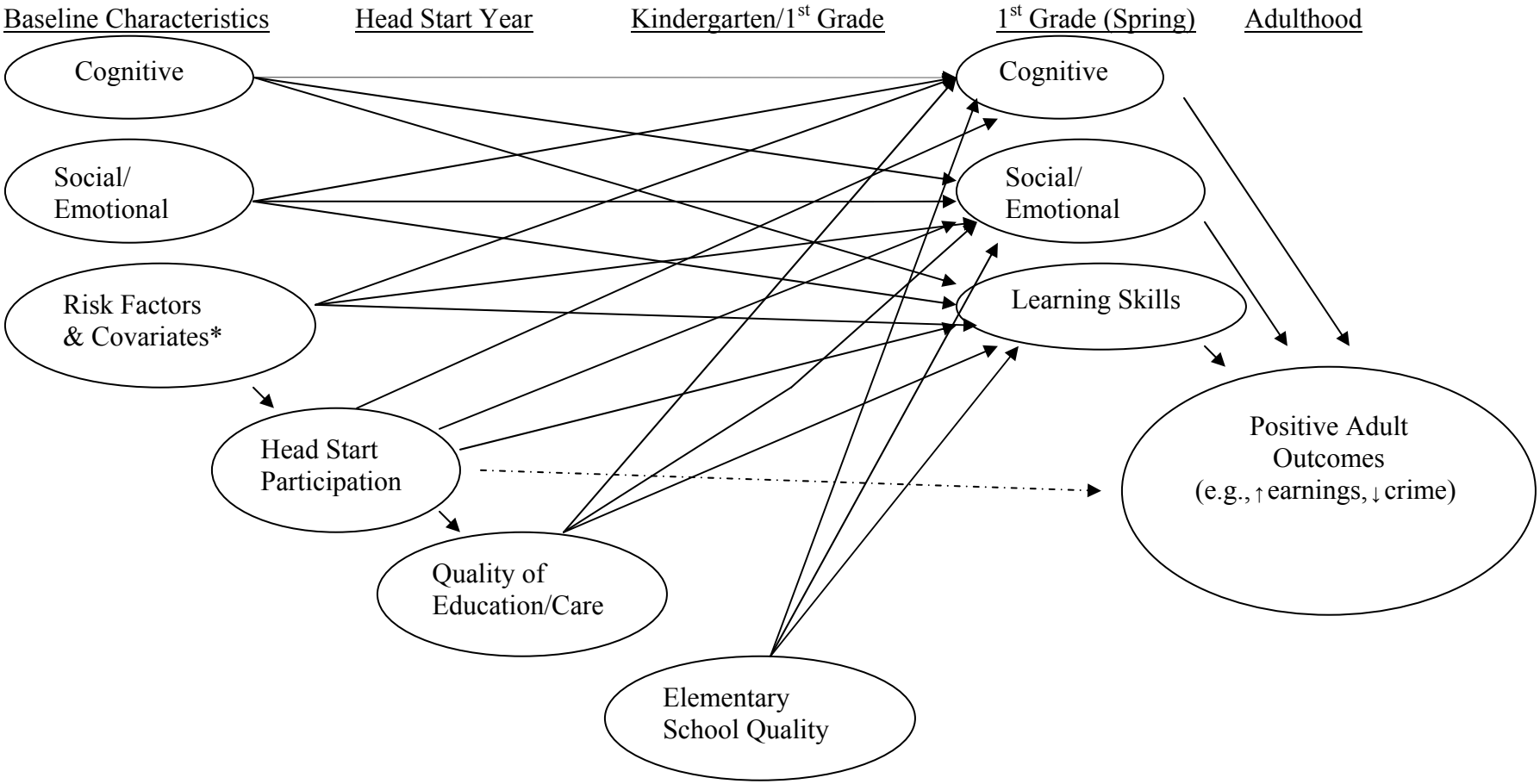

Figure 1: Conceptual Model

* For this study, the following risk factors and covariates are used: gender, race/ethnicity, primary language, language spoken at home, primary caregiver's age, number of parents in the home, mother's marital status, mother's status as a recent immigrant, being born to a teen mother, mother's level of education, urbanicity, and having a disability.

Sample

The analyses presented in the following chapters use data from the Head Start Impact Study. This study collected information on 4,667 children that applied to a nationally representative sample of Head Start programs in 2002 (Department of Health and Human Services, 2010a). Data on 4,442 of these children, excluding 225 from Puerto Rico, are available in a data set that is available to researchers who sign a data use agreement protecting participant confidentiality. A total of 383 representative Head Start centers were identified through a multi-stage sampling process that included stratification by geographic location, state early childhood policies, child race/ethnicity, and urbanicity. Within each of these centers, random assignment was used to establish a treatment group of children that were given access to Head Start services and a control group that was not granted such access. In the aggregate, the research team reported finding no statistically significant pre-existing differences between the treatment and control group on observed variables.

Baseline data on the children's development in various domains of cognitive and social/emotional development and their family background was collected in the fall of 2002 through parent interviews and independent child assessments. In the following chapters, these data are used as covariates in the analyses addressing each of the five research questions examined in this study. Parent interviews and independent child assessments were repeated each spring through the child's first grade year. During these subsequent rounds of data collection, data was also collected on the quality of the children's care and education environments through independent observations, interviews

with teachers and center directors, and secondary data sources. The present study uses data on the quality of preschool care and education to address the question of whether there is evidence that the quality of care and education provided in Head Start classrooms explains children's outcomes through first grade.

Data regarding children's early elementary classroom experiences is available from teacher reporting. These reports include such information as the number of children in the student's class and the number of years the teacher has taught. The present study uses this data to address the question of whether there is evidence that early elementary school experiences influence whether initial effects of Head Start endure.

Secondary data on the elementary schools attended by children in the sample is available through the Great Schools Database and the Private School Universe Survey. The Great Schools Database contains information on more than 90,000 public elementary, middle, and high schools in the United States, including charter schools, magnet schools, and alternative schools. The Private School Universe Survey, maintained by the National Center for Education Statistics, contains information on approximately 30,000 private elementary, middle, and high schools in the United States. In the present study, data from these sources regarding school proficiency in math and reading and the proportion of children in the school that are eligible for free or reduced price lunch are used to explore whether there is evidence that early elementary school experiences influence whether initial effects of Head Start endure.

Response rates vary by method of data collection, with most response rates being between 80 and 90 percent. Differences in response rates between the treatment and

control groups produced limited differences between the two groups on observable pre-existing characteristics.

Table 1 provides a summary of the primary care settings attended by children in the sample for the first year of the study, which is the only year for which access to Head Start was randomly assigned. Table 3 in the following section provides a summary of measures representing the quality of the preschool centers and elementary schools attended by children in the sample. Table 5, also in the following section, provides baseline demographic information for the sample.

Table 1: Primary Care Setting in First Year of Study, by Cohort and Treatment

	3-year-old cohort		4-year-old cohort	
	Treatment	Control	Treatment	Control
Covariate	*n = 1,464*	*n = 985*	*n = 1,182*	*n = 811*
Head Start	84%	18%	77%	14%
Center-based (non-Head Start)	7%	25%	11%	35%
Parent Care	7%	38%	9%	40%
Home-Based (non-parent)	2%	19%	3%	11%

Measures

Child Outcome Measures

Preacademic Skills: Independent assessors administered the Woodcock-Johnson III Tests of Achievement to each child in the fall of the baseline year and each spring thereafter through the child's first grade year. The analyses in the following chapters use the pre-academic skills cluster from these tests as a proxy for cognitive development. This cluster represents an overall measure that includes pre-reading and early math, writing, and spelling skills. The reliability for this measure is 0.97 or greater. In the typical administration of each of the subtests, children are given progressively difficult questions and their score is determined by how far they proceed before giving six consecutive incorrect responses. For the Head Start Impact Study, the tests were administered until a child provided three consecutive incorrect responses, in the interest of minimizing child discouragement and reducing the time required to administer the test. While this modification prevents the scores from being compared to national norms, it does not prevent comparisons between the treatment and control group. The preacademic cluster score represents the average of a child's scores on the Letter-Word Identification, Spelling, and Applied Problems subtests.

The W score of this measure, which is used for the analyses in the following chapters, represents how far the child proceeded in the test before giving three consecutive incorrect responses. This score has been calibrated on an equal-interval scale, meaning that any given interval represents the same difference in ability regardless of where it is on the scale, which makes it useful for analyzing an individual's growth

over time (Jaffe, 2009). This measure is centered on 500, which is the approximate average score for ten-year-olds on the tests as typically administered. Table 2 provides summary statistics for this measure. For both the three-year-old and four-year-old cohorts, the baseline differences in mean scores between the treatment and control groups are not statistically significant.

Social/emotional: As part of the study, parents rated their children on 14 problem behaviors that are indicative of aggressiveness, hyperactivity, and withdrawing. These ratings yielded an overall scale score that could vary from 0 to 28. Table 2 provides summary statistics for this measure, which is used in subsequent chapters to examine whether controlling for within-child variation facilitates the detection of previously unidentified social/emotional effects associated with being given access to Head Start and whether there is evidence of enduring effects associated with being given access to Head Start compared to a counterfactual of no preschool participation. For both the three-year-old and four-year-old cohorts, the baseline differences in mean scores between the treatment and control groups are not statistically significant.

Table 2: Summary of Child Preacademic Skills and Problem Behavior

		Treatment					Control				
Cohort	Measurement	n	m	sd	min	max	n	m	sd	min	max
3-yr old	Preacademic Skills	1,432	338	21	282	400	966	338	21	278	396
	Problem Behavior	1,464	6	4	0	20	985	6	4	0	19
4-yr old	Preacademic Skills	1,174	360	20	286	425	803	358	20	297	421
	Problem Behavior	1,182	6	4	0	21	811	6	4	0	20

For chi-square test for differences between treatment and control: * $p < .05$, ** $< .01$.

Learning Skills: The official report of the Head Start Impact Study did not address whether there is evidence of enduring effects associated with being given access to Head Start for measures of child attention, persistence, and confidence. In this study, I use the measures discussed in this subsection to conduct analyses to address this question.

As part of the Impact Study, the independent assessors rated children's attention, task persistence, and confidence during the assessment at each measurement period. For attention, the following ratings are used: easily distracted, some distraction with noise or movement, attends with assessor direction, and focuses attention voluntarily. For purposes of the present study, the two middle ratings are collapsed. For task persistence, the following ratings were used: refuses, attempts task after much encouragement, attempts task briefly, and persists with task. For purposes of the present study, the two middle ratings are collapsed to form a rating referred to as "attempts". For confidence, the following ratings are used: very uncertain, reluctant to try new or difficult things, attempts new things with encouragement, and very sure of self. Summary statistics for each of these measures, by cohort and treatment condition, are presented in Tables 12, 14, and 15.

In addition to these measures, the study collected three measures of child attention based on teacher and parent reporting. One of these measures represents the inattention/hyperactive scale from the Adjustment Scales for Preschool Intervention, which is based on teacher responses on ten items regarding the children's behavior over the past two months. Scores range from zero to 10, with higher scores meaning more inattentive/hyperactive behavior was reported for the child. For the next child attention

measure, the child's parent reported the degree to which it was true that their child usually cannot pay attention for long. The final child attention measure represents a parent report of whether the child's teacher had informed the parent that the child has difficulty concentrating. Summary statistics for each of these measures, by cohort and treatment condition, are presented in Tables 10, 11, and 13.

For child persistence, the independent assessment measure described above is the only measure of child persistence collected as part of the Head Start Impact Study. For child confidence, the only measure to supplement the independent assessment measure is one that reflects whether the child's parent reported that the child's teacher had communicated that the child lacks confidence in learning new things or taking part in new activities. Summary statistics for this measure, by cohort and treatment condition, is presented in Table 16.

Learning Environment Measures

Preschool: Trained independent assessors directly observed the quality of each child's primary care setting. In center-based settings, including Head Start, these assessors administered the Early Childhood Environment Rating Scale-Revised (ECERS-R), which captures the overall quality of the care environment and also has subscales for different aspects of quality. The score for this scale reflects the average of the assessors' ratings (1 to 7, with higher ratings representing higher quality) assigned on 43 items during their observations. Table 3 provides summary statistics for the overall scale by cohort and treatment group. This table shows that among children that attended center-based preschool, the mean quality, as measured by the ECERS-R, for the year of

treatment assignment was greater for those in the treatment group than for those in the control group. In this study, I use this measure to examine whether there is evidence that the quality of care and education provided in Head Start classrooms explains children's outcomes through first grade.

Elementary School: Elementary school quality was not assessed by direct observation as part of the Impact Study, but various proxies for school quality are available, including the percent of children in the school meeting state proficiency standards for math and reading, the percent of students in the school that are eligible for free or reduced price lunch, the size of the child's class, and the number of years the teacher has taught. Summary statistics presented in Table 3 indicate that differences in the elementary classes/schools attended by children in the treatment and control groups were not statistically significant.

Table 3: Preschool and Elementary Quality, by Cohort and Treatment

		Treatment			Control		
Cohort	Measurement	n	m	sd	n	m	sd
3-yr old	Preschool (ECERS-R)						
	Two Years Prior to Kindergarten	1,072	5.1 ***	1.0	234	4.7 ***	1.1
	Year Prior to Kindergarten	1,028	4.9	1.0	574	4.8	1.0
	Kindergarten						
	Teacher's Years of Experience	1,048	13.3	10	647	13	10
	School Math Proficiency %	898	66	22	553	67	20
	School Reading Proficiency %	895	64	26	552	65	25
	School Free/Reduced Lunch %	906	69	26	562	69	25
	Number of Students in Child's Class	1,055	20	5	655	20	5
	First Grade						
	Teacher's Years of Experience	1,051	13	10	674	13	10
	School Math Proficiency %	974	68	22	622	68	21
	School Reading Proficiency %	971	65	25	622	66	23
	School Free/Reduced Lunch %	934	67	26	585	66	26
	Number of Students in Child's Class	1,058	20	4	679	20	4
4-yr old	Preschool (ECERS-R)	854	5.3 ***	0.9	218	4.6 ***	1.1
	Kindergarten						
	Teacher's Years of Experience	645	14	10	394	14	10
	School Math Proficiency %	603	65	22	368	65	23
	School Reading Proficiency %	602	57	28	368	59	28
	School Free/Reduced Lunch %	575	66	26	362	66	25
	Number of Students in Child's Class	626	21	5	388	21	5
	First Grade						
	Teacher's Years of Experience	776	14	10	495	13	10
	School Math Proficiency %	709	67	20	456	67	20
	School Reading Proficiency %	703	61	28	453	62	27
	School Free/Reduced Lunch %	708	66	26	433	65	26
	Number of Students in Child's Class	787	20	4	500	20	4

For chi-square test for differences between treatment and control: * p < .05, ** < .01, *** < .001.

Covariates

Table 4 provides a description of the covariates used in the regression analyses presented in subsequent chapters. This list of covariates includes those used in analyses for the official report of the Head Start Impact Study (Department of Health and Human Services, 2010a) plus measures of urbanicity and whether the child's parent reported they have been told that their child has "special needs". The process for selecting the initial list of covariates for the study involved identifying likely predictors of child outcomes based on past research and removing or collapsing covariates that were highly correlated with other covariates and/or had small cell sizes.

Table 4: Description of Covariates

Covariate	Description	Source
Female	Dummy (Female = 1)	parent report
Black	Dummy (Black = 1)	parent report
Hispanic	Dummy (Hispanic = 1)	parent report
Spanish Primary Language	Dummy (Yes = 1)	assessment
Spanish Spoken at Home	Dummy (Yes = 1)	parent report
Age of Primary Caregiver	Continuous variable. Caregiver age in years.	parent report
Live with Both Parents	Dummy (Yes = 1)	parent report
Mom is Married	Dummy (Yes = 1)	parent report
Mom Recent Immigrant	Dummy (Yes = 1)	parent report
Born to Teen Mother	Dummy (Yes = 1)	parent report
Mother's Education	Ordinal variable. < HS = 1; HS or GED = 2, > HS = 3	parent report
Urban	Dummy (Urban = 1)	geo-coding
"Special Needs"	Dummy (Doctor said child has "special needs" = 1)	parent report

Table 5 provides a summary of baseline characteristics for these covariates by treatment condition for each cohort. With two exceptions, there are not statistically significant differences between the treatment and control groups on these measures. For the three-year-old cohort, 13 percent of the parents of treatment group children reported having been told that their child has "special needs" compared to 10 percent for the control group. For this same cohort, 18 percent of the children in the treatment group were born to a teenage mother compared to 14 percent for the control group.

The child assessments and parent interviews were not conducted at the exact same time for each of the children. For this reason, the multivariate analyses in the following chapters also include a covariate reflecting the number of weeks from the baseline date of September 1, 2002 to the date that the assessment or interview was completed.

Table 5: Baseline Covariate Summary, by Cohort and Treatment

	3-year-old cohort		4-year-old cohort	
	Treatment	Control	Treatment	Control
Covariate	*n = 1,464*	*n = 985*	*n = 1,182*	*n = 811*
Female	51%	49%	48%	49%
Black	37%	35%	23%	25%
Hispanic	34%	32%	43%	43%
Spanish Primary Language	22%	20%	32%	31%
Spanish Spoken at Home	27%	26%	33%	35%
Age of Primary Caregiver	29 (8)	28 (7)	29 (7)	29 (7)
Live with Both Parents	49%	48%	51%	51%
Mother is Married	43%	44%	46%	46%
Mother is Recent Immigrant	17%	16%	23%	22%
Born to Teen Mother	14%**	18%**	18%	18%
Mother's Education:				
Less than High School	34%	36%	42%	43%
High School or GED	36%	34%	31%	32%
Beyond High School	31%	30%	27%	26%
Urban (vs. Rural)	83%	82%	86%	87%
"Special Needs"	13%*	10%*	15%	13%

For chi-square test for differences between treatment and control: * $p < .05$, ** $< .01$.

Table 6 provides a summary of the average number of weeks from baseline for child assessments and parent interviews by cohort and treatment group. For both cohorts, for the spring assessment of the initial year of the study, children in the control group on average had later assessments and parent interviews than the treatment group.

Table 6: Child Assessment and Parent Interview Timing (weeks from 9/1/2002)

		Treatment			Control		
Cohort	Measurement	n	m	sd	n	m	sd
3-yr old	Assessment						
	spring of pre-K	1,299	32 ***	3	771	34 ***	3
	spring of next pre-K year	1,276	85	3	779	85	3
	spring of kindergarten	1,195	138	4	732	138	4
	spring of first grade	1,165	189	3	711	189	3
	Parent Interview						
	spring of pre-K	1,464	34 ***	4	985	35 ***	5
	spring of next pre-K year	1,264	86	4	769	86	4
	spring of kindergarten	1,238	138	4	761	138	3
	spring of first grade	1,219	189	3	739	189	3
4-yr old	Assessment						
	spring of pre-K	1,024	33 ***	3	614	34 ***	3
	spring of kindergarten	954	86	3	581	86	3
	spring of first grade	945	138	4	582	138	4
	Parent Interview						
	spring of pre-K	1,182	35 ***	4	811	36 ***	5
	spring of kindergarten	1,156	86	3	592	86	3
	spring of first grade	973	138	4	604	138	4

For chi-square test for differences between treatment and control: * p < .05, ** < .01, *** < .001.

Missing Values

As provided, the Head Start Impact Study data set contains imputed values in place of otherwise missing data for the covariates presented in Tables 4 and 5. These imputed values were derived through a hot-deck procedure by which a replacement value was randomly selected from a group of similar children identified using characteristics correlated with the variable being imputed (Department of Health and Human Services,

2010c). As reflected in Table 7, no more than 18 percent of the values for any covariate have been imputed.

Table 7: Imputed Values by Covariate

Measure	Total	Imputed	
		#	%
Covariates:			
Female	4,442	2	0.05
Black	4,442	45	1
Hispanic	4,442	45	1
Spanish Primary Language	4,442	56	1
Spanish Spoken at Home	4,442	56	1
Age of Primary Caregiver	4,442	35	1
Live with Both Parents	4,442	804	18
Mother is Married	4,442	696	16
Mother is Recent Immigrant	4,442	275	6
Born to Teen Mother	4,442	738	17
Mother's Education	4,442	70	2
Urban (vs. Rural)	4,442	--	--
"Special Needs"	4,442	663	15

Methods

Regression Analysis

The following chapters present the results of multiple ordinary least squares, logistic, ordered logistic, and cross-sectional time-series regression analyses conducted with Stata. Regression analysis is used to answer all five of the research questions addressed in this study because it facilitates the explanation of a single quantitative dependent variable from multiple quantitative independent variables, including dichotomous measures of qualitative variables such as gender and race (Kirk, 1994). Unlike analysis of variance, regression analysis utilizes the magnitude of differences in scores for continuous independent variables in explaining the dependent variable. Each of the following chapters introduces the regression equations used for the analyses therein.

Sampling Weights

The Head Start Impact Study data set contains sampling weights for use in producing estimates representative of the national population of newly entering Head Start participants in 2002 *(*Department of Health and Human Services, 2010c*)*. These weights reflect any given child's probability for being selected for the study, adjusted for nonresponse rates for each method of data collection. Weights corresponding with the dependent variable's data collection method and measurement period were applied for all of the cross-sectional regression analyses presented in the following chapters.

Accounting for Clustering

Observations collected through the Head Start Impact Study do not represent an independent sample, as these observations are clustered within those Head Start centers and other care and education settings attended by children selected for participation in the study and in those elementary schools subsequently attended. As a result, significance tests for relationships between variables derived from these observations are prone to higher chances of Type I error (i.e., rejecting the null hypothesis of no relationship), to the extent that there is a stronger correlation within these groups than across them. For the data used in this study, the variable representing the precise center the child attended has a large proportion of missing values (50% missing for the three-year-old cohort, 46% missing for the three-year-old cohort).[2] For this reason, I have not accounted for clustering by center. Given the resulting potential for bias toward Type I error, this is important context for any findings of statistical significance.

Propensity Score Analysis

Chapter 8 compares the outcomes of children that attended Head Start with those of comparable children that did not attend any preschool. The Head Start Impact Study was not designed to produce such groups of children, but a statistical technique known as propensity score analysis can be used to generate such groups (Gertler, Martinez, Premand, Rawlings, & Vermeersch, 2010; Guo & Fraser, 2009). Chapter 8 introduces this technique and presents the results of propensity score analysis used in conjunction with regression analysis.

[2] Data regarding the characteristics of these care settings is available for most of these observations, even though an identification number for the care setting is not available.

CHAPTER FOUR: PREACADEMIC SKILLS AND PROBLEM BEHAVIOR

The analyses presented in this chapter address the first research question presented in Chapter One: *Does controlling for within-child variation facilitate the detection of previously unidentified effects associated with being given access to Head Start?* The data collected through the Head Start Impact Study represents time-series cross-sectional data, in that measurements on the same variables were collected for the same children and their families at multiple points in time. One way to assess whether there are effects associated with giving children access to Head Start is to analyze the children's outcomes one year at a time, controlling for baseline status for these outcomes and observable covariates. This approach was used to produce the vast majority of the results presented in the official report of the Head Start Impact Study (Department of Health and Human Services, 2010a). A second approach is to analyze the data as a panel, which can enhance one's ability to detect effects by controlling for within-child variation in the interest of reducing overall unexplained variation (Wooldridge, 2008). For select cognitive outcomes, the official report of the Head Start Impact Study reports the results of such analysis for the period from the beginning of the study to the first grade measurement as a whole, but does not report the statistical significance of any potential treatment effects for each intervening year. The analyses presented in this chapter examine whether, for each intervening year, such an approach detects previously

unidentified effects associated with access to Head Start for a measure of preacademic skills and a measure of problem behavior.

The following cross-sectional time-series random-effects regression is used to estimate the effect of access to Head Start on preacademic skills and problem behavior over time:

$$OUTCOME_{it} = \beta_{1+}\beta_2 HS_i + \beta_3 INT_{it+}\beta_4 COV_i + T_t + \varepsilon_{it}, \quad (1)$$

where $OUTCOME_{it}$ is a measure of the outcome (preacademic skills or problem behavior) for child *i* at measurement time *t*, HS_i is a dummy variable representing assignment to the treatment group given access to Head Start, INT_i is a term interacting access to Head Start with time, COV_i is a set of covariates measuring the child, family, and assessment timing characteristics presented in Table 4, and T_t is the effect of time. The regression coefficient β_3 estimates the effect of access to Head Start for each measurement time.

For both the three-year-old and four-year-old cohorts, the official report of the Head Start Impact Study included results from annual cross-sectional analyses that indicated positive effects for preacademic skills for the spring of the Head Start year, but not for any subsequent measurement period through first grade. Results in the official report from longitudinal analysis controlling for within-child variation did not indicate a statistically significant difference between the treatment and control groups for preacademic skills for the period from baseline to the spring of first grade as a whole. For problem behavior, the annual cross-sectional results in the official report did not reveal a positive treatment effect for problem behavior for either cohort at any

measurement period. The official report did not include results for analyses controlling for within-child variation for problem behavior.

Results

Preacademic Skills

Table 8 contains the results from cross-sectional time-series random-effects regression for preacademic skills. There is not a statistically significant first order effect associated with being given access to Head Start, for either cohort. Time itself (year of measurement) does have strong effects, reflecting a developmental impact as students grow older over the duration of study.

More important, there is a statistically significant interaction between the treatment condition and the measurement period for the spring of the Head Start year for both cohorts (3.8 and 3.0 points, respectively). The finding of positive effects for preacademic skills for the spring of the Head Start year is consistent with the findings in the official report of the Head Start Impact Study. In addition, analyzing this data as a panel reveals a statistically significant treatment effect of 2.2 points received by the three-year-old cohort in their post-Head Start year prior to starting Kindergarten. This effect was not reported in the original study. Using a fixed effects model produces the same results.[3]

These single-digit positive effects associated with access to Head Start are small in comparison to the double-digit developmental and schooling effects on the order of 25 to 30 points for the kindergarten and first grade years. Accordingly, the modest gains

[3] The coefficient for the second pre-kindergarten year effects for 3-year olds changed to 2.35 ($p = 0.016$) from 2.20 ($p = 0.023$).

of Head Start may be overwhelmed by typical cognitive growth during the early school years.

Table 8: Predictors of Preacademic Skills [a b c]

	3-yr old cohort		4-yr old cohort	
Variable	Coefficient	P-Value	Coefficient	P-Value
Treatment	0.78	0.422	1.24	0.270
Spring of Head Start Year	2.37	0.002 **	1.54	0.068
Spring of Next Pre-K Year	11.77	< 0.001 ***	n/a	n/a
Spring of Kindergarten Year	35.29	< 0.001 ***	34.88	< 0.001 ***
Spring of First Grade Year	52.67	< 0.001 ***	60.72	< 0.001 ***
Treatment*Year (Head Start)	3.76	< 0.001 ***	2.97	0.005 **
Treatment*Year (Next Pre-K)	2.20	0.023 *	n/a	n/a
Treatment*Year (K)	-0.50	0.614	-1.53	0.164
Treatment*Year (G1)	-0.82	0.412	-0.51	0.642
Female	7.21	< 0.001 ***	5.25	< 0.001 ***
Black	-1.32	0.168	-4.06	0.001 **
Hispanic	-1.52	0.222	-2.43	0.105
Spanish Primary Language	0.31	0.856	3.59	0.100
Spanish Spoken at Home	0.84	0.593	-1.41	0.518
Age of Primary Caregiver	0.08	0.156	0.03	0.632
Live with Both Parents	-0.21	0.837	1.52	0.205
Mother is Married	3.12	0.002 **	-0.39	0.744
Mother is a Recent Immigrant	0.81	0.540	-2.27	0.109
Born to Teen Mother	0.66	0.552	-0.33	0.800
Mother's Education	3.54	< 0.001 ***	4.32	< 0.001 ***
Urban	-0.38	0.699	0.21	0.873
"Special Needs"	-4.91	< 0.001 ***	-7.69	< 0.001 ***
Testing Lag (week)	0.37	< 0.001 ***	0.24	0.001 ***
Constant	311.69	< 0.001 ***	341.50	< 0.001 ***

[a] Results from cross-sectional time-series random-effects regression.

[b] For 3-yr old cohort, 9,861 observations and 2,223 groups. For 4-yr old cohort, 6,306 observations and 1,740 groups.

[c] * p < .05, ** < .01, *** < .001.

Table 8 also reveals important influences of demographic and socioeconomic covariates, which are routinely found in studies of achievement. For both cohorts, being a female, having a more educated mother, and the absence of "special needs" are each statistically significant predictors of better preacademic skill scores, controlling for the other variables in the model. In addition, having a married mother is associated with a higher score for the three-year-old cohort, and being Black is associated with a lower score for the four-year old cohort. Not surprisingly, the coefficient for the testing lag variable is also statistically significant for both cohorts, indicating that an additional week of maturation and learning facilitates slightly higher scores.

Problem Behavior

Table 9 contains the results from the previously discussed cross-sectional time-series random-effects model for problem behavior. As with preacademic skills, for problem behavior there is not a statistically significant first order effect associated with being given access to Head Start, for either cohort. Unlike preacademic skills, for problem behavior there is not a statistically significant change over time with the exception of the four-year-old cohort for which there is a modest reduction in problem behavior associated with proceeding from the baseline to the end of the Head Start year. Using a fixed effects model produces the same results.

Table 9: Predictors of Problem Behavior [a b c]

	3-yr old cohort		4-yr old cohort	
Variable	Coefficient	P-Value	Coefficient	P-Value
Treatment	-0.18	0.270	-0.30	0.091
Spring of Head Start Year	0.03	0.800	-0.44	0.003 **
Spring of Next Pre-K Year	0.81	0.139	n/a	n/a
Spring of Kindergarten Year	1.53	0.155	0.21	0.748
Spring of First Grade Year	2.99	0.062	1.79	0.171
Treatment*Year (Head Start)	-0.32	0.068	0.17	0.359
Treatment*Year (Next Pre-K)	-0.27	0.129	n/a	n/a
Treatment*Year (K)	0.01	0.959	0.47	0.014 *
Treatment*Year (G1)	-0.11	0.522	0.31	0.106
Female	-0.83	< 0.001 ***	-0.99	< 0.001 ***
Black	-0.61	< 0.001 ***	-1.00	< 0.001 ***
Hispanic	-0.22	0.269	-0.39	0.088
Spanish Primary Language	1.06	< 0.001 ***	0.79	0.016 *
Spanish Spoken at Home	0.50	0.044 *	-0.07	0.841
Age of Primary Caregiver	0.00	0.966	0.00	0.648
Live with Both Parents	-0.45	0.004 **	-0.37	0.043 *
Mother is Married	-0.28	0.077	-0.25	0.172
Mother is a Recent Immigrant	0.73	< 0.001 ***	0.51	0.019 *
Born to Teen Mother	-0.08	0.626	0.07	0.709
Mother's Education	-0.51	< 0.001 ***	-0.45	< 0.001 ***
Urban	-0.48	0.002 **	0.47	0.015 *
"Special Needs"	1.30	< 0.001 ***	2.16	< 0.001 ***
Testing Lag (week)	-0.03	0.008 **	-0.03	0.015 *
Constant	8.72	< 0.001 ***	8.20	< 0.001 ***

[a] Results from cross-sectional time-series random-effects regression.

[b] For 3-yr old cohort, 9,839 observations and 2,223 groups. For 4-yr old cohort, 6,268 observations and 1,739 groups.

[c] * p < .05, ** < .01, *** < .001.

For the three-year-old cohort, the interaction between treatment condition and year of measurement does not have a statistically significant coefficient for any

measurement period. This is largely consistent with the annual cross-sectional results in the official report of the Head Start Impact Study, which had only found a small but statistically significant difference for the spring of the Head Start year indicating more problem behavior for the treatment group. For the four-year-old cohort, there is a statistically significant interaction between the treatment condition and the measurement period for the spring of the kindergarten year indicating more problem behavior for the treatment group. No such relationship was found for the prior or subsequent year. The cross-sectional results in the official report of the Head Start Impact Study had not found any statistically significant differences for this measure.

Several covariates also have statistically significant effects. For both cohorts, lower problem behavior scores are associated with being a female, being Black, not having Spanish as a primary language, living with both parents, not having a mother that is a recent immigrant, having a more educated mother, and not having "special needs". In addition, for the three-year-old cohort, having Spanish spoken at home is associated with more problem behavior. Being in an urban area is associated with less problem behavior for the three-year-old cohort, but associated with more problem behavior for the four-year-old cohort.

The coefficient for the testing lag variable is also statistically significant for both cohorts, indicating that an additional week of maturation and learning is associated with less problem behavior.

Discussion

The analyses presented in this chapter examined whether controlling for within-child variation facilitates the detection of effects associated with being given access to Head Start, for each incremental year. Consistent with the official report of the Head Start Impact Study, this analysis found favorable effects at the end of the Head Start year for preacademic skills (both cohorts) but not for problem behavior. In addition, for the three-year-old cohort this analysis found a previously unidentified positive effect for preacademic skills for the treatment group that was given access to Head Start, in the spring of their post-Head Start year prior to starting Kindergarten. However, this method did not identify favorable effects for preacademic skills or problem behavior in kindergarten or first grade for either cohort. To the contrary, for the four-year-old cohort a statistically significant unfavorable effect was found for problem behavior for the spring of the kindergarten year. This unfavorable effect was not found for the prior or subsequent year.

CHAPTER FIVE: LEARNING SKILLS

Evaluations of Head Start have tended to focus on the impact of the program on cognitive development, and given less attention to noncognitive skills that are important for future success (Currie & Thomas, 1995; Garces et al., 2002). While the official report of the Head Start Impact Study reports impacts for cognitive as well as social/emotional outcomes, it does not indicate whether there are statistically significant differences between the treatment and control groups for measures of child attention, persistence, and confidence, for which measures are available through the study. Theory and empirical evidence suggest these behaviors influence subsequent cognitive ability, achievement, school completion, and behavior (Heckman et al., 2006).

This chapter presents the results of analyses examining whether there is evidence of enduring effects associated with access to Head Start for outcomes not addressed in the official report of the Head Start Impact Study. Child attention, persistence, and confidence outcomes, as measured at the end of the Head Start year and each spring thereafter through first grade, are examined.

The results include comparisons of mean or percentage differences between treatment and control groups, with and without controls, with appropriate significance tests. Categorical (dichotomous and ordinal) measures were analyzed using chi-square tests for association, as well as ordered or logistic regression to control for covariates.

The lone continuous measure examined in this section was analyzed using t-tests, as well as multiple linear regression controlling for the covariates identified in Table 4.

For the six ordinal dependent variables examined in this chapter, the following regression model is used to estimate the effect of access to Head Start for each post-baseline measurement period:

$$SKILL_i = \beta_1 + \beta_2 HS_i + \beta_3 SKILL_{ib} + \beta_4 COV_i + \varepsilon_i, \quad (2)$$

where $SKILL_i$ is a measure of a given learning skill for child *i*, HS_i is a dummy variable representing assignment to the treatment group given access to Head Start, $SKILL_{ib}$ is a measure of a given learning skill for child *i* at baseline (represented by a separate dummy variable for each ordered level), and COV_i is a set of covariates measuring child, family, and assessment timing characteristics presented in Table 4. The regression coefficient β_2 estimates the effect of access to Head Start for each measurement time.

For the one continuous variable examined in this chapter, for which a baseline measurement is not available, the following regression model is used to estimate the effect of access to Head Start for each post-baseline measurement period:

$$ATTENT_i = \beta_1 + \beta_2 HS_i + \beta_3 COV_i + \varepsilon_i, \quad (3)$$

where $ATTENT_i$ is a measure of child attention for child *i*, HS_i is a dummy variable representing assignment to the treatment group given access to Head Start, and COV_i is a set of covariates measuring child, family, and assessment timing characteristics presented in Table 4. The regression coefficient β_2 estimates the effect of access to Head Start for each measurement time. It would be preferable to control for a baseline

measurement of this variable, but in the absence of data for such a measure it is assumed that random assignment produced treatment and control groups with similar baseline characteristics.

Results

Child Attention

Four measures of child attention were collected as part of the Head Start Impact Study. One of these measures was based on teacher assessments, two were reported by the child's parent, and one was reported by independent assessors. These variables are not aggregated into a single composite measure for the following analyses, given the substantial differences in content, evaluators, and scoring methods across measures.

Table 10 presents the mean scores, by treatment group, of the attention measure that was based on teacher assessments as well as a t-test for differences in these means before controlling for any covariates. Table 10 also shows the results of a multiple regression controlling for the covariates identified above; the standardized regression coefficients (β's) along with the p-values for these coefficients are reported.
No statistically significant differences were found between the treatment and control groups at any measurement period, for either cohort, under either analysis method.

Additional regression analyses, reported in Table 33 in the Appendix, in which the treatment variable was replaced with a variable that reflects participation in Head Start regardless of treatment assignment, did not result in statistically significant differences between the treatment and control groups at any measurement period, for either cohort.

Table 10: Child Attention (teacher report), by Treatment [b c]

Cohort	Measurement Period	Mean		T-test	Control for Covariates [a]	
		Control	Treatment		Beta	Sig.
3-yr old	spring of pre-K	1.89	1.88	p=.463	< 0.01	p=.937
	spring of next pre-K year	1.74	1.69	p=.336	< 0.01	p=.876
	spring of K	1.87	1.84	p=.399	< 0.01	p=.928
	spring of G1	1.89	1.87	p=.443	< 0.01	p=.914
4-yr old	spring of pre-K	1.68	1.66	p=.439	0.01	p=.782
	spring of K	1.81	1.70	p=.224	-0.02	p=.441
	spring of G1	1.75	1.73	p=.449	-0.01	p=.825

[a] Results from ordinary least squares linear regression.

[b] For treatment and control groups for both cohorts, standard deviations: $1.98 < sd < 2.33$.

[c] For 3-yr old cohort, treatment $1{,}022 < n < 1{,}076$, control $265 < n < 667$; for 4-yr old cohort, treatment $677 < n < 854$, control $241 < n < 495$.

Table 11 presents the results of analysis of a second child attention measure. This measure represents parents' assessment of whether their child usually cannot pay attention for long. For the uncontrolled comparisons, a chi square test is used for detecting significant differences between treatment and control groups. The covariate analysis utilizes ordered logistic regression.

Table 11: Child Attention (parent report), by Treatment [b]

Cohort	Measurement Period	Rating	Control %	Treatment %	Chi Square	Control for Covariates [a]	
						Odds Ratio	Sig.
3-yr old	fall of pre-k (baseline)	Very True	17	13			
		Sometimes True	45	46			
		Not True	37	41	**p=.048**		
	spring of pre-K	Very True	18	13			
		Sometimes True	49	44			
		Not True	33	43	**p<.001**	**1.51**	**p=.001**
	spring of next pre-K year	Very True	13	10			
		Sometimes True	43	43			
		Not True	44	47	p=.152	1.05	p=.708
	spring of K	Very True	12	11			
		Sometimes True	36	34			
		Not True	52	55	p=.448	1.27	p=.061
	spring of G1	Very True	12	11			
		Sometimes True	38	36			
		Not True	50	53	p=.420	1.02	p=.903
4-yr old	fall of pre-K (baseline)	Very True	15	13			
		Sometimes True	45	43			
		Not True	40	44	p=.172		
	spring of pre-K	Very True	14	12			
		Sometimes True	45	44			
		Not True	42	44	p=.415	1.17	p=.254
	spring of K	Very True	8	10			
		Sometimes True	34	37			
		Not True	58	54	p=.274	0.85	p=.263
	spring of G1	Very True	8	10			
		Sometimes True	37	36			
		Not True	54	54	p=.553	0.91	p=.523

[a] Results from ordered logistic regression.

[b] For 3-yr old cohort, treatment 1,215 < n < 1,276, control 733 < n < 784; for 4-yr old cohort, treatment 962 < n < 1,006, control 592 < n < 622.

For the cohort that entered the study as four-year-olds, no statistically significant differences were found between the treatment and control groups. For the cohort that entered the study as three-year-olds, despite random treatment assignment, a statistically significant difference was found between the treatment and control groups at baseline,

with parents of children in the treatment group reporting better attention among their children. Statistically significant differences were only found to endure through the spring of the Head Start year. At that point, children in the treatment group had 51 percent greater odds of their parent reporting they did not have problems paying attention during the spring of the Head Start year, controlling for covariates.

As reflected in Table 33 in the Appendix, an additional regression analysis differing from the analysis described above only in that the treatment variable was replaced with a variable reflecting participation in Head Start regardless of treatment assignment also yielded a statistically significant difference ($p = .007$) for the three-year-old cohort in the spring of the Head Start year. In this analysis, children in the treatment group had 41 percent greater odds of their parent reporting they did not have problems paying attention. This only occurred for the three-year-old cohort, and did not occur during the subsequent three spring follow-ups.

For the other child attention measures, one representing three ordinal categories of attention reported by an independent assessor and the other representing a parent report of whether the child's teacher had informed the parent that the child has difficulty concentrating, neither chi-square tests nor regression analyses produced statistically significant differences between the treatment and control groups at any measurement period, for either cohort. The results from these tests are presented in Tables 12 and 13 respectively.

Table 12: Child Attention (independent assessment), by Treatment [b]

Cohort	Measurement Period	Rating	Control %	Treatment %	Chi Square	Control for Covariates [a]	
						Odds Ratio	Sig.
3-yr old	fall of pre-k (baseline)	Easily Distracted	22	21			
		Some Distraction	49	47			
		Focuses Attention	29	32	p=.455		
	spring of pre-K	Easily Distracted	18	17			
		Some Distraction	47	48			
		Focuses Attention	35	35	p=.731	0.99	p=.928
	spring of next pre-K year	Easily Distracted	10	10			
		Some Distraction	44	44			
		Focuses Attention	46	46	p=.960	1.12	p=.385
	spring of K	Easily Distracted	7	7			
		Some Distraction	35	32			
		Focuses Attention	58	62	p=.298	1.15	p=.332
	spring of G1	Easily Distracted	5	6			
		Some Distraction	28	25			
		Focuses Attention	68	69	p=.431	1.07	p=.650
4-yr old	fall of pre-K (baseline)	Easily Distracted	8	7			
		Some Distraction	35	40			
		Focuses Attention	57	52	p=.089		
	spring of pre-K	Easily Distracted	8	8			
		Some Distraction	34	36			
		Focuses Attention	57	56	p=.756	1.01	p=.922
	spring of K	Easily Distracted	3	4			
		Some Distraction	24	25			
		Focuses Attention	73	70	p=.201	1.17	p=.373
	spring of G1	Easily Distracted	4	4			
		Some Distraction	24	23			
		Focuses Attention	72	73	p=.970	1.00	p=.992

[a] Results from ordered logistic regression.

[b] For 3-yr old cohort, treatment $1,253 < n < 1,280$, control $704 < n < 772$; for 4-yr old cohort, treatment $933 < n < 1,020$, control $568 < n < 602$.

Table 13: Parent Reports Teacher Indicated Child Lacks Concentration, by Treatment [b]

Cohort	Measurement Period	Parent Reports Teacher Has Told Them Child Can't Concentrate	Control %	Treatment %	Chi Square	Control for Covariates [a]	
						Odds Ratio	Sig.
3-yr old	spring of K	No	71	70			
		Yes	29	30	p=.538	1.06	p=.692
	spring of G1	No	66	66			
		Yes	34	34	p=.979	1.05	p=.719
4-yr old	spring of K	No	74	70			
		Yes	26	30	p=.117	1.35	p=.068
	spring of G1	No	71	70			
		Yes	29	30	p=.578	1.03	p=.846

[a] Results from logistic regression.

[b] For 3-yr old cohort, treatment 1,215 < n < 1,225, control 736 < n < 751; for 4-yr old cohort, treatment 940 < n < 968, control 578 < n < 599.

Child Persistence

Only one measure of child persistence was collected as part of the Head Start Impact Study. This measure was reported by independent assessors for the fall of the first year of the study and each spring thereafter through first grade. Table 14 presents the percentage of children rated in each of three categories reported by the independent assessors, by treatment group, as well as the results of a chi-square test for association between treatment group assignment and this measure. Table 14 also presents the p-value for the regression coefficient for the treatment group assignment variable, after controlling for covariates.

Table 14: Child Persistence (independent assessment), by Treatment [b]

Cohort	Measurement Period	Rating	Control %	Treatment %	Chi Square	Control for Covariates [a]	
						Odds Ratio	Sig.
3-yr old	fall of pre-k (baseline)	Refuses	5	4			
		Attempts	54	53			
		Persists	41	43	p=.569		
	spring of pre-K	Refuses	4	4			
		Attempts	50	50			
		Persists	46	47	p=.941	1.00	p=.991
	spring of next pre-K year	Refuses	2	1			
		Attempts	38	37			
		Persists	61	62	p=.896	1.00	p=.992
	spring of K	Refuses	1	1			
		Attempts	34	30			
		Persists	65	69	p=.174	1.02	p=.914
	spring of G1	Refuses	1	1			
		Attempts	26	27			
		Persists	73	72	p=.429	0.86	p=.335
4-yr old	fall of pre-K (baseline)	Refuses	2	2			
		Attempts	30	32			
		Persists	68	67	p=.833		
	spring of pre-K	Refuses	2	1			
		Attempts	32	32			
		Persists	66	67	p=.484	1.17	p=.361
	spring of K	Refuses	0.4	1			
		Attempts	22	25			
		Persists	78	74	p=.288	0.84	p=.375
	spring of G1	Refuses	0.2	0.2			
		Attempts	25	23			
		Persists	75	77	p=.778	1.08	p=.662

[a] Results from ordered logistic regression.

[b] For 3-yr old cohort, treatment $1{,}155 < n < 1{,}282$, control $702 < n < 773$; for 4-yr old cohort, treatment $933 < n < 1{,}021$, control $566 < n < 603$.

For this measure, no statistically significant differences were found between the treatment and control groups, with or without controlling for covariates. Additional regression analyses reported in Table 33 of the Appendix, in which the treatment variable was replaced with a variable that reflects participation in Head Start regardless of treatment assignment, did not find statistically significant differences between the treatment and control groups at any measurement period, for either cohort.

Child Confidence

Two measures of child confidence were collected through the Head Start Impact Study. One of these measures was reported by independent assessors for the fall of the first year of the study and each spring thereafter through first grade. The other measure reflected whether the child's parent reported that the child's teacher had communicated that the child lacks confidence in learning new things or taking part in new activities.

Table 15 presents the percentage of children rated in each of four categories reported on by independent assessors, as well as the results of a chi-square test for association between treatment group assignment and this measure. Table 15 also presents the p-value for the regression coefficient for the treatment assignment variable, controlling for the covariates introduced above. For this measure, no statistically significant differences were found between the treatment and control groups at any measurement period, for either cohort. Additional regression analyses reported in Table 33 in the Appendix, in which the treatment variable was replaced with a variable reflecting participation in Head Start, did not result in statistically significant differences between the treatment and control groups at any measurement period, for either cohort.

Table 15: Child Confidence (independent assessment), by Treatment [b]

Cohort	Measurement Period	Rating	Control %	Treatment %	Chi Square	Control for Covariates [a]	
						Odds Ratio	Sig.
3-yr old	fall of pre-k (baseline)	Very Uncertain	12	12			
		Reluctant	17	15			
		Confident	43	42			
		Very Sure	28	31	p=.420		
	spring of pre-K	Very Uncertain	12	12			
		Reluctant	13	14			
		Confident	42	42			
		Very Sure	33	32	p=.927	1.11	p=.403
	spring of next pre-K year	Very Uncertain	8	6			
		Reluctant	13	12			
		Confident	42	45			
		Very Sure	37	37	p=.316	1.00	p=.983
	spring of K	Very Uncertain	4	5			
		Reluctant	12	11			
		Confident	42	43			
		Very Sure	42	40	p=.576	0.93	p=.574
	spring of G1	Very Uncertain	4	3			
		Reluctant	11	13			
		Confident	38	41			
		Very Sure	46	43	p=.336	0.92	p=.520
4-yr old	fall of pre-K (baseline)	Very Uncertain	5	5			
		Reluctant	9	10			
		Confident	43	42			
		Very Sure	43	42	p=.778		
	spring of pre-K	Very Uncertain	8	5			
		Reluctant	7	9			
		Confident	39	41			
		Very Sure	46	45	p=.100	1.10	p=.509
	spring of K	Very Uncertain	1	3			
		Reluctant	10	10			
		Confident	38	36			
		Very Sure	52	51	p=.162	0.88	p=.392
	spring of G1	Very Uncertain	2	3			
		Reluctant	12	12			
		Confident	36	32			
		Very Sure	50	52	p=.299	0.93	p=.591

[a] Results from ordered logistic regression.

[b] For 3-yr old cohort, treatment $1,143 < n < 1,278$, control $697 < n < 765$; for 4-yr old cohort, treatment $922 < n < 1,019$, control $565 < n < 597$.

The second measure, for which parents reported whether the child's teacher had said the child lacks confidence in learning new things or taking part in new activities, was only collected during the fall of the children's kindergarten and first grade years.

As reported in Table 16, chi-square tests for association and regression analyses controlling for covariates, did not produce statistically significant differences between the treatment and control groups at any measurement period, for either cohort.

Table 16: Parent Reports Teacher Indicated Child Lacks Confidence, by Treatment

Cohort	Measurement Period	Parent Reports Teacher Has Told Them Child Lacks Confidence	Control %	Treatment %	Chi Square	Control for Covariates [a]	
						Odds Ratio	Sig.
3-yr old	spring of K	No	81	81			
		Yes	19	19	p=.940	1.05	p=.784
	spring of G1	No	82	80			
		Yes	18	20	p=.378	1.26	p=.150
4-yr old	spring of K	No	82	80			
		Yes	18	20	p=.440	1.22	p=.269
	spring of G1	No	82	82			
		Yes	18	18	p=.840	1.02	p=.917

[a] Results from logistic regression.

[b] For 3-yr old cohort, treatment 1,214 < n < 1,225, control 737 < n < 750; for 4-yr old cohort, treatment 941 < n < 966, control 576 < n < 600.

Discussion

The results of the analyses discussed in this chapter do not yield evidence that children granted access to Head Start, or those actually attending Head Start, fare better than their counterparts with respect to attention, persistence, or confidence, either at the end of the Head Start year or any subsequent spring through first grade. The only

statistically significant difference between these groups, for parent-reported child attention at the end of the Head Start year among the three-year old cohort, existed at the baseline measurement, despite random assignment. Cross-sectional time-series analysis, as presented in Chapter 4, of the variables presented in this chapter yields the same overall results. The results from this sensitivity analysis are not presented in this paper.

The measures used in the analysis presented in this chapter are admittedly imperfect proxies for their underlying constructs, which limits the confidence with which one can infer that Head Start had no impact on these learning skills. Nevertheless, no program impacts on child attention were found for any of the four separate measures, including one derived from a validated instrument, nor were program impacts found for either of the two separate child confidence measures.

The point of comparison for the treatment group includes a sizable proportion of children that attended a center-based preschool. It is possible that the attention, persistence, and confidence outcomes for children participating in Head Start would differ from those of comparable children that did not attend preschool. Chapter 8 presents the results of analyses examining whether there is evidence of such differences.

CHAPTER SIX: PRESCHOOL QUALITY

It is possible that children given access to higher quality Head Start settings demonstrate enduring gains relative to peers not given such access. If this is the case, then successful efforts to improve lower quality classrooms could contribute to a national program that demonstrates enduring gains on average. This chapter presents the findings of multiple ordinary least squares and ordered logistic regression analyses to examine whether the quality observed in Head Start classrooms is associated with enduring positive outcomes, controlling for the covariates identified in Table 4.

Unlike the analyses presented in the preceding two chapters, the analyses presented in this chapter only include children that attended Head Start. For the two continuous outcome measures analyzed in this chapter for which baseline data is available (preacademic skills and problem behavior), for each post-baseline measurement period the following regression model is used to estimate whether the quality observed in Head Start classrooms is associated with changes in the respective outcome:

$$OUTCOME_i = \beta_1 + \beta_2 PQ_i + \beta_3 COV_i + \varepsilon_i, \quad (4)$$

where $OUTCOME_i$ is a measure of the change in preacademic skill or problem behavior from baseline for child *i*, PQ_i is a measure of preschool quality (using the Early Childhood Education Rating System overall measure), and COV_i is a set of covariates measuring child, family, and assessment timing characteristics presented in Table 4.

The regression coefficient β_2 estimates the effect of preschool quality (as measured by the Early Childhood Education Rating System overall measure).

For the one continuous outcome measure analyzed in this chapter for which baseline data is not available (child attention), for each post-baseline measurement period the following regression model is used to estimate whether the quality observed in Head Start classrooms explains child attention:

$$ATTENT_i = \beta_1 + \beta_2 PQ_i + \beta_3 COV_i + \varepsilon_i, \quad (5)$$

where $ATTENT_i$ is a measure of child attention for child *i*, PQ_i is a measure of preschool quality (using the Early Childhood Education Rating System overall measure), and COV_i is a set of covariates measuring child, family, and assessment timing characteristics presented in Table 4. The regression coefficient β_2 estimates the effect of preschool quality (as measured by the Early Childhood Education Rating System overall measure). It would be preferable to control for a baseline measurement of this variable, but in the absence of data for such a measure it is assumed that random assignment produced treatment and control groups with similar baseline characteristics.

For the two ordinal outcome measures analyzed in this chapter (child persistence and child confidence), for each post-baseline measurement period the following regression model is used to estimate whether the quality observed in Head Start classrooms is associated with different outcomes:

$$SKILL_i = \beta_1 + \beta_2 PQ_i + \beta_3 SKILL_{ib} + \beta_4 COV_i + \varepsilon_i, \quad (6)$$

where $SKILL_i$ is a measure of child attention for child *i*, PQ_i is a measure of preschool quality (using the Early Childhood Education Rating System overall measure),

$SKILL_{ib}$ is a measure of persistence or confidence for child *i* at baseline, and COV_i is a set of covariates measuring child, family, and assessment timing characteristics presented in Table 4. The regression coefficient β_2 estimates the effect of preschool quality (as measured by the Early Childhood Education Rating System overall measure).

As indicated in Equation 4, for the continuous dependent variables for which baseline data is available (preacademic skills and problem behavior), the change from the baseline measurement to the subsequent outcome is used as the dependent variable. As indicated in Equation 5, for the lone continuous dependent variable for which baseline data is not available, the outcome itself for each post-baseline measurement is used. As indicated in Equation 6, for the two ordinal dependent variables (persistence and confidence), the rating for the measurement period of interest is used, controlling for the baseline rating through the inclusion of a dummy variable representing each of the variable's levels.

Empirical studies have found process features of quality to be better predictors of child outcomes than structural measures of quality (Howes et al., 2008; Mashburn et al., 2008; National Institute of Child Health and Human Development Early Child Care Research Network, 2002). Accordingly, I considered using the interaction subscale from the revised version of the Early Childhood Environment Rating Scale as administered by independent observers as part of the Impact Study, as the preschool quality measure for the analyses presented in this chapter. However, preliminary analysis suggested that neither this subscale nor other subscales were stronger predictors of child outcomes than the comprehensive scale. Accordingly, the analyses presented in this chapter use the

comprehensive scale, which has been found to be predictive of oral and written language skills (Mashburn et al., 2008) and social/emotional development (Montes, Hightower, Brugger, & Moustafa, 2005).

Results

Preacademic Skills

As described in Chapter 3, the preacademic outcome measure used in the analysis presented in this section is the W score for the pre-academic skills cluster of the Woodcock-Johnson III Tests of Achievement. Table 17 presents the simple correlation between preschool quality and child preacademic skill gains as well as the standardized coefficient and p-value for preschool quality as a predictor of preacademic skill gains, controlling for covariates.

Table 17: Preschool Quality as Predictor of Improvement in Child Preacademic Skills

Cohort	Measurement Period	Correlation	Control for Covariates [a]	
			Beta	Sig.
3-yr old	spring of pre-K	-0.01	0.06	p=.141
	spring of next pre-K year	-0.03	0.02	p=.605
	spring of K	-0.03	<0.01	p=.969
	spring of G1	-0.03	0.03	p=.948
4-yr old	spring of pre-K	-0.13	**-0.10**	**p=.031**
	spring of K	-0.04	-0.04	p=.351
	spring of G1	-0.05	-0.07	p=.186

[a] Results from ordinary least squares regression.

[b] For 3-yr old cohort, 948 < n < 1,051; for 4-yr old cohort, 725 < n < 804.

The measure of preschool quality used in this analysis is only a statistically significant predictor of child preacademic skill gains for the cohort that entered the study as four year olds, and only for the measurement at the end of the preschool year, at which point one standard deviation of higher quality is actually associated with 0.1 standard deviations less improvement in preacademic skills. This relationship was only found for one cohort at one measurement period and previous research has not found higher preschool quality as measured by this instrument to be associated with less cognitive development.

Behavior

As described in Chapter 3, the measure of child behavior used in this analysis is derived from parent ratings of their children on 14 problem behaviors that are indicative

of aggressiveness, hyperactivity and withdrawing. Table 18 presents the simple correlation between preschool quality and child behavior improvement as well as the standardized coefficient and p-value for preschool quality as a predictor of child behavior improvement, controlling for covariates. The measure of preschool quality used in this analysis is not a statistically significant predictor of behavior improvement for either cohort for any measurement period through first grade.

Table 18: Preschool Quality as Predictor of Improvement in Child Behavior [b]

Cohort	Measurement Period	Correlation	Control for Covariates [a]	
			Beta	Sig.
3-yr old	spring of pre-K	0.01	0.03	p=.698
	spring of next pre-K year	0.02	0.06	p=.410
	spring of K	0.03	0.02	p=.456
	spring of G1	-0.05	-0.05	p=.177
4-yr old	spring of pre-K	0.02	0.04	p=.503
	spring of K	-0.01	0.02	p=.511
	spring of G1	< 0.01	0.03	p=.574

[a] Results from ordinary least squares regression.

[b] For 3-yr old cohort, 730 < n < 790; for 4-yr old cohort, 976 < n < 1,051.

Attention

The results presented in this section reflect analysis using the attention measure that reflects the independent assessor's characterization of the children. Table 19 presents the simple correlation between preschool quality and child attention as well as the standardized coefficient and p-value for preschool quality as a predictor of child attention, controlling for covariates.

Table 19: Preschool Quality as Predictor of Child Attention [b]

Cohort	Measurement Period	Correlation	Control for Covariates [a]	
			Beta	Sig.
3-yr old	spring of pre-K	0.03	-0.01	p=.769
	spring of next pre-K year	-0.01	0.04	p=.365
	spring of K	-0.07	-0.06	p=.159
	spring of G1	-0.02	-0.06	p=.225
4-yr old	spring of pre-K	-0.09	-0.06	p=.167
	spring of K	0.03	0.04	p=.450
	spring of G1	0.02	0.04	p=.491

[a] Results from ordinary least squares regression.

[b] For 3-yr old cohort, $528 < n < 770$; for 4-yr old cohort, $804 < n < 1005$.

The measure of preschool quality used in this analysis is not a statistically significant predictor of child attention for either cohort for any measurement period through first grade. Analyses using the three other measures of attention presented in Chapter 5 yielded comparable results.

Persistence

Table 20 presents the simple correlation between preschool quality and child persistence as well as the standardized coefficient and p-value for preschool quality as a predictor of child persistence, controlling for persistence at baseline as well as the covariates introduced above. The measure of preschool quality used in this analysis is

not a statistically significant predictor of child persistence for either cohort for any measurement period through first grade.

Table 20: Preschool Quality as Predictor of Child Persistence [b]

Cohort	Measurement Period	Correlation	Control for Covariates [a]	
			Odds Ratio	Sig.
3-yr old	spring of pre-K	-0.05	0.86	p=.104
	spring of next pre-K year	-0.02	0.92	p=.434
	spring of K	< 0.01	1.09	p=.405
	spring of G1	-0.01	0.93	p=.449
4-yr old	spring of pre-K	-0.10	0.74	p=.064
	spring of K	-0.07	1.17	p=.337
	spring of G1	-0.06	0.79	p=.126

[a] Results from ordered logistic regression.

[b] For 3-yr old cohort, 691 < n < 771; for 4-yr old cohort, 906 < n < 1,005.

Confidence

Table 21 presents the simple correlation between preschool quality and child confidence as well as the standardized coefficient and p-value for preschool quality as a predictor of child confidence, controlling for child confidence at baseline as well as the covariates introduced above. The measure of preschool quality used in this analysis is only a statistically significant predictor of child confidence for the cohort that entered the study as three-year olds, and only for the measurement at the end of the year prior to kindergarten, at which point children given access to Head Start exhibit less confidence compared to children not given such access through the study.

Table 21: Preschool Quality as Predictor of Child Confidence [b]

Cohort	Measurement Period	Correlation	Control for Covariates [a]	
			Odds Ratio	Sig.
3-yr old	spring of pre-K	< 0.01	0.95	p=.613
	spring of next pre-K year	-0.05	0.78	p=.010
	spring of K	-0.05	0.91	p=.355
	spring of G1	< 0.01	0.91	p=.295
4-yr old	spring of pre-K	-0.06	0.87	p=.245
	spring of K	-0.05	0.93	p=.561
	spring of G1	-0.05	0.88	p=.332

[a] Results from ordered logistic regression.

[b] For 3-yr old cohort, 899 < n < 1,005; for 4-yr old cohort, 680 < n < 765.

Discussion

The results of the analyses presented in this chapter do not yield evidence that higher quality preschool experiences, as measured by the Early Childhood Environment Rating Scale-Revised, are associated with better child preacademic skills, behavior, attention, persistence, or confidence. Based on these analyses, it is not clear that efforts to improve lower quality Head Start classrooms would create a national program that demonstrates observable enduring gains beyond those experienced by comparable children not participating in the program.

While the revised version of the Early Childhood Environment Rating Scale has been found to be predictive of preacademic skills and social/emotional development (Mashburn et al., 2008; Montes et al., 2005), it is possible that in this study this instrument did not capture existing aspects of quality that would have been correlated with improved child outcomes. It is possible that other instruments for measuring preschool quality that were not administered as part of this study would have captured such correlations. For instance, the Classroom Assessment Scoring System (La Paro, Pianta, & Stuhlman, 2004; Pianta, La Paro, & Hamre, 2008), which has recently gained prominence, has been found to be a stronger predictor of preacademic skills than the Early Childhood Environment Rating Scale (Mashburn et al., 2008). The Classroom Assessment Scoring System has been found to explain child outcomes during the preschool year equally well for children in Head Start and non-Head Start classrooms (Burchinal, Vandergrift, Pianta, & Mashburn, 2010).

CHAPTER SEVEN: ELEMENTARY SCHOOL QUALITY

It is often suggested that improving the quality of early elementary education would enable initial gains experienced by Head Start participants to endure for a longer period of time. While this theory has intuitive validity, it seems to be contradicted by the findings of a study that found preschool achievement effects to diminish more quickly in smaller elementary classrooms and elementary classrooms with higher levels of academic instruction (Magnuson et al., 2007). The final report of the Impact Study did not identify observed differences in the characteristics of schools attended in kindergarten and first grade between those children that did and did not participate in Head Start, which suggests that school experiences would only play a role to the extent that they differentially impact children given access to Head Start and those not given such access.

In this chapter I use multiple ordinary least squares regression and ordered least squares regression to examine whether various proxies for elementary school quality explain the persistence of initial effects of Head Start participation, controlling for risk factors. These proxies include the percent of children in the school meeting state proficiency standards for math and reading, the percent of students in the school that are eligible for free or reduced price lunch, the size of the child's class, and the number of years the teacher has taught. Unlike the analyses presented in Chapter Six which solely

examined children that attended Head Start, the analyses presented in this chapter include all children in the study.

For the two continuous outcome measures analyzed in this chapter for which baseline data is available (preacademic skills and problem behavior), for each post-baseline measurement period the following regression model is used to estimate the interaction effect between access to Head Start and elementary quality:

$$OUTCOME_i = \beta_1 + \beta_2 HS_i + \beta_3 EQ_i + \beta_4 (HS)(EQ)_i + \beta_5 COV_i + \varepsilon_i , \quad (7)$$

where $OUTCOME_i$ is a measure of the change in preacademic skill or problem behavior from baseline for child *i*, HS_i is a dummy variable representing assignment to the treatment group given access to Head Start, EQ_i is a given proxy measure of elementary quality, $(HS)(EQ)_i$ is a term interacting access to Head Start with elementary quality, and COV_i is a set of covariates measuring child, family, and assessment timing characteristics presented in Table 4. The regression coefficient β_4 estimates the interaction effect between access to Head Start and elementary quality.

For the one continuous outcome measure analyzed in this chapter for which baseline data is not available (child attention), for each post-baseline measurement period the following regression model is used to estimate the interaction effect between access to Head Start and elementary quality:

$$ATTENT_i = \beta_1 + \beta_2 HS_i + \beta_3 EQ_i + \beta_4 (HS)(EQ)_i + \beta_5 COV_i + \varepsilon_i , \quad (8)$$

where $ATTENT_i$ is an absolute measure of child attention for child *i*, HS_i is a dummy variable representing assignment to the treatment group given access to Head Start, EQ_i is a given proxy measure of elementary quality, $(HS)(EQ)_i$ is a term interacting

access to Head Start with elementary quality, and COV_i is a set of covariates measuring child, family, and assessment timing characteristics presented in Table 4. The regression coefficient β_4 estimates the interaction effect between access to Head Start and elementary quality. It would be preferable to control for a baseline measurement of this variable, but in the absence of data for such a measure it is assumed that random assignment produced treatment and control groups with similar baseline characteristics.

For the two ordinal outcome measures analyzed in this chapter (child persistence and child confidence), the following regression model is used to estimate the interaction effect between access to Head Start and elementary quality:

$$SKILL_i = \beta_1 + \beta_2 HS_i + \beta_3 EQ_i + \beta_4 (HS)(EQ)_i + \beta_5 SKILL_{ib} + \beta_6 COV_i + \varepsilon_i , \quad (9)$$

where $SKILL_i$ is a measure of child attention for child *i*, HS_i is a dummy variable representing assignment to the treatment group given access to Head Start, EQ_i is a given proxy measure of elementary quality, $(HS)(EQ)_i$ is a term interacting access to Head Start with elementary quality, $SKILL_{ib}$ is a measure of persistence or confidence for child *i* at baseline, and COV_i is a set of covariates measuring child, family, and assessment timing characteristics presented in Table 4. The regression coefficient β_4 estimates the interaction effect between access to Head Start and elementary quality.

Consistent with the analyses presented in Chapter 6 and as indicated in Equations 7-9 in this chapter, for the continuous dependent variables the change from the baseline measurement to the subsequent outcome is used, for the lone continuous dependent variable without a baseline measure the absolute score for each post-baseline

measurement period is used, and for the ordinal measures the rating for the measurement period of interest is used, controlling for the baseline rating.

The regression analyses presented in this section were conducted separate from the analyses presented in the preceding chapter in the interest of maximizing the number of observations available for inclusion, as simultaneously including the measure of preschool quality and the proxies for elementary school quality markedly reduces the number of observations retained through the process of listwise deletion.

Results

Preacademic Skills

Table 22 presents the p-value for the regression coefficient for an interaction term reflecting the product of each proxy measure of elementary school quality and the treatment group, controlling for covariates. With one exception, these school quality proxies are not differentially associated with child preacademic gains between the two treatment groups. The exception is that for the four-year-old cohort, attending a school with a greater percentage of children eligible for free or reduced price lunch is associated with greater improvement in preacademic skills among children in the control group relative to children in the treatment group.

Table 22: Differential Impact of School Quality on Preacademic Gains by Treatment? [b]

School Quality Proxy	Cohort	Measurement Period	Interaction with Treatment [a]		
			Beta	Sig.	Condition with Greater Benefit from Quality
Child's Teacher's Years of Experience	3-yr old	spring of K	0.07	p=.264	-
		spring of G1	-0.09	p=.110	-
	4-yr old	spring of K	-0.01	p=.892	-
		spring of G1	0.05	p=.502	-
Math Proficiency in Child's School	3-yr old	spring of K	-0.14	p=.136	-
		spring of G1	-0.17	p=.139	-
	4-yr old	spring of K	-0.20	p=.076	-
		spring of G1	-0.22	p=.066	-
Reading Proficiency in Child's School	3-yr old	spring of K	-0.13	p=.120	-
		spring of G1	-0.15	p=.124	-
	4-yr old	spring of K	-0.18	p=.051	-
		spring of G1	-0.10	p=.263	-
% Child's School Eligible Free/Reduced Lunch	3-yr old	spring of K	0.02	p=.789	-
		spring of G1	0.01	p=.928	-
	4-yr old	**spring of K**	**0.39**	**p=.001**	**Control**
		spring of G1	0.15	p=.142	-
Number of Students in Child's Class	3-yr old	spring of K	0.21	p=.127	-
		spring of G1	0.13	p=.485	-
	4-yr old	spring of K	0.16	p=.346	-
		spring of G1	-0.04	p=.870	-

[a] Results from ordinary least squares regression.

[b] For 3-yr old cohort, 1,080 < n < 1,574; for 4-yr old cohort, 913 < n < 1,213.

Behavior

Table 23 presents the p-value for the regression coefficient for an interaction term reflecting the product of each proxy for elementary school quality and treatment group,

controlling for the covariates introduced above. These proxies are not differentially associated with improvement in child behavior between the two treatment groups.

Table 23: Differential Impact of School Quality on Behavior Improvement? [b]

School Quality Proxy	Cohort	Measurement Period	Interaction with Treatment [a]		
			Beta	Sig.	Condition with Greater Benefit from Quality
Child's Teacher's Years of Experience	3-yr old	spring of K	-0.04	p=.536	-
		spring of G1	0.06	p=.286	-
	4-yr old	spring of K	-0.08	p=.330	-
		spring of G1	<.01	p=.998	-
Math Proficiency in Child's School	3-yr old	spring of K	-0.16	p=.248	-
		spring of G1	0.02	p=.887	-
	4-yr old	spring of K	-0.10	p=.461	-
		spring of G1	0.08	p=.557	-
Reading Proficiency in Child's School	3-yr old	spring of K	-0.14	p=.196	-
		spring of G1	-0.08	p=.475	-
	4-yr old	spring of K	-0.03	p=.741	-
		spring of G1	0.06	p=.530	-
% Child's School Eligible Free/Reduced Lunch	3-yr old	spring of K	-0.17	p=.140	-
		spring of G1	-0.01	p=.902	-
	4-yr old	spring of K	0.14	p=.302	-
		spring of G1	0.10	p=.352	-
Number of Students in Child's Class	3-yr old	spring of K	-0.10	p=.452	-
		spring of G1	0.02	p=.891	-
	4-yr old	spring of K	-0.06	p=.770	-
		spring of G1	0.10	p=.592	-

[a] Results from ordinary least squares regression.

[b] For 3-yr old cohort, 1,371 < n < 1,597; for 4-yr old cohort, 890 < n < 1,219.

Attention

Table 24 presents the p-value for the regression coefficient for an interaction term reflecting the product of each proxy measure of elementary school quality and the treatment group, controlling for the covariates introduced above. These proxies are not differentially associated with improvement in child attention between the two treatment groups. Although not presented here, the same results were found for the other three attention measures examined in Chapter 5.

Table 24: Differential Impact of School Quality on Child Attention by Treatment? [b]

School Quality Proxy	Cohort	Measurement Period	Interaction with Treatment [a]		
			Beta	Sig.	Condition with Greater Benefit from Quality
Child's Teacher's Years of Experience	3-yr old	spring of K	-0.05	p=.475	-
		spring of G1	-0.09	p=.179	-
	4-yr old	spring of K	0.10	p=.180	-
		spring of G1	0.07	p=.334	-
Math Proficiency in Child's School	3-yr old	spring of K	0.02	p=.860	-
		spring of G1	0.05	p=.633	-
	4-yr old	spring of K	0.10	p=.391	-
		spring of G1	-0.05	p=.712	-
Reading Proficiency in Child's School	3-yr old	spring of K	< 0.01	p=.934	-
		spring of G1	-0.01	p=.911	-
	4-yr old	spring of K	0.10	p=.256	-
		spring of G1	-0.09	p=.336	-
% Child's School Eligible Free/Reduced Lunch	3-yr old	spring of K	-0.12	p=.255	-
		spring of G1	-0.05	p=.633	-
	4-yr old	spring of K	-0.09	p=.469	-
		spring of G1	-0.01	p=.923	-
Number of Students in Child's Class	3-yr old	spring of K	0.20	p=.133	-
		spring of G1	-0.14	p=.410	-
	4-yr old	spring of K	-0.12	p=.528	-
		spring of G1	-0.32	p=.114	-

[a] Results from ordinary least squares regression.

[b] For 3-yr old cohort, 1,295 < n < 1,546; for 4-yr old cohort, 867 < n < 1,178.

Persistence

Table 25 presents the p-value for the regression coefficient for an interaction term reflecting the product of each proxy measure of elementary school quality and the treatment groups, controlling for the covariates introduced above.

Table 25: Differential Impact of School Quality on Child Persistence by Treatment? [b]

School Quality Proxy	Cohort	Measurement Period	Interaction with Treatment [a]		
			Odds Ratio	Sig.	Condition with Greater Benefit from Quality
Child's Teacher's Years of Experience	3-yr old	spring of K	1.01	p=.475	-
		spring of G1	1.00	p=.850	-
	4-yr old	**spring of K**	**0.95**	**p=.010**	**Control**
		spring of G1	1.00	p=.841	-
Math Proficiency in Child's School	3-yr old	spring of K	0.99	p=.549	-
		spring of G1	1.00	p=.894	-
	4-yr old	spring of K	1.01	p=.571	-
		spring of G1	1.01	p=.202	-
Reading Proficiency in Child's School	3-yr old	spring of K	1.00	p=.917	-
		spring of G1	1.00	p=.462	-
	4-yr old	spring of K	1.00	p=.872	-
		spring of G1	1.00	p=.583	-
% Child's School Eligible Free/Reduced Lunch	3-yr old	spring of K	0.70	p=.610	-
		spring of G1	2.08	p=.290	-
	4-yr old	spring of K	2.68	p=.303	-
		spring of G1	0.39	p=.236	-
Number of Students in Child's Class	3-yr old	spring of K	1.03	p=.429	-
		spring of G1	1.00	p=.959	-
	4-yr old	spring of K	1.04	p=.351	-
		spring of G1	1.04	p=.453	-

[a] Results from ordered logistic regression.

[b] For 3-yr old cohort, 1,202 < n < 1,400; for 4-yr old cohort, 797 < n < 1,077.

With one exception, these school quality proxies are not differentially associated with improvement in child persistence between the two treatment groups. The one exception is that, among the four-year-old cohort, children in the control group benefit from having a teacher with greater experience more than children in the treatment group.

Confidence

Table 26 presents the p-value for the regression coefficient for an interaction term reflecting the product of various proxy measures of elementary school quality and the treatment groups, controlling for the covariates introduced above.

With one exception, these school quality proxies are not differentially associated with improvement in child confidence between the two treatment groups, for either cohort, at either measurement period. The one exception is that, among the four-year-old cohort, children in the control group benefit from having a teacher with greater experience more than children in the treatment group.

Table 26: Differential Impact of School Quality on Child Confidence by Treatment?

School Quality Proxy	Cohort	Measurement Period	Interaction with Treatment [a]		
			Odds Ratio	Sig.	Condition with Greater Benefit from Quality
Child's Teacher's Years of Experience	3-yr old	spring of K	1.01	p=.438	-
		spring of G1	0.99	p=.314	-
	4-yr old	**spring of K**	**0.96**	**p=.046**	**Control**
		spring of G1	1.00	p=.588	-
Math Proficiency in Child's School	3-yr old	spring of K	1.00	p=.518	-
		spring of G1	1.00	p=.916	-
	4-yr old	spring of K	1.01	p=.272	-
		spring of G1	1.00	p=.241	-
Reading Proficiency in Child's School	3-yr old	spring of K	1.00	p=.256	-
		spring of G1	1.00	p=.773	-
	4-yr old	spring of K	1.00	p=.339	-
		spring of G1	1.00	p=.402	-
% Child's School Eligible Free/Reduced Lunch	3-yr old	spring of K	1.03	p=.960	-
		spring of G1	1.45	p=.522	-
	4-yr old	spring of K	1.55	p=.581	-
		spring of G1	0.77	p=.696	-
Number of Students in Child's Class	3-yr old	spring of K	1.05	p=.127	-
		spring of G1	1.00	p=.889	-
	4-yr old	spring of K	1.05	p=.200	-
		spring of G1	1.00	p=.993	-

[a] Results from ordered logistic regression.

[b] For 3-yr old cohort, 1,197 < n < 1,391; for 4-yr old cohort, 786 < n < 1,062.

Discussion

The results of the analyses presented in this chapter do not yield evidence that higher quality elementary school experiences disproportionately benefit children given access to Head Start in comparison to those not granted such access. For both cohorts,

child preacademic skills, behavior, persistence, and confidence are either not differentially influenced by proxies for elementary quality, or children not granted access to Head Start during preschool benefit more from higher quality than those granted access to Head Start.

Based on this analysis, using admittedly imperfect proxies for elementary school quality, it not evident that improving early elementary learning environments would extend the period over which positive impacts for children given access to Head Start are sustained.

CHAPTER EIGHT: HEAD START VERSUS NO PRESCHOOL

Given it would be unethical to prohibit children from participating in early childhood education, evaluations of Head Start do not compare child outcomes to a counterfactual of no preschool participation. In the case of the Head Start Impact Study, 40-50 percent of the control group attended some form of center-based care, including 15-20 percent that attended Head Start. It would also be unethical to mandate that children in a study's treatment group participate in the program. Roughly 15-20 percent of those in the treatment group of the Head Start Impact Study did not participate in the program. The official report of the Impact Study reported effect size estimates that were adjusted using instrumental variable analysis (using the Head Start participation variable) to account for infidelity to the assigned treatment condition. The point of comparison for these estimates still included a sizable proportion of children that attended care settings of comparable quality to Head Start, many of which were publicly funded. This is the most relevant comparison for determining the impact of federal funding for Head Start if one assumes that state and local policymakers make annual funding decisions based on a judgment of the overall number of children in their jurisdiction that should benefit from publicly funded preschool experiences. However, if one assumes that these policymakers make funding decisions based on a judgment of how much preschool they can afford within the context of broader fiscal constraints, which is a more realistic assumption, then

the impact of Head Start participation relative to no preschool experience becomes more salient.

This section presents the results of multiple ordinary least squares regression and ordered logistic regression analyses that examine the effect of Head Start participation relative to no preschool at all. In order to estimate the impact of attending Head Start versus not attending any preschool, one must establish comparable groups for which this is the only difference. While the Head Start Impact Study was not designed to produce such groups, a statistical technique known as propensity score analysis can be used to generate such groups (Guo & Fraser, 2009). In general terms, this technique uses observable characteristics to identify each study participant's propensity for having been exposed to a given experience. These propensities are then used to identify participants with comparable propensities for being exposed to that experience, some of whom actually were exposed to it and some of whom were not. While this technique is traditionally used with observational data, in this analysis it is used to identify a sample of children from the treatment group that attended Head Start that is comparable to a sample of children from the control group that did not attend a preschool program. Conceptually, these groups represent children in the treatment group that attended Head Start but likely would not have attended a preschool program if they had not been assigned to the treatment group, and children in the control group that did not attend a preschool program but are likely to have attended Head Start if they had been assigned to the treatment group. Limiting the analysis to such children reduces the likelihood of selection bias, in which differences between these two groups would be attributable to a

characteristic other than participation in Head Start. It does, however, reduce the sample size and limit the generalizability of any findings to children in these categories.

The sample of children included in the analysis presented in this chapter was identified through Mahalanobis matching using a publicly available Stata add-on (Leuven & Sianesi, 2003). Mahalanobis matching, as opposed to other matching methods, is used in this study as it is the conventional matching method (Guo & Fraser, 2009; Rubin, 1980) and works well in scenarios without the need to use a large number of matching covariates (Guo & Fraser, 2009). The covariates presented in Table 27 were used to identify children with similar propensities for not attending preschool. Ordinary least squares regression and ordered logistic regression was then conducted to examine differences in outcomes for children that attended Head Start with those that did not experience preschool, controlling for the covariates presented in Table 4.

For the two continuous outcome measures analyzed in this chapter for which baseline data is available (preacademic skills and problem behavior), for each post-baseline measurement period the following regression model is used:

$$OUTCOME_i = \beta_1 + \beta_2 TC_i + \beta_3 COV_i + \varepsilon_i, \quad (10)$$

where $OUTCOME_i$ is a measure of the change in preacademic skill or problem behavior from baseline for child *i*, TC_i is a dummy variable representing the treatment contrast between a child that participated in Head Start through the study and a child that was not given access to Head Start through the study and did not attend preschool, and COV_i is a set of covariates measuring child, family, and assessment timing characteristics

presented in Table 4. The regression coefficient β_2 estimates the effect of attending Head Start relative to a counterfactual of receiving no preschool.

For the one continuous outcome measure analyzed in this chapter for which baseline data is not available (child attention), for each post-baseline measurement period the following regression model is used:

$$ATTENT_i = \beta_1 + \beta_2 TC_i + \beta_3 COV_i + \varepsilon_i, \quad (11)$$

where $ATTENT_i$ is an absolute measure of child attention for child *i*, TC_i is a dummy variable representing the treatment contrast between a child that participated in Head Start through the study and a child that was not given access to Head Start through the study and did not attend preschool, and COV_i is a set of covariates measuring child, family, and assessment timing characteristics presented in Table 4. The regression coefficient β_2 estimates the effect of attending Head Start relative to a counterfactual of receiving no preschool. It would be preferable to control for a baseline measurement of the dependent variable, but in the absence of data for such a measure it is assumed that random assignment produced treatment and control groups with similar baseline characteristics.

For the two ordinal outcome measures analyzed in this chapter (child persistence and child confidence), the following regression model is used:

$$SKILL_i = \beta_1 + \beta_2 TC_i + \beta_3 SKILL_{ib} + \beta_4 COV_i + \varepsilon_i, \quad (12)$$

where $SKILL_i$ is a measure of child attention for child *i*, TC_i is a dummy variable representing the treatment contrast between a child that participated in Head Start through the study and a child that was not given access to Head Start through the study and did

not attend preschool, $SKILL_{ib}$ is a measure of persistence or confidence for child *i* at baseline, and COV_i is a set of covariates measuring child, family, and assessment timing characteristics presented in Table 4. The regression coefficient β_2 estimates the effect of attending Head Start relative to a counterfactual of receiving no preschool.

The term "treatment contrast" is used in this chapter to refer to participation in Head Start versus no attendance of a preschool program. This terminology intentionally differs from use of the term "treatment group" and "treatment assignment" in the preceding chapters, which refer to the group to which children were assigned as part of the experimental design of the Head Start Impact Study.

Elementary school quality variables were not included as covariates in the analyses presented in this chapter, in the interest of maximizing the sample size, within the context of these variables not being found to be statistically significant predictors of child outcomes in the analyses presented in the preceding two chapters.

Results

One key criterion for judging the quality of propensity score matching estimation is whether it balances characteristics between the treatment and comparison groups (Heinrich, Maffioli, & Vazquez, 2010). In the ideal scenario, there would be no additional matching covariates that could be added that would improve the estimation. For this analysis, no additional theoretically relevant covariates could be added without substantially reducing the available sample as a result of missing values. It should be acknowledged that the omission of unobserved characteristics could lead to bias in the resulting estimation. For instance, varying levels of motivation among children's parents

that is not captured by the included covariates could yield a biased result. In the ideal situation, there would also be no post-match differences in the baseline characteristics between the treatment and comparison groups on the observed characteristics. Table 27 provides a summary of the baseline characteristics for children within the common support region identified through the propensity score analysis.

Table 27: Baseline Covariate Summary, Among Children in Common Support Region [a]

	3-year-old cohort		4-year-old cohort	
Covariate	No Preschool *n = 326*	Head Start *n = 820*	No Preschool *n = 265*	Head Start *n = 642*
Female	52%	50%	50%	47%
Black	29%*	36%*	18%	17%
Hispanic	40%*	33%*	43%	46%
Spanish Primary Language	24%	21%	29%	34%
Spanish Spoken at Home	28%	25%	30%*	38%*
Age of Primary Caregiver	29 (7)	28 (7)	29 (7)	29 (7)
Live with Both Parents	51%	47%	53%	51%
Mother is Married	48%*	40%*	50%	45%
Mother is Recent Immigrant	16%	15%	20%	24%
Born to Teen Mother	14%	14%	21%*	15%*
Mother's Education:				
Less than High School	39%	34%	48%	43%
High School or GED	34%	35%	29%	31%
Beyond High School	27%	30%	23%	27%
Urban (vs. Rural)	81%	83%	87%	85%
"Special Needs"	8%**	15%**	14%	14%

[a] For chi-square test for differences between treatment and control: * p < .05, ** < .01.

Consistent with the original sample, for both cohorts there are more children in the Head Start group than in the no preschool group. While there are not statistically significant differences between these two groups for most of these covariates, there are for a few variables. For the three-year-old cohort, children that attended Head Start are more likely to be Black, more likely to have been reported to have "special needs", less likely to be Hispanic, and less likely to have an unmarried mother. For the four-year-old cohort, children that attended Head Start are more likely to live in a home where Spanish is spoken and are less likely to have been born to a teen mother. For context, within the original control group the subset of children included in this analysis (those that did not attend preschool) are more likely to have both parents at home, have a married mother, have a more educated mother, have a mother that is a recent immigrant, have Spanish spoken at home, and not have had a physician tell their parent they have "special needs".

A second key criterion for judging the quality of propensity score matching estimation is whether there is satisfactory overlap in the propensity scores of the generated treatment and comparison groups (Heinrich et al., 2010). Figure 2 displays propensity scores by treatment and comparison group. This figure shows substantial overlap among the scores for these two groups.

A third criterion is the propensity score model fit, which can be measured by the pseudo-R^2 statistic (Guo & Fraser, 2009). While this statistic cannot be interpreted independently or across data sets, it can be used to identify the best model fit for a given analysis. For both cohorts, 0.03 is the greatest pseudo-R^2 statistic could be achieved without notably reducing the sample size or comprising theoretical validity.

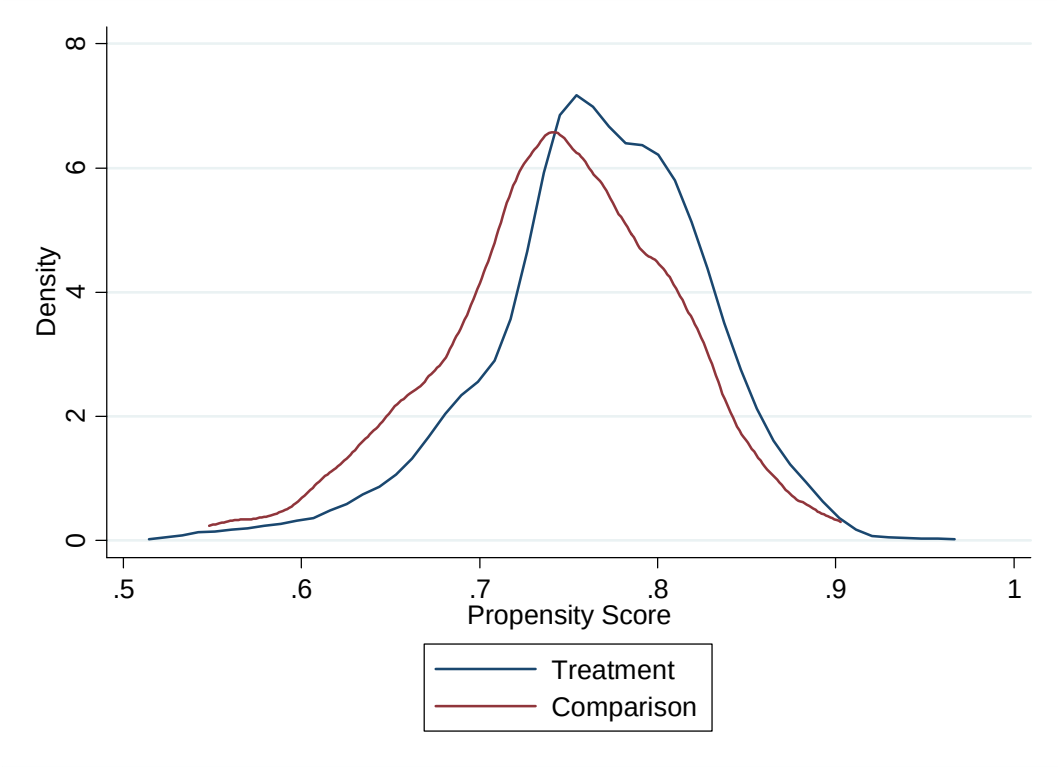

Figure 2: Propensity Scores by Treatment and Comparison Group

Preacademic Skills

Table 28 presents the mean score for preacademic skill gains, for comparable children in the Head Start and no preschool groups. This table also reports the results of t-tests examining differences in the means for this measure, as well as the p-value for the regression coefficient for the treatment contrast, controlling for the covariates described above.

For both the three-year-old and four-year-old cohorts, at the end of the Head Start year children participating in Head Start improved by roughly 0.1 standard deviations

more than comparable children that did not experience preschool, controlling for covariates. For the four-year-old cohort, there were no statistically significant differences on this measure in kindergarten or first grade, controlling for covariates. For the three-year-old cohort, in kindergarten and first grade, children that did not experience preschool actually demonstrated roughly 0.1 standard deviations more improvement than comparable children that attended Head Start, controlling for covariates.

Table 28: Preacademic Skill Gains for Head Start versus No Preschool [b]

Cohort	Measurement Period	Mean		T-test	Control for Covariates [a]	
		No Preschool	Head Start		Beta	Sig.
3-yr old	spring of pre-K	0	7	p=<.001	**0.09**	**p=.008**
	spring of next pre-K year	28	33	p=.002	0.05	p=.161
	spring of K	74	74	p=.936	**-0.08**	**p=.024**
	spring of G1	112	110	p=.266	**-0.11**	**p=.006**
4-yr old	spring of pre-K	2	4	p=.213	**0.10**	**p=.001**
	spring of K	51	45	p=.002	-0.07	p=.148
	spring of G1	90	86	p=.073	-0.03	p=.486

[a] Results from ordinary least squares linear regression.

[b] For 3-yr old cohort, 954 < n < 1,053; for 4-yr old cohort, 758 < n < 832.

Behavior

Table 29 presents the mean behavior improvement, for comparable children in the Head Start and no preschool groups. This table also reports the results of t-tests examining differences in means for this measure, as well as the p-value for the regression

coefficient for the treatment contrast, controlling for the covariates described above. With one exception, no statistically significant differences were found between the treatment and control groups at any measurement period, for either cohort, under either analysis method. The exception is that for the four-year-old cohort, when not controlling for covariates, in the spring of kindergarten children in the treatment group showed less behavior improvement than the control group. No such difference was found in the prior or subsequent year.

Table 29: Behavior Improvement for Head Start versus No Preschool [b]

		Mean			Control for Covariates [a]	
Cohort	Measurement Period	No Preschool	Head Start	T-test	Beta	Sig.
3-yr old	spring of pre-K	0.07	0.18	p=.623	0.04	p=.345
	spring of next pre-K year	0.59	0.80	p=.403	0.07	p=.100
	spring of K	1.21	1.18	p=.911	0.01	p=.764
	spring of G1	1.27	1.34	p=.813	-0.01	p=.796
4-yr old	spring of pre-K	0.41	0.42	p=.994	0.05	p=.263
	spring of K	1.72	1.00	**p=.017**	-0.04	p=.442
	spring of G1	1.42	1.36	p=.854	0.06	p=.271

[a] Results from ordinary least squares linear regression.

[b] For 3-yr old cohort, 969 < n < 1,069; for 4-yr old cohort, 762 < n < 833.

Attention

Table 30 presents the mean score for child attention as assessed by their teacher, for comparable children in the Head Start and no preschool groups. Whereas gain scores

were used for preacademic skills and behavior in the two preceding sections, the analysis in this section uses absolute attention scores because teachers did not assess child attention during the baseline measurement. This table also reports the results of t-tests examining differences in the means for this measure, as well as the p-value for the regression coefficient for the treatment contrast, controlling for the covariates described above. No statistically significant differences were found between the treatment and control groups for either cohort at any measurement period under either analysis method. Although not presented here, no effects were found for the other three attention measures examined in Chapter 5.

Table 30: Child Attention for Head Start versus No Preschool [b]

		Mean			Control for Covariates [a]	
Cohort	Measurement Period	No Preschool	Head Start	T-test	Beta	Sig.
3-yr old	spring of pre-K	1.64	1.87	p=.503	0.02	p=.781
	spring of next pre-K year	1.37	1.64	p=.105	0.04	p=.307
	spring of K	1.80	1.90	p=.574	-0.01	p=.886
	spring of G1	1.76	1.97	p=.227	0.05	p=.258
4-yr old	spring of pre-K	1.81	1.81	p=.997	-0.06	p=.432
	spring of K	1.85	1.76	p=.680	0.01	p=.912
	spring of G1	1.80	1.75	p=.809	-0.05	p=.392

[a] Results from ordinary least squares linear regression.

[b] For 3-yr old cohort, $726 < n < 823$; for 4-yr old cohort, $554 < n < 594$.

Persistence

Table 31 presents the percentage of children rated in each of three persistence categories reported by the independent assessors, by treatment contrast, as well as the results of a chi-square test for association between treatment contrast and this measure. Table 31 also presents the p-value for the regression coefficient for the treatment contrast variable, after controlling for covariates.

For both the three-year-old and four-year-old cohorts, at the end of the Head Start year children participating in Head Start were 61 percent and 75 percent, respectively, more likely to be rated at a higher level of persistence, controlling for covariates. It is worth noting that for the three-year-old cohort the matching process yielded a modest baseline persistence advantage for the treatment group. No statistically significant differences were found for either cohort at any subsequent measurement period, under either analysis method.

Table 31: Child Persistence for Head Start versus No Preschool [b]

Cohort	Measurement Period	Rating	No Preschool	Head Start	Chi Square	Control for Covariates [a]	
						Sig.	Odds Ratio
3-yr old	fall of pre-k (baseline)	Refuses	5	4			
		Attempts	59	52			
		Persists	37	44	**p=.020**		
	spring of pre-K	Refuses	3	3			
		Attempts	55	48			
		Persists	41	49	p=.091	**p=.010**	**1.61**
	spring of next pre-K year	Refuses	2	1			
		Attempts	38	37			
		Persists	60	61	p=.679	p=.658	1.09
	spring of K	Refuses	0	0			
		Attempts	34	29			
		Persists	66	71	p=.200	p=.792	0.95
	spring of G1	Refuses	1	1			
		Attempts	26	26			
		Persists	73	73	p=.996	p=.433	0.85
4-yr old	fall of pre-K (baseline)	Refuses	2	2			
		Attempts	34	34			
		Persists	64	65	p=.964		
	spring of pre-K	Refuses	2	1			
		Attempts	42	33			
		Persists	56	66	**p=.037**	**p=.020**	**1.75**
	spring of K	Refuses	0	1			
		Attempts	25	29			
		Persists	75	70	p=.228	p=.159	0.67
	spring of G1	Refuses	0	0			
		Attempts	23	25			
		Persists	76	75	p=.720	p=.593	1.15

[a] Results from ordered logistic regression.

[b] For 3-yr old cohort, 938 < n < 1,047; for 4-yr old cohort, 734 < n < 817.

Confidence

Table 32 presents the percentage of children rated in each of four confidence categories reported by the independent assessors, by treatment contrast, as well as the results of a chi-square test for association between the treatment contrast and this measure. Table 32 also presents the p-value for the regression coefficient for the treatment contrast variable, controlling for covariates.

For the three-year-old cohort, at the end of the Head Start year children participating in Head Start were 53 percent more likely to be rated at a higher level of confidence, controlling for covariates. It is worth noting that for this cohort the matching process yielded a modest baseline confidence advantage for the treatment group. For both cohorts, no statistically significant differences were found at any subsequent measurement period, controlling for covariates.

Table 32: Child Confidence for Head Start versus No Preschool [b]

Cohort	Measurement Period	Rating	No Preschool	Head Start	Chi Square	Control for Covariates [a]	
						Odds Ratio	Sig.
3-yr old	fall of pre-k (baseline)	Very Uncertain	14	12			
		Reluctant	20	14			
		Confident	38	43			
		Very Sure	28	30	**p=.011**		
	spring of pre-K	Very Uncertain	14	11			
		Reluctant	14	14			
		Confident	41	42			
		Very Sure	31	33	p=.591	**1.53**	**p=.020**
	spring of next pre-K year	Very Uncertain	9	7			
		Reluctant	12	12			
		Confident	43	44			
		Very Sure	35	37	p=.531	1.32	p=.109
	spring of K	Very Uncertain	5	4			
		Reluctant	12	11			
		Confident	42	44			
		Very Sure	41	42	p=.842	0.91	p=.604
	spring of G1	Very Uncertain	5	3			
		Reluctant	11	13			
		Confident	35	41			
		Very Sure	49	43	p=.239	0.83	p=.316
4-yr old	fall of pre-K (baseline)	Very Uncertain	6	5			
		Reluctant	11	11			
		Confident	44	43			
		Very Sure	40	41	p=.889		
	spring of pre-K	Very Uncertain	11	4			
		Reluctant	9	10			
		Confident	39	43			
		Very Sure	41	43	**p=.004**	1.39	p=.124
	spring of K	Very Uncertain	1	3			
		Reluctant	11	13			
		Confident	43	36			
		Very Sure	45	49	p=.212	0.87	p=.532
	spring of G1	Very Uncertain	2	3			
		Reluctant	10	14			
		Confident	32	33			
		Very Sure	56	50	p=.369	0.88	p=.568

[a] Results from ordered logistic regression.

[b] For 3-yr old cohort, 932 < n < 1,044; for 4-yr old cohort, 724 < n < 810.

Discussion

The results presented in this chapter do not yield evidence that children attending Head Start experience better outcomes in kindergarten or first grade than comparable children not experiencing preschool. It is important to note that the subsample used in this analysis is only representative of the subset of study participants who were likely to attend Head Start if given access, but were not otherwise likely to experience preschool.

For the three-year-old cohort, children that had participated in Head Start actually demonstrated roughly 0.1 standard deviations less improvement on preacademic skills in kindergarten and first grade than comparable children that did not experience preschool, controlling for covariates. This difference could be attributable to random error given that it only occurred for the younger cohort and in the spring of their preschool year those participating in Head Start had outperformed those that did not experience preschool.

CHAPTER NINE: CONCLUSIONS & FUTURE RESEARCH

In some media outlets and policy circles, findings from the Head Start Impact Study (Department of Health and Human Services, 2010a) have been cited as evidence that the program is ineffective and its funding should be diverted to uses believed to be more effective (Klein, 2011). Citing these results along with program integrity concerns, the appropriations committee in the House of Representatives proposed a $1 billion reduction in funding for the program for fiscal year 2011.

Some analysts rebut such proposals by citing previous studies that have found positive long term high school completion effects associated with Head Start participation (Garces et al., 2002; Ludwig & Miller, 2007) and broader meta-analyses of early childhood education studies carried out in the United States that have consistently found positive cognitive and social/emotional outcomes for children participating in the intervention relative to a comparison group of children with similar characteristics (Camilli et al., 2010; Gorey, 2001; Karoly et al., 2005; Nelson et al., 2003). In the interest of identifying early elementary mediators for potential long term effects, the analyses presented in Chapter 4 examined whether controlling for within-child variation facilitates the detection of previously unidentified enduring effects associated with being given access to Head Start, and the analyses presented in Chapter 5 examined whether access to Head Start is associated with program effects for learning skills such as

attention, persistence, and confidence. Controlling for within-child variation did reveal a previously unidentified positive effect for preacademic skills among the cohort that entered the study as three-year-olds, in their post-Head Start year prior to starting Kindergarten. However, this method did not identify favorable effects for preacademic skills or problem behavior in kindergarten or first grade for either cohort. With respect to attention, persistence, and confidence, no enduring positive effects were identified. It should be acknowledged that it is possible that other measures of these constructs, other constructs altogether, or alternative analyses of these same constructs could ultimately be found to mediate any long term benefits associated with the program. It should also be acknowledged that benefits could still reemerge later in life (Peisner-Feinberg et al., 2001). While there is not a strong theoretical explanation for such reemergence, evidence suggestive of it has been identified in at least three studies (Broberg et al., 1997; Magnuson et al., 2007; Schweinhart & Weikart, 1998).

In response to criticisms of the effectiveness of the Head Start program, its supporters also note that the program has been improved since the period covered by the most recent evaluation and ongoing improvements will make it even more effective. Within this context, the analyses presented in Chapter 6 examined whether there is evidence that the quality of care and education provided in Head Start classrooms explains children's outcomes through first grade. The results from these analyses do not lend empirical support to the suggestion that improving program quality within the range of quality provided by current Head Start programs would yield enduring effects for children given access to Head Start (relative to comparable peers not given such access).

While the measure of preschool quality used in this study has previously been found to be predictive of preacademic skills and social/emotional development (Mashburn et al., 2008; Montes et al., 2005), it is possible that quality as measured by other instruments would yield different results. It should also be acknowledged that past research has found the association between preschool quality and positive child outcomes to be stronger at higher levels of quality (Zaslow et al., 2010) that may not be met by a number of the Head Start centers included in the present study.

It is also commonly suggested that improving the nation's early elementary education system would enable initial Head Start program effects to endure. Within this context, the analyses presented in Chapter 7 examined whether there is evidence that early elementary school experiences influence whether initial effects of Head Start endure. The analyses presented in this chapter do not find evidence suggesting that improving the quality of elementary schools subsequently attended by Head Start participants can be expected to produce an average program effect that is sustained through the early elementary years. That said, the proxies for elementary quality used in this study are imperfect, and it is possible that quality as measured by other instruments (particularly those assessing actual classroom interactions) would yield different results.

Some analysts also correctly note that a sizable proportion of children that make up the comparison group for evaluations of Head Start attend other publicly funded preschool programs and in some cases Head Start itself. As a result, assessments of the impact of the program are not made against a "no-services" counterfactual. Within this context, the analyses presented in Chapter 8 examined whether there is evidence of

enduring effects associated with being given access to Head Start compared to a counterfactual of no preschool participation. The results from the analyses presented in this chapter do not reveal evidence that comparing Head Start participation to a counterfactual of no preschool attendance yields program effects through the early elementary years.

The absence of evidence that initial positive Head Start effects are sustained in the early elementary years is consistent with the findings reported in the official report of the Head Start Impact Study (Department of Health and Human Services, 2010a) and previous evaluations of the program (McKey et al., 1985). These results suggest that the influence of access to the Head Start early childhood services pales in comparison to the powerful influence of a child's family and community environment and resources on their development (Armor, 2003; Brooks-Gunn & Duncan, 1997). Nevertheless, it is worth considering how consistent findings of initial positive effects not being sustained can be reconciled with broader evidence of stability in cognitive and social/emotional competencies over time. As indicated in the literature review in Chapter 2, one interpretation of stability in cognitive and social/emotional competencies over time is that early development facilitates subsequent development. This would suggest that initial gains should translate into a continued advantage through a child's elementary years. In the cognitive domain, one potential explanation for initial positive Head Start gains not being sustained is that the initial gains (less than 5 points) are rather modest compared to broader development and schooling effects (50-60 points by first grade). While children's cognitive/academic performance in preschool or kindergarten has been found

to account for 25 percent of the variance in these domains a year or two later, this leaves 75 percent to be explained by other factors (La Paro & Pianta, 2000). Children's social development and behavior in preschool or kindergarten has been found to account for 10 percent of the variance in these domains a year or two later, leaving an even greater proportion to be explained by other factors.

An alternative interpretation of stability in cognitive and social/emotional competencies is that children are generally exposed to similar environmental factors, like parenting behaviors, over time (Entwisle et al., 2005). This latter interpretation is more consistent with the findings of the present study and past evaluations of the Head Start program and with other empirical evidence regarding the powerful influence of a child's family and community environment and resources on their development (Armor, 2003; Brooks-Gunn & Duncan, 1997). For cognitive development, the most important risk factors are parents' intelligence, parents' educational attainment, family income, family structure, age of the mother when the child is born, the number of siblings, the child's nutrition and birth weight, and parental instruction and nurturing (Armor, 2003). For social development, the strongest predictors are maternal education, maternal depressive symptoms, and sensitivity of mothering (National Institute of Child Health and Human Development Early Child Care Research Network, 2003). With respect to these key risk factors, the Head Start program endeavors to positively influence child nutrition and parental instruction and nurturing, but the remaining risk factors are largely beyond the reach of the program. Despite the program's efforts to support improved

parenting, the official report of the Head Start Impact Study did not find favorable parenting impacts in kindergarten or first grade.

The positive high school completion effects found in longer-term studies could be the result of early development influencing subsequent development in yet unidentified or imperfectly measured domains. Studies of other early childhood education programs have reported evidence of positive long-term effects despite undetected effects in the early elementary years (Broberg et al., 1997; Magnuson et al., 2007; Schweinhart & Weikart, 1998). Attention skills are an example of a competency that has shown promise to predict subsequent achievement outcomes (Barriga et al., 2002; Duncan et al., 2007; Howse et al., 2003), but has not been emphasized in evaluations of Head Start. The present study did not find evidence of Head Start influencing attention, persistence, or confidence, but this does not preclude future studies from finding competencies or learning behaviors, such as response to novelty and error, reflectivity, and flexibility (Yen, Konold, & McDermott, 2004), that could be influenced by early childhood education and in turn influence subsequent development.

One way in which Head Start may have a positive impact through the early elementary years and beyond, in the absence of Head Start participants outperforming comparable nonparticipants, would be if they contribute to the wellbeing of their early elementary classmates (Ohio University and Westinghouse Learning Corporation, 1969) by virtue of arriving at kindergarten with the improved outcomes that have consistently been found upon completion of the Head Start year. As examples, this could occur through less problem behavior that would otherwise distract classmates, less

individualized attention being required of the teacher, and more positive relationships with their peers. The child-centered nature of the data used for this research project precludes the examination of whether there is evidence to support this potential contribution from Head Start. The community-based design used by Ludwig and Miller (2007) that found long-term benefits was structured in a way that included such indirect benefits.

Future Research

At least two notable studies using data from the Head Start Impact Study are currently underway and can be expected to further inform policy and practice. The first is the analysis of recently collected data on experiences and outcomes through third grade of children participating in the Head Start Impact Study. The results of this analysis are expected to be published in the coming months, with restricted use data likely becoming available sometime thereafter. The second study, which is being carried out by a collaboration of researchers, is using the currently available data through first grade to examine how various features of Head Start centers are associated with variation in program impacts on cognitive and social/emotional child outcomes. This research is seeking to identify what differentiates programs that are effective from those that are less effective.

Ongoing research is also examining the relationship between the quality of child care and education and child outcomes (Zaslow et al., 2010). It would be particularly helpful for future research on this relationship to utilize data on the quality of Head Start classrooms that is now being collected for approximately one-third of Head Start grantees

annually using the Classroom Assessment Scoring System (La Paro et al., 2004; Pianta et al., 2008). The three quality domains captured by this instrument have been found to be associated with child outcomes, but it would be valuable to learn the particular circumstances under which quality improvements as measured by this instrument yield initial and sustained positive outcomes.

Future research could build on the present study by examining potential endogeneity issues, such as the possibility of nonrandom attrition and nonrandom selection of elementary schools that cannot be accounted for through observable characteristics. Future research could also seek to find yet unidentified competencies and characteristics exhibited by children in their early elementary years that may serve as mediators between receipt of quality early childhood education and positive longer-term outcomes such as high school completion. Such competencies could include learning behaviors not examined in the present study, such as response to novelty and error, reflectivity, and flexibility (Yen et al., 2004). Future research could also attempt to utilize research designs that account for potential positive effects to classrooms that could result from Head Start children arriving with greater cognitive and social/emotional development than they would have had in the absence of benefiting from high quality early childhood education.

APPENDIX

Table 33 summarizes the results from multiple ordinary least squares, logistic, and ordered logistic regression models structured as follows:

$$SKILL_i = \beta_1 + \beta_2 A_i + \beta_3 SKILL_{ib} + \beta_4 COV_i + \varepsilon_i , \quad (13)$$

where $SKILL_i$ is a measure of a given learning skill for child *i*, A_i is a dummy variable representing whether a child attended Head Start, $SKILL_{ib}$ is a measure of a given learning skill for child *i* at baseline (where such measurement is available), and COV_i is a set of covariates measuring child, family, and assessment timing characteristics presented in Table 4. The regression coefficient β_2 (for which the results of significance tests are provided in Table 33) estimates the effect of attending Head Start.

Table 33: Participation as Predictor of Learning Skills, Controlling for Covariates [a]

Cohort	Measurement Period	Attention				Persistence	Confidence	
		Teacher Report	Parent Report	Parent Report	Independent Assessment	Independent Assessment	Independent Assessment	Parent Report
3-yr old	spring of pre-K	.580	**.007**	n/a	.366	.328	.061	n/a
	spring of next pre-K year	.742	.648	n/a	.998	.532	.225	n/a
	spring of K	.608	.164	.828	.564	.990	.162	.073
	spring of G1	.480	.871	.141	.591	.238	**.049**	.584
4-yr old	spring of pre-K	.553	.243	n/a	.727	.118	.267	n/a
	spring of K	.520	.340	.092	.315	.239	.275	.775
	spring of G1	.625	.509	.519	.674	.304	.642	.237

[a] Values in this table are the p-values for the regression coefficient for Head Start participation, controlling for the covariates identified in Table 4.

REFERENCES

REFERENCES

Abbott-Shim, M., Lambert, R., & McCarty, F. (2003). A Comparison of School Readiness Outcomes for Children Randomly Assigned to a Head Start Program and the Program's Wait List. *Journal of Education for Students Placed at Risk (JESPAR)*, *8*(2), 191–214.

Allhusen, V., Belsky, J., Booth-LaForce, C. L., Bradley, R., Brownwell, C. A., Burchinal, M., Campbell, S. B., et al. (2004). Does Class Size in First Grade Relate to Children's Academic and Social Performance or Observed Classroom Processes? *Developmental Psychology*, *40*(5), 651–664.

Armor, D. J. (2003). *Maximizing Intelligence*. New Brunswick: Transaction Publishers.

Barnett, W. S., & Masse, L. N. (2007). Comparative benefit-cost analysis of the Abecedarian program and its policy implications. *Economics of Education Review*, *26*(1), 113–125.

Barriga, A. Q., Doran, J. W., Newell, S. B., Morrison, E. M., Barbetti, V., & Dean Robbins, B. (2002). Relationships Between Problem Behaviors and Academic Achievement in Adolescents. *Journal of Emotional and Behavioral Disorders*, *10*(4), 233 –240.

Belfield, C. R., Nores, M., Barnett, S., & Schweinhart, L. (2006). The High/Scope Perry Preschool Program: Cost-Benefit Analysis Using Data from the Age-40 Followup. *The Journal of Human Resources*, *41*(1), 162–190.

Berk, L. E. (2009). *Child Development* (8th ed.). Boston: Allyn & Bacon.

Betts, J. R., & Shkolnik, J. L. (1999). The Behavioral Effects of Variations in Class Size: The Case of Math Teachers. *Educational Evaluation and Policy Analysis*, *21*(2), 193–213.

Birch, S. H., & Ladd, G. W. (1997). The teacher-child relationship and children's early school adjustment. *Journal of School Psychology*, *35*(1), 61–79.

Broberg, A. G., Wessels, H., Lamb, M. E., & Hwang, C. P. (1997). Effects of day care on the development of cognitive abilities in 8-year-olds: A longitudinal study. *Developmental Psychology*, *33*(1), 62–69.

Bronfenbrenner, U. (1979). *The Ecology of Human Development: Experiments by Nature and Design*. Cambridge: Harvard University Press.

Brooks-Gunn, J., & Duncan, G. J. (1997). The Effects of Poverty on Children. *The Future of Children*, *7*(2), 55–71.

Bryant, D. M., Clifford, R. M., & Peisner, E. S. (1991). Best Practices for Beginners: Developmental Appropriateness in Kindergarten. *American Educational Research Journal*, *28*(4), 783–803.

Burchinal, M. R., Vandergrift, N., Pianta, R. C., & Mashburn, A. J. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly*, *25*(2), 166–176.

Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *The Teachers College Record*, *112*(3).

Campbell, F. A., Pungello, E. P., Miller-Johnson, S., Burchinal, M., & Ramey, C. T. (2001). The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. *Developmental Psychology*, *37*(2), 231–242.

Currie, J., & Thomas, D. (1995). Does Head Start Make a Difference? *The American Economic Review*, *85*(3), 341–364.

Currie, J., & Thomas, D. (2000). School Quality and the Longer-Term Effects of Head Start. *The Journal of Human Resources*, *35*(4), 755–774.

Denham, S. A. (1998). *Emotional development in young children*. Guilford Press.

Department of Health and Human Services. (2005). Head Start Impact Study, First Year Findings.

Department of Health and Human Services. (2010a). Head Start Impact Study, Final Report.

Department of Health and Human Services. (2010b). Early Head Start Children in Grade 5: Long-term Follow-Up of the Early Head Start Research and Evaluation Study Sample.

Department of Health and Human Services. (2010c). Head Start Impact Study. Technical Report. Washington DC.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K. A., Huston, A. C., Klebanov, P., Pagani, L. S., et al. (2007). School Readiness and Later Achievement. *Developmental Psychology*, *43*(6), 1428–1446.

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2005). First Grade and Educational Attainment by Age 22: A New Story. *American Journal of Sociology*, *110*(5), 1458–1502.

Finn, J. D., & Achilles, C. M. (1999). Tennessee's Class Size Study: Findings, Implications, Misconceptions. *Educational Evaluation and Policy Analysis*, *21*(2), 97–109.

Garces, E., Thomas, D., & Currie, J. (2002). Longer-Term Effects of Head Start. *American Economic Review*, American Economic Review, *92*(4), 999–1012.

Gertler, P., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2010). *Impact Evaluation in Practice*. World Bank Publications.

Gorey, K. M. (2001). Early Childhood Education: A Meta-Analytic Affirmation of the Short- and Long-Term Benefits of Educational Opportunity. *School Psychology Quarterly*, *16*(1), 9–30.

Guo, S. Y., & Fraser, M. W. (2009). *Propensity Score Analysis: Statistical Methods and Applications* (1st ed.). Sage Publications, Inc.

Hamre, B. K., & Pianta, R. C. (2001). Early Teacher–Child Relationships and the Trajectory of Children's School Outcomes through Eighth Grade. *Child Development*, *72*(2), 625.

Hanushek, E. A. (1999). Some Findings from an Independent Investigation of the Tennessee STAR Experiment and from Other Investigations of Class Size Effects. *Educational Evaluation and Policy Analysis*, *21*(2), 143–163.

Heckman, J. J. (2000). Policies to Foster Human Capital. *Research in Economics*, *54*, 3-56.

Heckman, J. J. (2006). Skill Formation and the Economics of Investing in Disadvantaged Children. *Science*, *312*(5782), 1900–1902.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. (2010). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics*, *94*(1–2), 114–128.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, *24*(3), 411–482.

Heinrich, C., Maffioli, A., & Vazquez, G. (2010). *A Primer for Applying Propensity-Score Matching*. Impact-Evaluation Guidelines: Technical Notes.

Howse, R. B., Lange, G., Farran, D. C., & Boyles, C. D. (2003). Motivation and Self-Regulation as Predictors of Achievement in Economically Disadvantaged Young Children. *The Journal of Experimental Education*, *71*(2), 151–174.

Jaffe, L. (2009). Development, Interpretation, and Application of the W Score and the Relative Proficiency Index. Riverside Publishing. Retrieved from http://www.assess.nelson.com/pdf/asb-11.pdf

Karoly, L. A., Kilburn, R. M., & Cannon, J. S. (2005). *Early Childhood Interventions: Proven Results, Future Promise*. Santa Monica, CA: The RAND Corporation. Retrieved from http://www.rand.org/pubs/monographs/MG341.html

Kirk, R. E. (1994). *Experimental Design: Procedures for Behavioral Sciences* (3rd ed.). Wadsworth Publishing.

Klein, J. (2011, undefined). Time to Ax Public Programs That Don't Yield Results. *Time*. Retrieved from http://www.time.com/time/nation/article/0,8599,2081778,00.html

La Paro, K. M., & Pianta, R. C. (2000). Predicting Children's Competence in the Early School Years: A Meta-Analytic Review. *Review of Educational Research*, *70*(4), 443 –484.

La Paro, K. M., Pianta, R. C., & Stuhlman, M. (2004). The Classroom Assessment Scoring System: Findings from the Prekindergarten Year. *The Elementary School Journal*, *104*(5), 409–426.

Lee, V. E., Brooks-Gunn, J., Schnur, E., & Liaw, F.-R. (1990). Are Head Start Effects Sustained? A Longitudinal Follow-Up Comparison of Disadvantaged Children Attending Head Start, No Preschool, and Other Preschool Programs. *Child Development*, *61*(2), 495–507.

Lee, V. E., & Loeb, S. (1995). Where Do Head Start Attendees End up? One Reason Why Preschool Effects Fade out. *Educational Evaluation and Policy Analysis*, *17*(1), 62–82.

Leuven, E., & Sianesi, B. (2003). *PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing*.

Ludwig, J., & Miller, D. L. (2007). Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design. *The Quarterly Journal of Economics*, *122*(1), 159 –208.

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.

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., Burchinal, M., et al. (2008). Measures of Classroom Quality in Prekindergarten and Children's Development of Academic, Language, and Social Skills. *Child Development*, *79*(3), 732–749.

McKey, R. H., Condelli, L., Ganson, H., Barrett, B., McConkey, C., & Plantz, M. (1985). *The Impact of Head Start on Children, Families and Communities. Final Report of the Head Start Evaluation, Synthesis and Utilization Project.* Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402.

McLanahan, S., & Sandefur, G. D. (1994). *Growing up with a single parent: what hurts, what helps*. Cambridge: Harvard University Press.

Meyer, L. A., Wardrop, J. L., Hastings, C. N., & Linn, R. L. (1993). Effects of Ability and Settings on Kindergartners' Reading Performance. *Journal of Educational Research*, *86*(3), 142.

Milesi, C., & Gamoran, A. (2006). Effects of Class Size and Instruction on Kindergarten Achievement. *Educational Evaluation and Policy Analysis*, *28*(4), 287–313.

Molnar, A., Smith, P., Zahorik, J., Palmer, A., Halbach, A., & Ehrle, K. (1999). Evaluating the SAGE Program: A Pilot Program in Targeted Pupil-Teacher Reduction in Wisconsin. *Educational Evaluation and Policy Analysis*, *21*(2), 165-177.

Montes, G., Hightower, A. D., Brugger, L., & Moustafa, E. (2005). Quality child care and socio-emotional risk factors: No evidence of diminishing returns for urban children. *Early Childhood Research Quarterly*, *20*(3), 361–372.

National Institute of Child Health and Human Development Early Child Care Research Network. (2003). Social Functioning in First Grade: Associations with Earlier

Home and Child Care Predictors and with Current Classroom Experiences. *Child Development*, *74*(6), 1639–1662.

National Institute of Child Health and Human Development, & Network, E. C. C. R. (2002). The Relation of Global First-Grade Classroom Environment to Structural Classroom Features and Teacher and Student Behaviors. *The Elementary School Journal*, *102*(5), 367–387.

National Research Council, Committee on Early Childhood Pedagogy. (2000). *Eager to Learn: Educating Our Preschoolers*. National Academies Press.

Nelson, G., Westhues, A., & MacLeod, J. (2003). A Meta-Analysis of Longitudinal Research on Preschool Prevention Programs for Children. *Prevention & Treatment*, *6*(1).

O'Connor, E., & McCartney, K. (2007). Examining Teacher-Child Relationships and Achievement as Part of an Ecological Model of Development. *American Educational Research Journal*, *44*(2), 340–369.

Ohio University and Westinghouse Learning Corporation. (1969). The Impact of Head Start: An Evaluation of the Effects of Head Start on Children's Cognitive and Affective Development (executive summary).

Ou, S. R. (2005). Pathways of Long-Term Effects of an Early Intervention Program on Educational Attainment: Findings from the Chicago Longitudinal Study. *Journal of Applied Developmental Psychology: An International Lifespan Journal*, *26*(5), 578–611.

Ou, S. R., & Reynolds, A. J. (2010). Childhood predictors of young adult male crime. *Children and Youth Services Review*, *32*(8), 1097–1107.

Patall, E. A., Cooper, H., & Allen, A. B. (2010). Extending the School Day or School Year A Systematic Review of Research (1985–2009). *Review of Educational Research*, *80*(3), 401–436.

Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Culkin, M. L., Howes, C., Sharon Lynn Kagan, & Yazejian, N. (2001). The Relation of Preschool Child-Care Quality to Children's Cognitive and Social Developmental Trajectories through Second Grade. *Child Development*, *72*(5), 1534–1553.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, Measurement, and Improvement of Classroom Processes: Standardized Observation Can Leverage Capacity. *Educational Researcher*, *38*(2), 109–119.

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). Classroom Assessment Scoring System Manual, Pre-K. Paul H Brookes Pub Co.

Pianta, R. C., Paro, K. M. la, Payne, C., Cox, M. J., & Bradley, R. (2002). The Relation of Kindergarten Classroom Environment to Teacher, Family, and School Characteristics and Child Outcomes. *The Elementary School Journal*, *102*(3), 225–238.

Pianta, R. C., & Stuhlman, M. W. (2004). Teacher-Child Relationships and Children's Success in the First Years of School. *School Psychology Review*, *33*(3), 444–458.

Reynolds, A. J., & Ou, S. R. (2011). Paths of Effects From Preschool to Adult Well‐Being: A Confirmatory Analysis of the Child‐Parent Center Program. *Child Development*, *82*(2), 555–582.

Reynolds, A. J., Temple, J. A., & Ou, S. R. (2010). Preschool education, educational attainment, and crime prevention: Contributions of cognitive and non-cognitive skills. *Children and Youth Services Review*, *32*(8), 1054–1063.

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*.

Reynolds, A. J., Temple, J. A., White, B. A. B., Ou, S. R., & Robertson, D. L. (2011). Age 26 Cost–Benefit Analysis of the Child Parent Center Early Education Program. *Child Development*, *82*(1), 379–404.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica*, *73*(2), 417–458.

Rubin, D. B. (1980). Bias Reduction Using Mahalanobis-Metric Matching. *Biometrics*, *36*(2), 293–298.

Schweinhart, L. J., & Weikart, D. P. (1998). High/Scope Perry Preschool Program Effects at Age Twenty-Seven. In J. Crane (Ed.), *Social programs that work*. Russell Sage Foundation.

Shonkoff, J. P., & Phillips, D. A. (Eds.). (2000). *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Washington, D.C.: The National Academies Press.

Sweet, M. A., & Appelbaum, M. I. (2004). Is Home Visiting an Effective Strategy? A Meta-Analytic Review of Home Visiting Programs for Families with Young Children. *Child Development*, *75*(5), 1435–1456.

Vygotskii, L. S. (1994). *The Vygotsky Reader*. Oxford, UK: Blackwell.

White, B. L. (1993). *The First Three Years of Life* (1st ed.). New York: Fireside Book.

Wooldridge, J. M. (2008). *Introductory Econometrics: A Modern Approach* (4th ed.). South-Western College Pub.

Yen, C.-J., Konold, T. R., & McDermott, P. A. (2004). Does learning behavior augment cognitive ability as an indicator of academic achievement? *Journal of School Psychology*, *42*(2), 157–169.

Zaslow, M., Anderson, R., Redd, Z., Wessel, J., Tarullo, L., & Burchinal, M. R. (2010). *Quality Dosage, Thresholds, and Features in Early Childhood Settings: A Review of the Literature*. Office of Planning, Research and Evaluation, Administration for Children and Families.

Zill, N. (2010). Investing in young children. In R. Haskins & B. W. Steven (Eds.), *Investing in Young Children: New Directions in Federal Preschool and Early Childhood Policy*. (pp. 39–48). Washington, DC: Brookings Institution, Center on Children and Families.

CURRICULUM VITAE

Pete M. Bernardy received a Bachelor of Arts degree in Economics from the University of Saint Thomas in 2002. He received a Master of Public Policy degree from the University of Minnesota's Humphrey School of Public Affairs in 2004. Upon completion of this degree, Pete entered federal service as a Presidential Management Fellow in the central budget office of the United States Department of Health and Human Services (HHS) where he has served for the past eight years. Pete became a member of the federal government's Senior Executive Service in 2011.