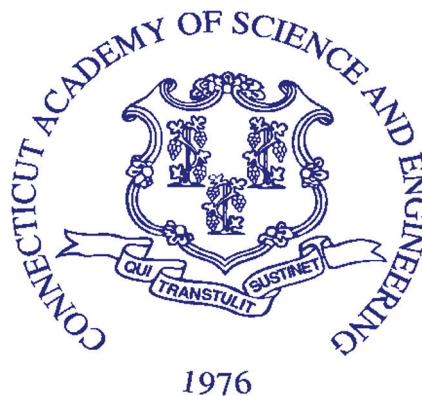


EARLY CHILDHOOD REGRESSION DISCONTINUITY STUDY

JUNE 2016

A REPORT BY

THE CONNECTICUT
ACADEMY OF SCIENCE
AND ENGINEERING



FOR

THE CONNECTICUT GENERAL ASSEMBLY
AND THE
EDUCATION COMMITTEE

EARLY CHILDHOOD REGRESSION DISCONTINUITY STUDY

A REPORT BY

THE CONNECTICUT ACADEMY
OF SCIENCE AND ENGINEERING

ORIGIN OF INQUIRY:

THE CONNECTICUT
GENERAL ASSEMBLY

DATE INQUIRY

ESTABLISHED:

MAY 9, 2014

DATE RESPONSE

RELEASED:

JUNE 6, 2016

This study was initiated at the request of the Connecticut General Assembly on May 9, 2014. The project was conducted by an Academy Study Committee with the support of faculty from the Neag School of Education, University of Connecticut with Bianca Montrosse-Moorhead, PhD serving as study manager. W. Steven Barnett, PhD, National Institute for Early Education Research, Rutgers University and Mary Beth Bruder, PhD, UConn Health served as study advisors. The content of this report lies within the province of the Academy's Education and Human Resources, and Technology Technical Boards. The report has been reviewed by Academy Members Theodore Holford, PhD and Nalini Ravishanker, PhD. Martha Sherman, the Academy's Managing Editor, edited the report. The report is hereby released with the approval of the Academy Council.

Richard H. Strauss
Executive Director

COMMITTEE MEMBERS FOR THE EARLY CHILDHOOD REGRESSION DISCONTINUITY STUDY

Elizabeth Aschenbrenner

Education Consultant; School Readiness
Liaison: Killingly, Plainfield, Putnam; Director
of Early Childhood Initiatives, EASTCONN

Regina S. Birdsell

Assistant Executive Director
Connecticut Association of Schools

Gary Henry, PhD

Patricia and Rodes Hart Professor of Public
Policy and Education, Dept. of Leadership,
Policy & Organizations, Vanderbilt University

Jessica Powell, PhD

Assistant Professor, Elementary Programs
Southern Connecticut State University

Chin Reyes, PhD

Associate Research Scientist
The Edward Zigler Center in Child
Development & Social Policy, Yale School of
Medicine

Wendy Rayack, PhD

Associate Professor of Economics
Wesleyan University

Sudha Swaminathan, PhD

Professor, Early Childhood Education
Eastern Connecticut State University

William Teale, EdD

Professor, Director of Center for Literacy
University of Illinois at Chicago

RESEARCH TEAM

NEAG SCHOOL OF EDUCATION, UNIVERSITY OF CONNECTICUT

STUDY MANAGER

Bianca Montross-Moorhead, PhD, Assistant Professor,
Measurement, Evaluation and Assessment

Shaun Dougherty, EdD, Assistant Professor, Educational Policy and Leadership

Hannah Dostal, PhD, Assistant Professor, Literacy Education

Tamika La Salle, PhD, Assistant Professor, School Psychology Program

Jennie Weiner, PhD, Assistant Professor, Educational Leadership

STUDY ADVISORS

W. Steven Barnett, PhD, Director, National Institute for Early Education Research,
Rutgers University

Mary Beth Bruder, PhD, Professor of Pediatrics; Director, The A.J. Pappanikou Center for
Excellence and Developmental Disabilities Research, Education, and Service,
UConn Health Center

ACADEMY PROJECT STAFF

Richard H. Strauss, Executive Director

Terri Clark, Associate Director

Ann G. Bertini, Assistant Director for Programs

TABLE OF CONTENTS

TABLE OF CONTENTS.....	v
EXECUTIVE SUMMARY.....	vii
1.0 STUDY BACKGROUND.....	1
2.0 INTRODUCTION AND STUDY METHODS.....	3
3.0 IMPLEMENTATION PLAN.....	27
4.0 DATA COLLECTION AND ANALYSIS.....	37
5.0 DISCUSSION AND IMPLICATIONS	49
REFERENCES.....	59
APPENDICES.....	67
APPENDIX A: THE RESEARCH TEAM.....	67
APPENDIX B: SUMMARY OF AGE CUT-OFF RD STUDIES INVESTIGATING THE EFFECTS OF PUBLICLY FUNDED PREKINDERGARTEN PROGRAMS ON CHILDREN’S ACADEMIC AND SOCIAL-EMOTIONAL SKILLS.....	69
APPENDIX C: PARENT AND TEACHER QUESTIONNAIRES	79
APPENDIX D: UCONN’S INSTITUTIONAL REVIEW BOARD (IRB) APPROVALS.....	99
APPENDIX E: ANALYSIS OF POPULATION AND SAMPLES, AND OF TREATMENT AND CONTROL GROUPS.....	108
APPENDIX F: FINAL POWER ANALYSIS.....	114
APPENDIX G: ANALYSIS OF BASC-3 RESPONSES	115
APPENDIX H: ADDITIONAL TESTS OF ESTIMATE ROBUSTNESS.....	116

EXECUTIVE SUMMARY

WHAT WAS THE PURPOSE OF THIS STUDY?

This evaluation study was conducted by the Connecticut Academy of Science and Engineering (CASE) on behalf of the Connecticut General Assembly (CGA) at the request of the Education Committee. The purpose of this study is to investigate the immediate effects associated with children who attend Connecticut's state-funded School Readiness full-day or school-day prekindergarten program. The primary research questions include:

1. Do children who attend full-day or school-day, state-funded preschool programs enter kindergarten with better language and literacy skills than if they had not attended the program?
2. Do children who attend full-day or school-day, state-funded preschool programs enter kindergarten with better mathematics skills than if they had not attended the program?
3. Do children who attend full-day or school-day, state-funded preschool programs enter kindergarten with better social skills than if they had not attended the program?

WHAT DID THE STUDY FIND?

The findings show evidence that attending state-funded prekindergarten in Connecticut, as delivered through the School Readiness program funding stream, positively impacts students' early literacy and early numeracy skills. An overview of results is included in Table ES.1.

TABLE ES.1. OVERVIEW OF IMPACT OF STATE-FUNDED PRESCHOOL PROGRAMS

Claim	Test Evidence	Measures	What Students Do on this Test	Skill Focus
Large, positive and statistically significant effects on a subset of student's early literacy skills (0.69 SD)	Basic reading	WJ-IV: Letter-word identification	Recognizing and naming printed letters and words	Letter/word recognition
		WJ-IV: Word attack	Reading made-up words that conform to conventional spelling rules	Phonemic awareness
Large, positive and statistically significant effects for most student's early numeracy skills (0.48 SD)	Broad math	WJ-IV: Calculations	Arithmetic computation with paper and pencil	Writing numbers to numerical operations
		WJ-IV: Math Fluency	Simple calculations for three minutes	Quickly solving numerical operations
		WJ-IV: Applied Problems	Oral, math "word problems," solved with paper and pencil	Math problem solving
Suggested positive, but non-statistically significant, effects on student's early vocabulary skills	Picture vocabulary	PPVT-IV	Listening to a word describing one of four pictures and then pointing to the picture that the word describes	Picture-to-word recognition
Suggested positive, but non-statistically significant, effects on student's early oral language skills	Oral Comprehension	WJ-IV: Picture Vocabulary	Listening to a word describing one of four pictures and then pointing to the picture that the word describes	Picture-to-word recognition
		WJ-IV: Oral Comprehension	Listening to an oral passage and identifying a missing key word that makes sense	Listening comprehension

TABLE ES.1. (CONTINUED)

Claim	Test Evidence	Measures	What Parents and Teachers Do on this Test	Skill Focus
Unknown effects for student's early <i>social</i> skills	Social Development	BASC-3	Answer survey questions	Student Externalizing Problems
				Student Internalizing Problems
				Student Behavioral Symptoms Index
				Student Adaptive Skills
<i>Note: Effect sizes are included in the second column of this table in parentheses only for outcomes that are statistically and practically significant. Woodcock-Johnson, Fourth Edition (WJ-IV). Peabody Picture Vocabulary Tests, Fourth Edition (PPVT-IV). Behavior Assessment Scale for Children, Third Edition (BASC-3)</i>				

Considering Some Trade-Offs

Though these results are promising, as is typical with any study, there are trade-offs regarding the scope of the conclusions based on the research design. For this study, the Regression Discontinuity design (RD) that was used allowed for a causal interpretation of the data. This made the findings far more powerful than a simple correlative study. However, the ability to make such claims came with some trade-offs regarding the conclusions that could be drawn from the findings. Table ES.2 highlights what can and cannot be concluded for the present RD study.

TABLE ES.2: WHAT CAN AND CANNOT BE CONCLUDED FROM THIS STUDY

Can Conclude from RD	Cannot Conclude from RD
<ul style="list-style-type: none"> On average, the School Readiness full-day or school-day prekindergarten programs makes a positive difference in the areas identified as statistically significant. Specifically, prekindergarten students who attend School Readiness full-day or school-day programs do better, on average, in early literacy and early numeracy. 	<ul style="list-style-type: none"> What about the School Readiness program makes a positive impact? What is the best of all possible School Readiness programs? Which aspects of the School Readiness program generated the most valuable outcomes?

However, it is important to note that the findings indicate that being enrolled in the School Readiness full-day or school-day program produces positives results in early literacy and early numeracy skills for students, on average.

HOW IMPORTANT ARE THESE FINDINGS?

Although it has been standard practice for researchers, policymakers, educators, program staff, and other key stakeholder groups to use Cohen's (1988) benchmarks to draw inferences about whether the size of an effect is substantively important, this study follows that of methodological leaders (Cooper, Hedges, & Valentine, 2009; Hill, Bloom, Black, & Lipsey, 2008) who argue that more appropriate inferences can be drawn using other benchmarks. To assist readers of this report in drawing inferences regarding the importance of findings, effect sizes for this study are compared to effect size results from past research in three different ways.

First, effect size benchmarks calculated by Hattie (2009) were used. As illustrated in Figure ES.1, both benchmarks further support the claims that large, positive, and statistically significant effects on student's early literacy and numeracy skills were detected and are noteworthy, with early literacy and early numeracy skills effect sizes both above these benchmarks.

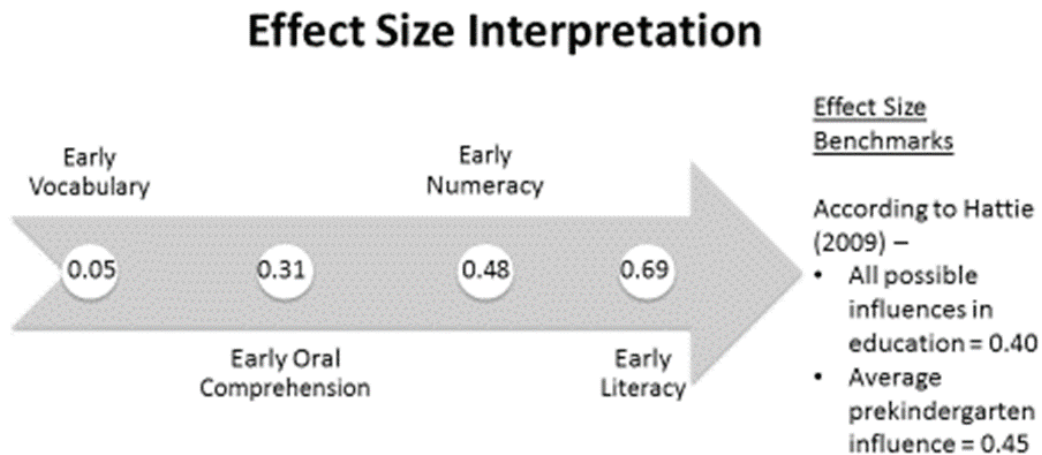


FIGURE ES.1: INTERPRETING CONNECTICUT EFFECTS IN RELATION TO HATTIE'S (2009) BENCHMARKS

Second, effect sizes were descriptively contextualized in relation to those found in other statewide prekindergarten evaluations (Figure ES.2). In examining results from Figure ES.2, it should be noted that this study was the first to use composite outcome measures. This is a strength of this study because it represents an outcome that assesses a wider content area than has been assessed in prior prekindergarten, statewide impact studies that used an RD design. At the same time, this creates a situation where one-to-one effect size comparisons are unavailable, as other states only looked at one sub-test (e.g., applied problems). Effect sizes included in Figure ES.2 for early numeracy are those that came only from the applied problems sub-test, as opposed to the Broad Math outcome used in this study. Similarly, as opposed to the Basic Reading outcome used in this study, effect sizes from other states for early literacy are those that came only from the letter-word sub-test.

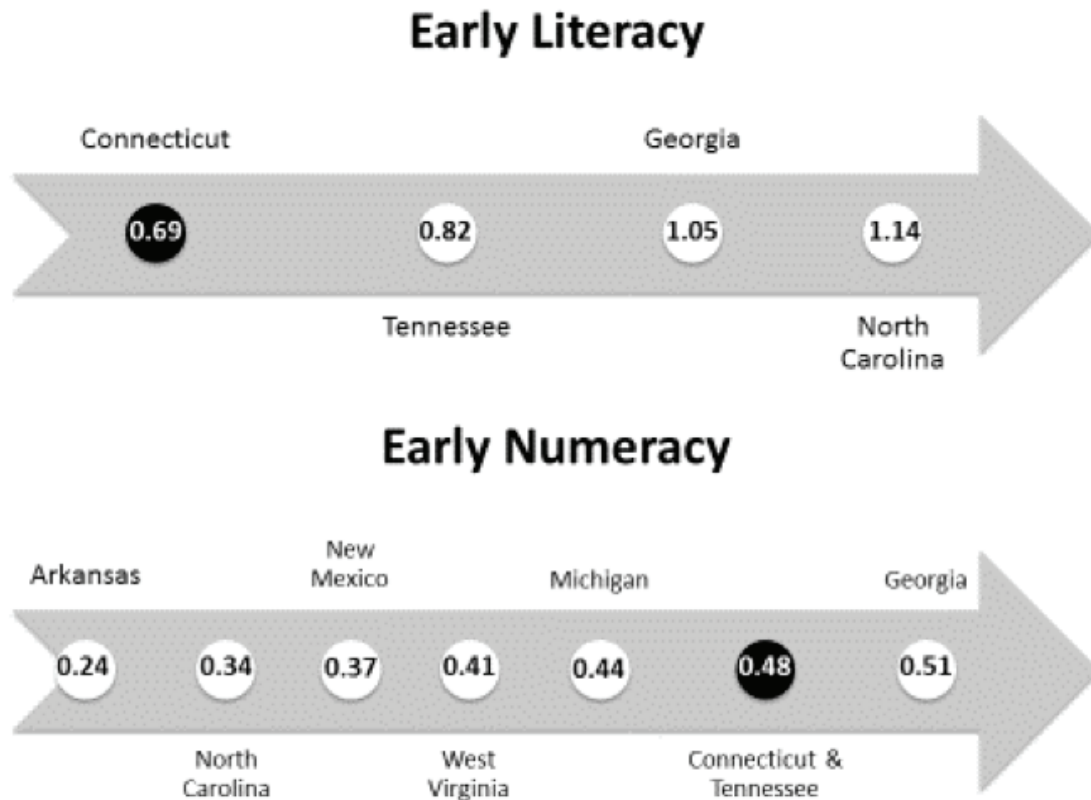


FIGURE ES.2: INTERPRETING CONNECTICUT EFFECTS IN RELATION TO PREKINDERGARTEN EFFECTS FOUND IN OTHER STATES

Third, effect sizes reported in Figures ES.1 and ES.2 can be descriptively compared to other prior research studies. For example, effect sizes reported for other state-funded prekindergarten programs range from .23–.53 (Gilliam & Zigler, 2001), and prekindergarten programs generally from .10 to .13 (Magnuson, Ruhm, & Waldfogel, 2004). Those reported for high-quality childcare programs seldom exceed .10 (NICHD Early Child Care Research Network & Duncan, 2003; Peisner-Feinberg et al., 2011). The Abecedarian project, widely acknowledged as a highly successful early intervention program, reported effect sizes of .73 and .79 for children ages 4 and 5 years old (Ramey, Campbell, Burchinal, Skinner, Gardner, & Ramey, 2000), and the highly praised Perry Preschool program reported effect sizes of .60 (Ramey, Bryant, & Suarez, 1985).

WHAT SHOULD BE DONE AS A RESULT?

Recommendations for Future Evaluation Questions

The findings from this study suggest the need for further studies regarding some of the mechanisms that helped to produce these results, as well as the non-findings. Table ES.3 includes recommendations for further exploration regarding these findings that may be of value.

TABLE ES.3: FUTURE EVALUATION QUESTIONS FOR CONNECTICUT BASED ON STUDY FINDINGS

Question Category	Potential Evaluation Questions
What works?	<ul style="list-style-type: none"> • Do replication studies support impact study findings across different cohorts of students? • Do longitudinal replication studies support impact study findings long-term? • Do children who attend full-day or school-day, state-funded preschool programs enter kindergarten with better social skills than if they had not attended the program?
What works for whom?	<ul style="list-style-type: none"> • Do results vary by state-funded preschool program type? • Do results vary by student characteristics (e.g., gender, race/ethnicity, income)? • Do results vary by student skill level (e.g., English proficiency)?
What works, for whom, and under what conditions?	<ul style="list-style-type: none"> • Do results vary by program quality? • Do results vary by the amount of school/system instructional support?
Which aspects are valuable?	<ul style="list-style-type: none"> • What is the relationship between program costs and outcomes observed? • Which aspects of the School Readiness program generated the most valuable outcomes?

Recommendations for Commissioning Future Statewide Prekindergarten Studies

This study represented the first statewide study of prekindergarten in Connecticut. The following lessons learned from this study will be useful in conducting future evaluation studies.

LOW PARTICIPATION

Historically, the state has been committed to a governance model that allows for local decision-making regarding participation, and ethical concerns regarding the ability of a parent/guardian to determine the participation of their child participate in this type of study. As this study demonstrated, this model contributed to a reduction in participation rates of districts, schools, centers, and students. Options for strategies to be considered to increase participation rates are included in Table ES.4:

TABLE ES.4: STRATEGIES FOR ADDRESSING LOW PARTICIPATION

Strategy	Pro	Con
Mandate that schools and centers receiving state funding for prekindergarten programs and kindergarten participate in state-mandated studies	<ul style="list-style-type: none"> • Greater participation 	<ul style="list-style-type: none"> • Would require schools and centers to notify parents/legal guardians of testing for such studies in the ways that are consistent with how parents are notified of annual state assessment testing periods • Would require a change in Connecticut's governance model in which decision-making would be shifted back to the state
Require schools and centers to file a letter of cooperation with OEC and CSDE indicating their willingness to participate in state-mandated studies	<ul style="list-style-type: none"> • Potential for greater participation 	<ul style="list-style-type: none"> • Limits generalizability to those that filed a letter. • Does not necessarily address the question site-based opt outs.
Incorporate potential for low participation rates in the study planning process	<ul style="list-style-type: none"> • Minimize concerns about study power beforehand 	<ul style="list-style-type: none"> • Increased monetary costs by increasing overall sample.
Consider alternative levels of stipends or alternative stipend disbursement methods	<ul style="list-style-type: none"> • Potential for greater participation for parent/guardian and teachers 	<ul style="list-style-type: none"> • Increased monetary costs • No research to inform incentive amount • No research specific to teachers or parent/guardian
Alternative parent/legal guardian data collection strategies (e.g., telephone administration of surveys)	<ul style="list-style-type: none"> • Potential for greater participation for parent/guardian 	<ul style="list-style-type: none"> • Increased monetary costs • Shorter, potentially less informative surveys • Requires sharing parental contact and address information

STUDENT-LEVEL DATA

OEC and CSDE eliminated the Prekindergarten Information Management System (PKIS) as of the summer 2014 in anticipation of replacing it with a new data collection system to be administered by OEC. The PKIS information previously collected was paramount to this study, a short term alternative student data collection process was developed by the Research Team/CASE in cooperation with OEC. The following suggestions should be considered to support future evaluation studies:

- Provide user-friendly mechanisms to facilitate the efficient transfer of school and center student data for both prekindergarten and kindergarten in a timely manner. For this study, having information such as classroom rosters (for students who attended prekindergarten last year that are attending kindergarten this year) and student demographic data (e.g. race/ethnicity, socio-economic status, gender, etc.) much sooner would have accelerated the data collection timeline and may have increased participation rates. It is noted that for the analyses conducted for this study, statistical controls were included to address shifts in the study's timeline, but it would have been preferable to be able to begin data collection in schools and centers much earlier in the school year.
- Include student demographic information (e.g. race/ethnicity, socio-economic status, gender, etc.) necessary to conduct this type of evaluation study in the new student data collection system to be administered by OEC.

Recommendations for Funding Future Research Studies

In addition to state-funded research studies, federal funding sources may be available to support studies to answer the questions cited in Table ES.3 including the U.S. Department of Education's Institute for Education Sciences (Evaluation of State and Local Education Programs and Policies program, Preschool Curriculum Evaluation Research program, and Early Learning Programs and Policies program), and the U.S. Department of Health and Human Service's National Institutes of Health.

HOW WERE THESE CONCLUSIONS AND RECOMMENDATIONS DRAWN?

A random sample of 529 students (40.7 % compliance rate) who attended the full-day or school-day, state-funded School Readiness prekindergarten program during the 2015-16 and 2014-15 were assessed using two standardized, psychometrically sound instruments: Woodcock-Johnson, Fourth Edition (WJ-IV) and Peabody Picture Vocabulary Tests, Fourth Edition (PPVT-IV).

Data were then statistically analyzed within a RD framework. Specifically, an RD approach can be used when there is a clear external means of distinguishing between two groups in such a way that the only difference between these groups is that some get a "treatment" and some do not. In other words, the two groups are treated as if they were randomly assigned and that the individuals within them are "equal in expectation" (i.e., they are only different because some are assigned to the treatment and some are not).

For this to occur, first, there has to be a treatment to which individuals are selected (e.g., prekindergarten). Second, the selection criteria have to be externally created values on a numeric rating (e.g., an age limit or requirement). By properly controlling for the value of the rating variable in the RD design, any unobserved differences between the treatment and comparison group can be accounted for. In other words, it can be assumed that the children very close to the cut-off for the rating variable are the same in all ways but their numeric score.

In essence, because the RD design relies on the use of some type of cut-off, it makes it both feasible and ethical to implement across a wide-variety of situations and allows for answers to the question, on average, across a group of people, of “what works?” In the last nine years, 16 studies have employed an age cut-off RD approach in evaluations of state-funded prekindergarten programs (Appendix B). Figure ES.3 illustrates how this process worked for this study of prekindergarten students using birth date as the cut-off mechanism, which is consistent with best practice in prekindergarten impact studies.

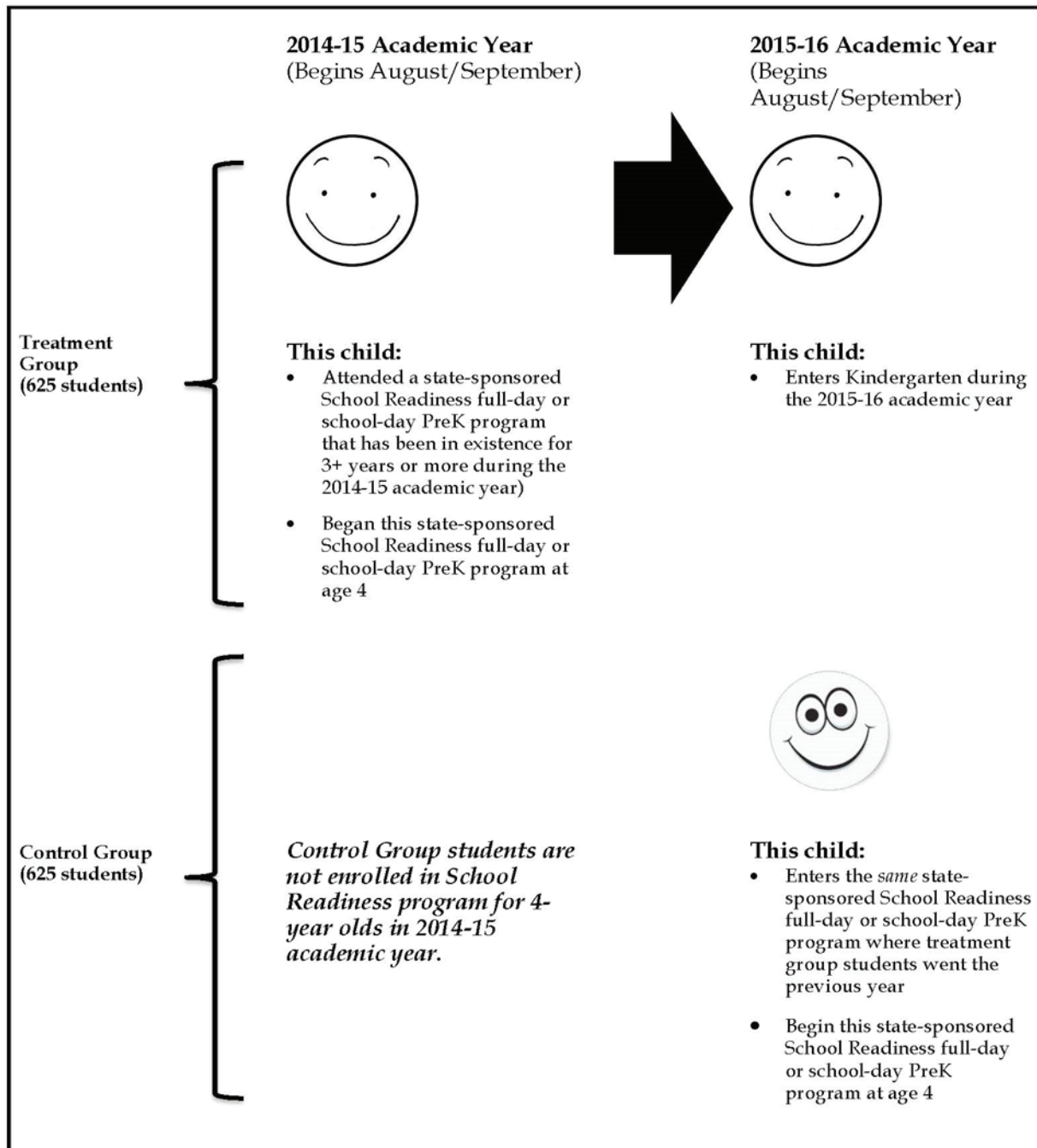


FIGURE ES.3: HOW THE RD PROCESS WORKS IN THIS STUDY

1.0 STUDY BACKGROUND

This study was conducted by the Connecticut Academy of Science and Engineering (CASE) on behalf of the Connecticut General Assembly (CGA) at the request of the Education Committee. The purpose of the study is to identify the effect that full-day/school-day, state-funded preschool has on children's academic achievement and social skills at kindergarten entry. Additionally, other CGA committees and state agencies with an interest in the study included the Appropriations Committee, Committee on Children, the Office of Early Childhood (OEC) and the Connecticut State Department of Education (CSDE). Other stakeholders included parents of children served by state-funded preschool programs and personnel employed by state-funded preschool programs, as well as other interested parties.

In 2008, a concept for an Early Childhood Education Study was developed as a component of a multi-part early childhood research agenda articulated in the *Connecticut Early Childhood Investment Plan, Part I* (November 2006). In preparation for conducting this study on behalf of the state, the National Institute for Early Education Research (NIEER) prepared a report entitled "*Prekindergarten Effectiveness Research Designs for the State of Connecticut*" (Riley-Ayers, Barnett, and Frede, 2008). The report examined the use of methodologies to be considered based on the questions the state was seeking to answer and available funding. A Regression Discontinuity research design (RD design) was selected as the most feasible and effective methodology for the study and is further described in section 2.3.3.1. This research design was intended to estimate the effects of state-funded prekindergarten programs on children's academic achievement and social skills at kindergarten entry (Schechter, 2008). A request for proposals (RFP #DHE0901, 2008) was developed but not issued due to funding issues.

The referenced NIEER report and additional information developed in preparation for the 2008 proposed study was used as a foundation for formulating the research methodology and work plan for this study.

1.1 PROJECT OVERVIEW

The major tasks for this study include research plan preparation, implementation, data collection and analysis, and production of a final report.

Study process included

- Study committee meetings
- Research using identified methods
- Development of a final report including study findings and recommendations based on the research

The study report contains the following sections

- Study Background
- Introduction and Study Methods
- Implementation
- Data Collection and Analyses
- Discussion and Implications

1.2 STUDY COMMITTEE AND RESEARCH TEAM

A Study Committee was appointed by CASE to oversee and guide the study effort. A research team from the Neag School of Education at the University of Connecticut was engaged by CASE to conduct the research for the study, with support from CASE study advisors. See Appendix A for information about the UConn Research Team.

In accordance with the CASE Agreement with the General Assembly (JCLM14PSA0018 as amended to extend the contract term to June 1, 2016), Public Act 15-5, Section 473 provided a carryover of the balance due on the contract; and Public Act 15-244, Section 44 provided an increase in the total cost of the study for additional work effort and study costs. The leadership of CGA committees and state agencies with an interest in this study were kept informed about the study. They were invited to all study committee meetings and provided with access to study research materials and audio recordings of most study committee meetings, including expert guest speakers.

2.0 INTRODUCTION AND STUDY METHODS

2.1 INTRODUCTION

Does prekindergarten matter for student success? That question has been fiercely debated and the subject of numerous evaluations since the 1960s (c.f., Lipsey, Weiland, Yoshikawa, Wilson, & Hofer, 2015; Whitehurst, 2014). From a policy perspective, two competing arguments have dominated much of this discourse. The first argument maintains that prekindergarten promotes school readiness, especially for children who are most in need and whose parents cannot afford high-quality private providers, and thereby minimizes the opportunity gap for children entering kindergarten (Pianta, Cox, & Snow, 2007). Advocates of this position argue that for this reason, universal state-funded prekindergarten access should be guaranteed, if not required, for all children (c.f., Strong Start for America's Children Act of 2013)¹ Others maintain that the benefits of prekindergarten have yet to be universally documented and call for a moratorium on legislation to extend prekindergarten programming, universal or otherwise (Whitehurst, 2014). Clearly, more research is needed not only to answer questions concerning whether or not prekindergarten matters, but also for whom, under what conditions it matters most, and for how long effects endure.

Although interest in prekindergarten has received national attention recently, spurred in part by President Obama's call for universal prekindergarten for all 4-year-olds, interest in prekindergarten can be traced back to 1965 with the launch of Head Start programs. Currently, 40 states and the District of Columbia offer some type of publicly funded prekindergarten programming (Barnett, Carolan, Fitzgerald, & Squires, 2012). Further, since 2003, the National Institute for Early Education Research (NIEER) at Rutgers University has annually chronicled the state of state-funded prekindergarten programs across the nation. In 2012, NIEER documented that there is downward trend nationally in prekindergarten program spending and quality (Barnett, Carolan, Fitzgerald, & Squires, 2012).

Results from the most recent report highlight areas of continued concern, as well as areas of progress. According to Barnett et al. (2013), compared to the prior academic year, in participating states, the number of children enrolled in prekindergarten programs during the 2012-13 academic year decreased. This was the first time a decrease in state-funded prekindergarten programs was observed since NIEER's first annual report in 2003. Though alarming, it is important to note that while the total number of children served decreased, the proportion of children enrolled did not change. In total, a little over a quarter (28% or 1.1 million) of 4-year-olds and approximately 4% (or 175,999) of 3-year-olds attended state-funded prekindergarten programs during the 2012-13 academic year. This percentage has held steady since 2010-11 for 4-year-olds and since 2007-08 for 3-year-olds.

In examining expenditure trends compared to the previous academic year, total state expenditures during the 2012-13 academic year increased approximately 1%; however, this increase was not enough to rebound fully from the half billion-dollar reduction in spending observed nationally in 2011-12. Lastly, for the first time, all 50 states and the District of

¹ <http://beta.congress.gov/bill/113th-congress/house-bill/3461>.

Columbia met the prekindergarten program quality standards² of implementing comprehensive early learning standards. This indicates that, for the first time since 2003, all states “have comprehensive early learning standards covering all areas identified as fundamental by the National Education Goals Panel – children’s physical well-being and motor development, social/emotional development, approaches toward learning, language development, and cognition and general knowledge” (Barnett, Carolan, Squires, & Clarke Brown, 2013, p. 23).

2.1.1 Overview of Connecticut’s Prekindergarten Program

The latest NIEER report (2013-14) on Connecticut’s prekindergarten programs suggests that the state follows national trends despite efforts to provide greater access to programming. The report indicates that during the 2013-14 academic year, 5,381 4-year-olds and 2,347 3-year-olds were enrolled in state prekindergarten programs. This represented approximately 10% of the total 4- and 3-year-old population living in Connecticut (or approximately 14% of 4-year-olds and 6% of 3-year-olds). Similar to national trends, the NIEER report documents a 1% increase in 3- and 4-year-olds served as compared to 2012-13. Also, in terms of prekindergarten program spending, in 2013-14, Connecticut ranked third when examining only state expenditures, as well as total expenditures. Additionally, in 2013-14, Connecticut met six of the NIEER prekindergarten program quality standards: comprehensive early learning standards; specialized training in prekindergarten; class sizes no larger than 20 children; a staff-child ratio of 1:10 or better; vision, hearing, and health screenings and referrals; and site visits and other monitoring.

Prekindergarten program expansion is poised to expand in Connecticut over the next five years³. Governor Dannel P. Malloy recommended that new prekindergarten program slots be added beginning in FY 2015 and ending in FY 2019. Specifically, 1,020 slots were added for FY 2015, and expansion is planned to continue through FY 2019, at which time 4,010 additional slots will have been added. The intent of this expansion is to move Connecticut one step closer toward universal access to prekindergarten programs, with low-income children given preference.

The prekindergarten market is diverse in Connecticut. Options are provided through School Readiness funding, Head Start funding, and choice options (e.g., magnet schools, charter schools, and private school options). An overview of each of these options follows.

2.1.2 Current Prekindergarten Options in Connecticut

According to the Connecticut State Department of Education (CSDE), Public Act 97-259, An Act Concerning School Readiness and Child Day Care⁴ encoded in the Conn. Gen. Stat. §10-16p⁵ established a grant program to provide prekindergarten program spaces for eligible children residing in 19 priority school districts and 50 designated competitive grant municipalities⁶. Municipalities are eligible for competitive funds if located in one of the 19 priority school districts or are ranked among the 50 lowest-wealth districts in the state.

² NIEER has developed a set of 10 research-based prekindergarten program quality standard benchmarks.

³ <http://www.ct.gov/oec/cwp/view.asp?a=4547&q=540158>.

⁴ <http://www.cga.ct.gov/ps97/Act/pa/1997PA-00259-R00HB-05461-PA.htm>.

⁵ <http://law.justia.com/codes/connecticut/2012/title-10/chapter-164/section-10-16p>.

⁶ <http://www.sde.ct.gov/sde/lib/sde/PDF/DEPS/Readiness/sroverview.pdf>.

The School Readiness program offers four types of slots, with student exposure being defined as follows:

- Full-day/full-year: 10 hours per day for 50 weeks per year
- School-day/school-year: 6 hours per day for 180 days per year
- Part-day/part-year: 2.5 to 5.5 hours per day for between 180 to 250 days per year
- Extended-day: provides extended services to meet full-day/full year requirements

The national Head Start program, established in 1965, provides services to children (age birth to 5) and their families. The federal Head Start program is housed within the Administration for Children and Families in the US Department of Health and Human Services. The aim of this program is to promote school readiness for children from low-income families. In the context of Connecticut, Head Start Centers, regardless of whether they receive federal or state monies, are operated by community agencies, local education agencies, and nonprofit agencies⁷.

The Connecticut Head Start program offers five types of slots with student exposure being defined as follows:

- Federal/full-time: more than 6 hours per day for 4 to 5 days per week year-round
- Federal/part-time: between 3.5 and 6 hours per day for 4 to 5 days per week for a minimum of 32 weeks
- Federal/extended-day: provides extended services to meet federal/full-time requirements
- State/full-time: more than 6 hours per day for 4 to 5 days per week year-round
- State/part-time: between 3.5 and 6 hours per day for 4 to 5 days per week year-round

Even Start is a federally funded program, first authorized in 1998, designed to improve family literacy⁸. Low-income, at-risk families are the target demographic for this program. Eligibility criteria include having a child birth to age eight, at least one parent lacking a high school diploma and/or basic English literacy skills, and documentation of low-income status. Program components include services for children, including early childhood/school-age education and parent/child playgroups, as well as services for their families, including adult basic education, parenting education and support, and home visits to assist families with creating a literacy-rich home learning environment. In terms of the prekindergarten component of Even Start in Connecticut, 2.5- to 3-year-olds are required to attend an average of 36 hours or more per month. Children ages 4 and 5 are required to attend an average of 60 hours or more per month⁹.

In addition to these federally funded initiatives, there are also state-based options for prekindergarten. According to the CSDE, “students in interdistrict magnet schools come together to learn in educational settings that offer a range of themes or teaching philosophies,

7 <http://www.sde.ct.gov/sde/cwp/view.asp?a=2626&q=320766>.

8 According to the U.S. Department of Education website, “the program became State administered in 1992” (<http://www2.ed.gov/programs/evenstartformula/index.html>).

9 <http://www.sde.ct.gov/sde/cwp/view.asp?A=2620&Q=320690>

including science, technology, engineering and math (STEM); environmental science; arts; Montessori; and global and international studies.”¹⁰ As of 2013-14, a total of 35 magnet schools, out of a total of 43 elementary schools (PK3-Grade 5) in operation across the state, offer prekindergarten options. Charter schools, which according the CSDE, “are publicly funded schools that are run independently of the school district... [whose] purpose is to establish an alternative means of education within the existing public school system...” also provide options. Currently, six schools, out of 26 Connecticut charter schools offered prekindergarten options during the 2013-14 academic year.¹¹ The type of slots available and the number of hours children attend interdistrict magnet and charter school prekindergarten programs is, at this time, unknown. Finally, parents may also choose to enroll their children in faith-based and private school (non-charter) prekindergarten programs. The number of these types of programs, the demographics of the children who attend them, and the type of slots available are unknown.

2.1.3 Purpose of the Evaluation Study

The purpose of this study’s evaluation is to investigate the immediate effects associated with children who attend state-funded preschool in Connecticut. As a first step to answer this question, the study was restricted to the fullest implementation of Connecticut’s state-funded School Readiness program – that is, the effect of prekindergarten for 4-year-old children who attend full-day or school-day, state-funded preschool in Connecticut. Further, the study includes only prekindergarten programs with full-day or school-day programming¹² funded through the School Readiness program, as they represent the fullest dose of the treatment (i.e., prekindergarten programming) available. An important caveat to note is that the focus of this study, in accordance with the scope of work authorized by the CASE/CGA project agreement, includes addressing only the following primary questions:

1. Do children who attend full-day or school-day, state-funded preschool programs enter kindergarten with better language and literacy skills than if they had not attended the program?
2. Do children who attend full-day or school-day, state-funded preschool programs enter kindergarten with better mathematics skills than if they had not attended the program?
3. Do children who attend full-day or school-day, state-funded preschool programs enter kindergarten with better social skills than if they had not attended the program?

Hence, this study should be viewed as a useful *initial* step in understanding the impact of prekindergarten programs in the state. The study will not be able, nor will it attempt, to make claims about prekindergarten programs across the state beyond full-day or school-day school-readiness preschool programs.

10 <http://www.sde.ct.gov/sde/cwp/view.asp?a=2681&q=335070&sdePNavCtr=|#45453>

11 <http://www.sde.ct.gov/sde/cwp/view.asp?a=2681&q=335076#Bridgeport>

12 The definition of full-day or school-day used in this study does not include extended-day programming. While this program option is intended to extend program duration for students already enrolled in Part-Day/Part-Year programming, this study’s research indicates that programming services vary widely (e.g., some include additional enrichment activities, while others simply provide socialization activities for children). Thus, it cannot be guaranteed that children who participate in these programs experience the “fullest dose” in the same way as students that attend full-day/full-year and school-day/school-year programming.

2.2 REVIEW OF PRIOR PREKINDERGARTEN EVALUATION STUDIES USING A REGRESSION DISCONTINUITY DESIGN APPROACH

Interest in the effects of prekindergarten generally can be traced back to the 1960s, while investigating the effects of state-funded prekindergarten programs through rigorous methods has been a more recent area of research conducted in response to policy questions. Much of this recent literature employs the use of RD design in an effort to provide claims about the *causal* impact of prekindergarten (See section 2.3.3.1 of this methodology). Because this evaluation study also uses this type of research design, the review of extant literature conducted was restricted to studies that have employed an age cut-off RD approach.

In the last nine years, 16 studies have employed an age cut-off RD approach in evaluations of state-funded prekindergarten programs. Gormley and colleagues (Gormley, 2008; Gormley & Gayer, 2005; Gormley, Phillips, & Gayer, 2008; Gormley, Gayer, Phillips, & Dawson, 2005) were the first to use the design in Tulsa, Oklahoma. Not only did Gormley and colleagues show that, on average, children that attended prekindergarten demonstrated improved outcomes, but also that the effects varied by subgroup. Specifically, larger, positive effects were shown for children in low-wealth families and for students from historically marginalized racial and ethnic groups.

Gormley and colleagues were not alone in documenting positive effects. In total, across all 16 studies published between 2005 and 2014, 84 outcome measures have been assessed. These outcome measures include literacy and language, mathematics, social skills, emotional skills, behavior, and general school readiness indices. Out of the 84 outcome measures assessed, 67 were both statistically significant and practically meaningful (as shown through effect size estimates) for students attending prekindergarten, and 17 were neither statistically significant nor practically meaningful. Importantly, none were shown to produce statistically significant and negative results, meaning that there was no indication that students who attended prekindergarten did *worse* than those who did not attend prekindergarten.

Appendix B summarizes key aspects of the 16 studies that employed an age cut-off RD approach to assess publicly funded prekindergarten programs within and across states.

2.2.1 *Non-Technical Review of the RD Approach*

Policy makers, teachers, parents, and other stakeholders, when considering educational improvement, commonly pose the question of “what works?” For example, policy makers frequently ask questions about whether an enacted policy is having the intended effect. Teachers often question whether using a particular instructional tool will help their students grasp the content better. Parents often question the best enrichment activities or interventions to meet their child’s needs and development. All of these questions are “what works?” questions and, given the multitude of competing views on each of these issues, it can be difficult to answer such questions effectively or accurately.

Advances in research methods and statistics offer a number of enhanced strategies to answer these types of questions in the real world. In fact, these strategies are used in a variety of studies, including medicine, public health, and education. One such strategy that is effective for answering questions related to “what works?” is the RD design. A non-technical description of the RD design that includes how this method is able to provide insights into what works for

most children, and in the case of this study, the impact, on average, of prekindergarten on a child's kindergarten readiness¹³, is presented in the next section.

2.2.1.1 WHAT IS AN RD STUDY?

An RD study is meant to address the fact that, despite wanting to know the true impact of a particular treatment, program or intervention, often it is not practical or appropriate to randomly assign the population to it and then compare groups. For example, it would not be appropriate to assign one large group of children to have access to libraries and another group not - just to see whether libraries “work” in improving children's learning and development. An RD study enables this issue to be directly addressed through statistical means. Specifically, an RD approach can be used when there is a clear external means of distinguishing between two groups in such a way that the only difference between these groups is that some get a “treatment” and some do not. In other words, the two groups are treated as if they were randomly assigned and that the individuals within them are “equal in expectation” (i.e., they are only different because some are assigned to the treatment and some are not).

For this to occur, first, there has to be a treatment to which individuals are selected (e.g., prekindergarten). Second, the selection criteria have to be externally created values on a numeric rating (e.g., a test score cut-off, an age limit or requirement, etc.). For example, a school's average achievement score on a statewide exam might determine assignment to a treatment group. Schools scoring below a certain threshold would be selected to be in the treatment group, and schools scoring above the threshold would be the comparison group (or vice versa).

The reason a numeric external rating is helpful is that it can be assumed that though the cut-off point is firm, the people or schools close to it may not be all that different. For example, the cut-off for an enrichment program for students was a score of 90 on a test. Is the student who scored an 89 really that different from one scoring a 90? How about a child scoring a 91? Or as we assume in an RD design, could these students really very similar such that on a different day the student originally scoring the 89 might earn a 91 and vice-versa? By properly controlling for the value of the rating variable (which, in this case, is the average achievement score) in the RD design or the regression equation, any unobserved differences between the treatment and comparison group can be accounted for. In other words, it can be assumed that the children vying for selection in the enrichment program and scoring very close to the cut-off are the same in all ways but their score.

This issue, that an RD design really can only assume that those close to the cut-off (i.e., closely above or below) are the same other than their selection status, is one of the limitations of the technique. However, even with this limitation, predictions can still be made regarding how well the intervention would impact all children on average, but these predictions might be considered somewhat less discerning for students scoring furthest away from the cut-off.

In essence, because the RD design relies on the use of some type of cut-off, it makes it both feasible and ethical to implement across a wide-variety of situations and allows for answers to the question, on average, across a group of people, of “what works?” Figure 2.1 illustrates how this

¹³ Readers interested in the technical description of RD are referred to Appendix B of this report, Bloom (2012), Cook (2008), Cook, Shadish, and Wong (2008), Schochet, Cook, Imbens, Lockwood, Porter, and Smith (2010), and Shadish, Cook, and Campbell (2001).

process worked for this study of prekindergarten students using birth date as the cut-off mechanism, which is consistent with best practice in prekindergarten impact studies.

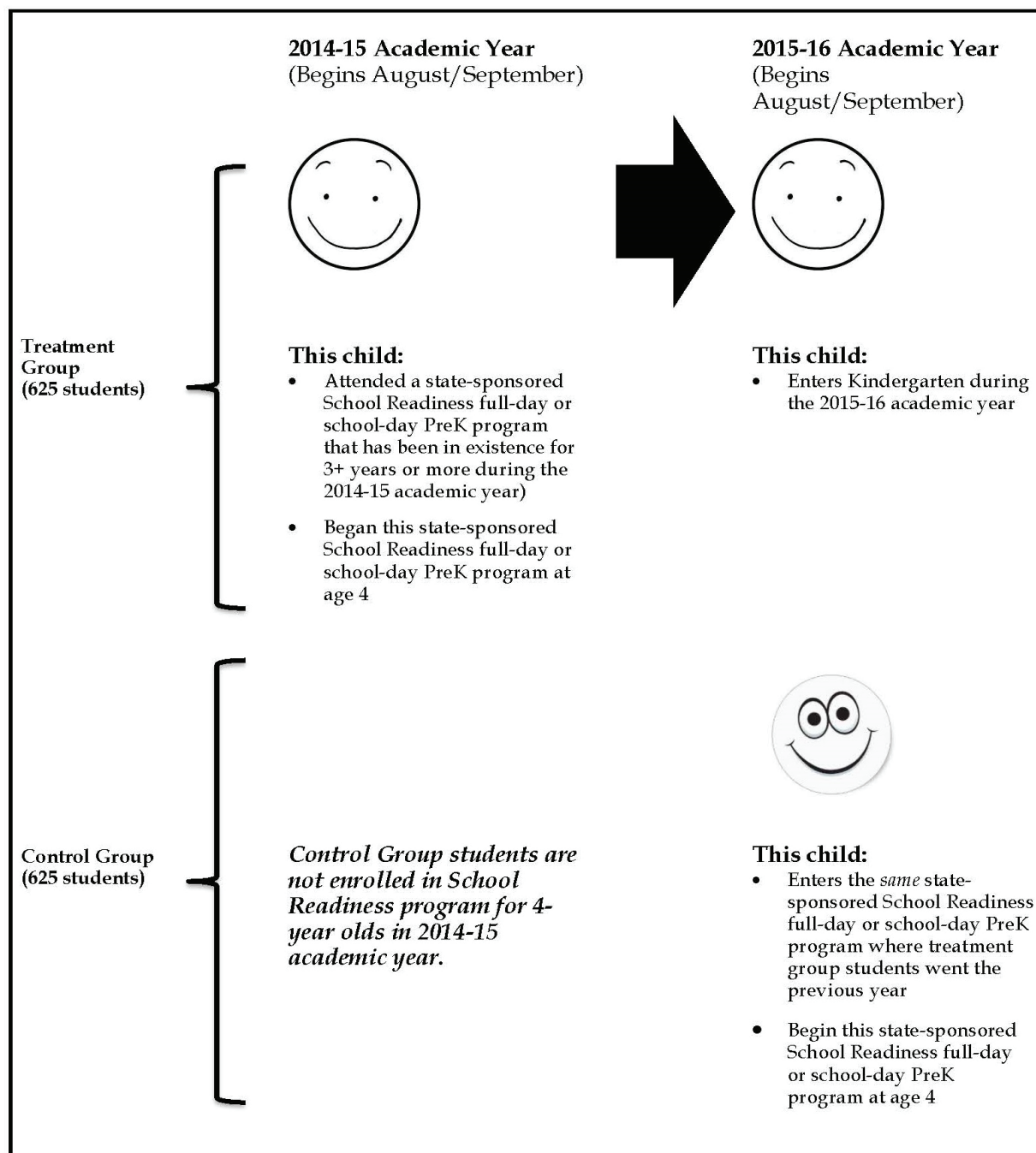


FIGURE 2.1: HOW THE RD PROCESS WORKS IN THIS STUDY

2.2.1.2 WHAT CAN WE CONCLUDE AS A RESULT OF AN EFFECTIVELY COMPLETED RD STUDY?

When an RD design study is successfully completed, it is time to interpret the results and see whether what you hoped worked did indeed work. Like any study, there are constraints on the

types of questions an RD study can answer. Turning to the present study, assuming successful completion, what can we conclude if we were to find statistically significant positive results of our treatment with a large enough effect size? Table 2.1 highlights what we can and cannot conclude for the present RD study.

TABLE 2.1: WHAT WE CAN AND CANNOT CONCLUDE FROM THE PRESENT RD STUDY

Can Conclude from RD	Cannot Conclude from RD
On average, that treatment makes a positive difference.	What about the treatment makes the difference.
This treatment causes better results, on average.	That this is the best of all possible treatments.
The tested treatment does produce positive results, on average, in the population.	That this is the most efficient (resources or cost) treatment.

2.3 STUDY METHODS

The following is a presentation of the plan proposed *at the beginning* of the study, prior to the feeder analysis and study implementation. Deviations from this proposed plan are noted in Chapter 4.

2.3.1 Study Participants

The NIEER report (Barnett et al., 2013) provided the best available initial estimate of the number of potential participants for this study. Although this report only includes students who attended prekindergarten programming during the 2012-13 academic year, a total of 5,302 were 4-year-olds and 2,517 were 3-year-olds. These numbers represent 13% of 4-year-olds and 6% of 3-year-olds (10% of the total 4- and 3-year-old population) living in the state. The CASE study was restricted to 4-year-old students enrolled in the state-funded School Readiness program and kindergarten students.

Multiple options are available for the recruitment of study sites. The two most promising options for this study are identified below, with the second being the preferred approach. The internal and external validity of the study is the key tradeoff between these two approaches—trading off the analytical rigor of the study against an ability to generalize the results of the study across a wider group of settings within the state.

2.3.1.1 OPTION 1: A FOCUS ON REPRESENTATIVE, BUT NOT ESTABLISHED PROGRAMS

One approach is to select a subset of sites that is representative of the geographic and socio-demographic makeup of the state, but is blind to program maturity. As programs across the state will be included in the sample, this sampling approach allows for the extension of inferences from the findings to make fairly robust generalizations to the population. However, a significant proportion of programs in the state are in their start-up phase. Therefore, if we were to include them in the sample, we also increase in the risk of null findings driven by random noise associated with program immaturity rather than potential program impact. For this reason, and based on study committee feedback, it was decided not to pursue this option.

2.3.1.2 OPTION 2: A FOCUS ON ESTABLISHED PROGRAMS (PREFERRED APPROACH)

The preferred approach is to sample only districts and schools that have had full-day and/or school-day prekindergarten programs in existence for three years or more. The guiding assumption, based on the work of Fixsen, Naoom, Blase, Friedman, and Wallace (2005) is that sites with established programs (three or more years) are more likely to have effectively functioning programs. Focusing on these sites will enable better identification of the effects of the program versus potential noise associated with program start-up. This approach was selected for use in this study.

2.3.1.3 UTILIZING FEEDER PATTERN ANALYSES TO INFORM SITE SAMPLING

A set of feeder analyses was conducted to identify mature sites and develop a sense of whom these sites have been serving in prior years. These analyses assessed the characteristics of students who participated in full-day and school-day, state-funded prekindergarten programs during the 2011-2013 school years. Specifically, the study seeks to understand 1) the characteristics of students who have participated in state-funded prekindergarten in this time frame, and 2) how these students have sorted subsequently into public schools offering kindergarten. These analyses enabled the choice of both kindergarten and prekindergarten sites selected for data collection (i.e., treatment and control sites) to allow for credible treatment and control comparisons.

By using administrative data from two cohorts of prekindergarten children (i.e., those that began prekindergarten in 2011-12 and those that began prekindergarten in 2012-13) and following them into kindergarten, the study was able to assess the feeder patterns between state-funded prekindergarten and kindergarten, assess attrition between public prekindergarten and kindergarten participation, and determine whether feeder patterns are stable over time. In addition, being able to observe the overall range of birthdates among children in a given site helps to better estimate how many sites and classrooms need to be sampled to achieve the desired sample size. For this study, the key assumption in conducting these analyses was that the historical data provide a reasonable indicator of what might be expected to occur during the 2014-2015 and 2015-2016 school years, the years in which students in the RD study will attend prekindergarten, and that this pattern may remain persistent. While this assumption may be highly flawed in some settings, it is expected that by identifying mature programs, these patterns may be more consistent and therefore allow for more credible inferences.

2.3.1.4 STUDENT SAMPLE

Feeder analyses were needed to provide a precise overview of the targeted study sample. However, a general outline of current thinking concerning the parameters that guided sampling procedures follows.

School programs were recruited from across the state with an emphasis on cities and municipalities with larger shares of families with lower incomes, and hence more likely to qualify for and utilize the state-funded full-day or school-day prekindergarten program. Sampling occurred at the program level with teachers and students then invited within sampled programs. The feeder analyses provided the data on where students went to prekindergarten. Treatment group participants were randomly selected from the students who attended a full-day or school-day, state-funded prekindergarten program during the 2014-15 academic year. Then, a matched sample from their prekindergarten locations was obtained, which constituted

the control group¹⁴. The aim was to include 1,250 students in the study sample to ensure sufficient statistical power to detect moderately sized effects of prekindergarten participation.

The completed feeder analyses were used to finalize the bounds of the study's age cut-off window. The preferred approach assumes the study is adequately powered to recruit students who were born within three months of the date used to determine eligibility for state-funded prekindergarten¹⁵. As an illustrative example, to achieve the desired sample size of 1,250 students equally split between the treatment and control conditions, roughly 25 students would be needed from each site at each of 25 treatment and 25 control sites (50 sites total). Sampling in this manner allows for detecting effect sizes as small as 0.25 standard deviations at a power level of at least 0.8.¹⁶ As another illustrative example, if too few students are available at each site in the desired birthdate range, the study would need to sample fewer students (e.g. 15) at each site and would then need at least 84 sites. With similar assumptions, the study would be able to detect a nearly identical effect size at an identical level of power. Based on these estimates, the larger target of 1,250 was used to randomly select students to ensure the study was adequately powered to detect an effect as small as 0.25 standard deviations at a power level of at least 0.8.

2.3.2 Measures and Procedures

2.3.2.1 CHILD OUTCOME MEASURES

Participants were assessed in the following areas: Emergent Reading Skills, Emergent Oral Expression, Receptive Vocabulary, Emergent Mathematics Skills, and Social Development. Emergent Reading Skills, Emergent Oral Expression, and Emergent Mathematics Skills were measured using subtests from the Woodcock-Johnson, Fourth Edition (WJ-IV). Receptive Vocabulary was assessed using the Peabody Picture Vocabulary Tests, Fourth Edition (PPVT-IV). Participants' social development was assessed using the Behavior Assessment Scale for Children, Third Edition (BASC-3). A description of the measures, and associated psychometric properties, are described below.

It should be noted that the administration time for each subtest from the WJ-IV described below is approximately 5 minutes. On average, the PPVT-IV test takes 10 to 15 minutes. In total, the administration of the WJ-IV and the PPVT-IV to prekindergarten and kindergarten children takes 45 to 50 minutes. The administration time for both the teacher and parent versions of the BASC-3 is 10 to 20 minutes per child.

2.3.2.1.1 Emergent Reading Skills

To assess Emergent Reading, two subtests from the WJ-IV were used including Word Attack and Letter-Word Identification. The reliability statistics for individual tests range from .90

¹⁴ As Steven Barnett (personal communication, June 18, 2014) indicated, an alternative identification approach exists. The alternative approach includes starting with those who participated in prekindergarten during the 2013-14 academic year and tracking them down in kindergarten, which would represent the treatment group. The control group would constitute a sample of new entrants during the 2014-15 academic year to the same prekindergarten programming.

¹⁵ For a fuller analysis of age cut-off windows, see Gormley, Gayer, Phillips, and Dawson (2005).

¹⁶ It is assumed sufficient covariates (4) to explain roughly 20% of the overall variation in the outcome variable and an equal split of students between treatment and control conditions. It is also assumed only 15% attrition spread equally on either side of the birthdate cut-off.

to .94 (LaForte, McGrew, & Schrank, 2014). The Letter-Word Identification subtest measures pre-reading skills including letter and word recognition and identification skills. The subtest contains 76 items and requires students to identify letters and pronounce words of increasing difficulty. The Word Attack subtest measures phonics and decoding skills. The subtest contains 32 items. Initially, participants are asked to produce letter sounds for individual letters, and then they are required to produce sounds for letter combinations. The WJ-IV is adaptive and requires the examiner to establish a baseline and a ceiling in order to limit the number of items to be administered, while still being able to estimate the participant's true score with high probability as though all items were administered.

2.3.2.1.2 Emergent Oral Expression

The Emergent Oral Expression composite includes two subtests from the WJ-IV including Picture Vocabulary and Oral Comprehension. The Picture Vocabulary subtest primarily assesses expressive vocabulary, though early items do provide some information about receptive vocabulary skills. This subtest contains 44 items and initially, participants are asked to select the correct object as named by the examiner. As items become more difficult, participants are asked to identify and name pictured objects. The individual test reliability for Picture Vocabulary is .88 (LaForte, McGrew, & Schrank, 2014). The Oral Comprehension subtest measures the ability to understand short passages that are orally presented. Participants are required to apply a missing word to the end of a sentence or related group of sentences. This subtest includes 34 items. The individual test reliabilities for Oral Comprehension is .82 (LaForte, McGrew, & Schrank, 2014).

2.3.2.1.3 Receptive Vocabulary

The PPVT-4 was selected to measure participants' receptive vocabulary. The PPVT-4 is a 228 -item test administered by having children point to one of four pictures shown when the test administrator presents a given word. The PPVT-4 age-norm and grade-norm data were designed to resemble the English- proficient population from ages 2:6-90+ and closely resembles the 2004 US Census data for demographic variables. Spearman-Brown corrected, split-half, and alpha reliabilities are reported across age, grade, and the fall and spring seasons and consistently fall above .94, indicating solid internal consistency evidence (Pearson Education, Inc., 2014; Collaborative Center for Literacy Development, 2014). Further, all items in the scale were drawn from the most recent editions of the Merriam-Webster Collegiate Dictionary and Webster's New Collegiate Dictionary.

2.3.2.1.4 Emergent Mathematics Skills

Early mathematical skills were assessed with three measures from the WJ-IV including the Applied Problems, Calculation, and Math Fluency subtests. The Applied Problems subtest was used to measure math problem solving skills. The subtest consists of 63 items and requires students to listen to math problems, identify the appropriate procedure to be followed, and then perform simple calculations to solve. Initially, the subtest assesses one-to-one correspondence and then increasingly requires participants to apply basic math procedures to solve problems. Reliability for this subtest is .96 (LaForte, McGrew, & Schrank, 2014). The Calculation subtest consists of 42 items and measures the participants' ability to complete items ranging from writing numbers to performing numerical operations. Reliability for this subtest is .97 (LaForte, McGrew, & Schrank, 2014). The Math Fluency subtest measures one's ability to solve problems

containing numerical operations quickly. Participants are given a three-minute time limit to complete a series of math facts. Reliability for this subtest was not available for the targeted age group.

2.3.2.1.5 *Social Development*

The Behavior Assessment System for Children, Second Edition (BASC-3) was administered to parents and teachers to provide a measure of adaptive and problem behaviors in the school, community, and home settings. The BASC was first published in 1992 as a multi-method, multidimensional tool that included a number of components: Teacher Rating Scales (TRS), Parent Rating Scales (PRS), Self-Report of Personality, Student Observation System, and Structured Developmental History. The test developers argue that the strength in their assessment battery is that it offers a triangulated view of behavior that is tied to the Diagnostic and Statistical Manual of Mental Disorders published by the American Psychiatric Association and the Individuals with Disabilities Education Act (Bracken, Keith, & Walker, 1994; McNamara, Hollmann, & Riegel, 1994; Sandoval & Echandia, 1994; Witt & Jones, 1998). Only TRS-Preschool and PRS-Preschool scales, both of which have been validated for use with preschool aged children up to age 5, will be used in the present evaluation. Teachers completed the assessment online. Parents or guardians were given the option to complete this assessment online or in a paper-and-pencil format. The TRS-Preschool and PRS-Preschool scales range from 100-160 items (depending on the age of the child) and describes specific behaviors on a four-point scale (i.e., from *never seen* to *almost always seen*). The TRS-Preschool and PRS-Preschool forms contain 13 scales that provide overall scores in the following four areas: Externalizing Problems, Internalizing Problems, Behavioral Symptoms Index, and Adaptive Skills. BASC-3 psychometric information is not yet available. However, psychometric information was available for the BASC-2. The BASC-2 standardization sample was matched to the 2001 US Census data. Internal consistency coefficients for the parent and teacher forms are generally near .90 for composite scales and are centered around .80 for the individual scales across both forms (McClendon, Warren, Green, Burlingam, Eggett, & McCledon, 2011). The scores on the BASC-3 are best interpreted as scale scores for decision-making purposes because of the increased reliability data. The average test-retest reliability coefficient was .80. The interrater reliability for the Teacher and Parent reports ranged from .57 to .74 for composite scores (McClendon et al., 2011).

2.3.2.1.6 *Testing Accommodations*

How best to include and assess English Language Learners (ELLs) has been intensely debated since standards-based reform took hold in the 1990s. Much of this debate has centered on issues of educational equity and issues of assessment reliability and validity. While it is beyond the scope of this study to shed light on these debates, they are relevant to the present study. In an effort to bypass anecdotal evidence, perceived notions of “best practice,” and ideological or political views on the subject, the decision concerning whether or not to assess students using translated versions of the WJ-IV and PPVT-4 is grounded in the best available evidence. At the same time, it is recognized that “judgment comes into play in determining whether a particular individual needs accommodation and the nature of such accommodation [because]... the overarching concern is the validity of the inference made from the scores on the modified test: fairness to all parties is best served by a decision about test modification that results in the most accurate measure possible of the construct of interest” (American Educational Research Association, 1999, p. 102).

Options for test accommodations (e.g., extra time, offering the assessment in the students' native language) vary widely, but all focus on "support provided [to] students for a given testing event either through modification of the test itself or through modification of the testing procedure to help students access the content in English and better demonstrate what they know" (Butler & Stevens, 1997, p. 5). All accommodations aim to provide a fair and accurate assessment of what all students know and can do by attempting to minimize measurement error for ELL students, while not disadvantaging non-ELL students (Baker, Linn, Herman, & Koretz, 2002; Shepard, Taylor, & Betebenner, 1998; Thurlow, Liu, Erickson, Spicuzza, & El Sawaf, 1996).

Further, evidence to guide decisions about when to offer accommodations is mixed (Abedi, Hoffstetter, & Lord, 2004). The validity of Home Language Surveys and English proficiency tests (e.g., Language Assessment Scales) has been called into question. For example, Abedi (2004) argued that parents might underreport the dominant language spoken in the home on Home Language Surveys for fear that their child will receive less instruction at school or held to lower standards, fear that their undocumented status may be uncovered and reported to authorities, or because of a lack of parental literacy. Extant research examining the correlations between different measures of English proficiency explain less than 5% of common variance, which is significantly lower than would be expected (Abedi, 2004; Zehler, Hopstock, Fleischman, & Greniuk, 1994). There is also little evidence of the alignment between ELL proficiency standards and English proficiency tests, which calls into question the content and construct validity of the tests (Bailey & Butler, 2003).

In examining evidence on the effectiveness of accommodations, there is little evidence that a one-size-fits-all approach (i.e., always offering accommodations or never offering them) is an appropriate position to embrace. Abedi et al. (2004) reviewed much of the empirical literature, focusing solely on "scientifically based research," and concluded that the extant literature supported four positions related to ELL accommodations:

1. Translating test items from English into other languages does not appear to be an effective accommodation strategy when students have studied the subject in a classroom where English is used. The language of assessment should match students' primary language of instruction.
2. Some accommodations are more effective with certain student groups than with others, depending on background factors such as English reading proficiency and length of time in the United States.
3. The performance gap between English learners and other students has been narrowed by modifying the language of the test items to reduce the use of low-frequency vocabulary and complex language structures that are incidental to the content knowledge being assessed. This accommodation strategy is effective; it is also valid, because it does not appear to affect the performance of English-proficient students.
4. Customized dictionaries can be an effective and valid alternative to commercial dictionaries; they have been found to help English learners while not affecting the scores of English-proficient students. (Abedi et al., pp. 17).

Based on the first recommendation, there is little evidence to suggest that kindergarten students (i.e., treatment group) included in the study should be given the option to take the Spanish version of the WJ-IV and the PPVT-4. These students will have completed prekindergarten programming, and because Connecticut is an English-only instruction state, they will have only learned content included on these assessments in English. However, this does not hold true for the prekindergarten sample (i.e., control group). In deciding whether or not to administer the WJ-IV and PPVT-4 to prekindergarten students in Spanish, assuming that is their native language, the research suggests that assessing them in their native language is the best way to gauge what they know and are able to do. It is important to note that this only holds true assuming that prekindergarten students have had no other formal early educational experiences prior to prekindergarten entry. While this information was gained through other means described below, there was no way to glean this information prior to Parent/Legal Guardian questionnaire administration. Without knowing in advance a student's prior exposure to formal education, there was no way to recruit students based on this potential information; therefore, the first recommendation cannot be employed. This represents a study constraint.

While the second recommendation dictates that decisions concerning when prekindergarten students should be assessed in English and Spanish is predicated on their length of time in the United States, it was not possible to assess this information prior to the Parent/Legal Guardian questionnaire administration. And, even if there were, as highlighted above, there are validity concerns associated with Home Language Surveys. This was also a constraint of the current study.

The third and fourth strategies cannot be adopted by this study since the WJ-IV and PPVT-4 are existing, standardized assessments, as opposed to assessments created specifically for this evaluation. This study used the specified assessments, as they are the best available measures of the outcomes of interest. Also, the current versions of the WJ-IV and PPVT-4 do not come with customized dictionaries. Were dictionaries available, they could not be used as they would provide a simplified definition of the word being assessed, which would lead to an invalid assessment.

2.3.2.2 OTHER DATA SOURCES

2.3.2.2.1 Administrative Data

State administrative data from CSDE and OEC provided information on student-level characteristics, and school-level characteristics. In addition, administrative data information provided by OEC identified the number of years programs funded through the School Readiness program have been in existence.

Furthermore, OEC and CSDE eliminated PKIS as of the summer 2014 in anticipation of replacing it with a new data collection system to be administered by OEC. Because information on prekindergarten students was paramount to the conduct of this study, the Research Team/CASE worked with OEC to develop a "PKIS-like" alternative data collection system. Fields in this dataset include School Readiness program identification (school-level variable), site accreditation status (school-level variable), and lead teacher name (class-level variable). The dataset also includes a number of student-level fields, such as student name, date of birth, address, gender, free/reduced lunch status, days enrolled, prekindergarten program funding

code, date of prekindergarten entry and exit, health insurance type, annual family income, size of family, and Individual Education Plan (IEP) status.

The PKIS-like data collected monthly from September through December 2014 will be used by the Research Team. December 2014 was chosen as the final month of data collection for the following reasons

- According to OEC, prekindergarten enrollment patterns stabilize by December, likely due to the fact the December 31 is used as a cut-off date for kindergarten enrollment eligibility
- Given that enrollment patterns stabilize, the burden on School Readiness program sites to provide this information was perceived as more important than the need for continued monthly prekindergarten enrollment data

The PKIS-like data collected for the months of September 2015 and October 2015 will be used by the Research Team for the purpose of identifying and randomly selecting prekindergarten students for data collection.

2.3.2.2.2 Student and Parent/Legal Guardian Student and Demographic Information

A basic demographic questionnaire developed by the Research Team and reviewed by the CASE Study Committee was administered online or in a paper-and-pencil format, whichever the parent preferred. This questionnaire asks for parents/legal guardians to self-report their respective demographic information, including gender, race and ethnicity, an indicator for whether their first language is English, multilingual status, disability status, as well as information on their educational background and employment history including highest degree earned, total years of education, employment status, and approximate income. The questionnaire also asks parents to provide information on how long their child has lived in the United States, and prior and supplementary early education opportunities their child has experienced. This questionnaire was available in both English and Spanish, with special consideration given to the fact that much of the Spanish-speaking population in Connecticut uses a Puerto Rican dialect.

Development of the Parent/Legal Guardian Questionnaire (see Appendix C) was based on principles for web-based survey research, including Dillman, Smyth, and Christian (2008). Briefly, these principles, as stated by Dillman et al., include:

1. The intro welcome screen should be motivational, emphasize the ease of responding, and instructs respondents about how to proceed to the next page.
2. Provide a unique PIN number to limit access and link data sources.
3. Choose a first item that is likely to be interesting to respondents, easily answered, and fully visible on the welcome screen.
4. Present each question in a conventional format, similar to that used in paper-based surveys.
5. Restrain the use of color to improve consistency, readability, navigational flow, and measurement properties of the questions.

6. Avoid problems that result from different devices, screen configurations, etc., by testing the survey from a variety of devices with varying settings.
7. Provide specific instructions on how to take each necessary computer action for responding to the question and give other necessary instructions when needed.
8. Avoid drop-down boxes to the extent possible.
9. Limit skip patterns.
10. Construct web questionnaires so respondents scroll from question to question, unless order effects are a concern.
11. Limit check-all-that-apply and open-ended questions.

2.3.2.2.3 Teacher Demographic Information

A basic demographic questionnaire developed by the Research Team and reviewed by the CASE Study Committee was administered to teachers. Teachers completed this questionnaire online. This questionnaire requested teachers to self-report their educational background, specifically which colleges and universities they attended for all degrees earned, as well as the types of in-service professional development opportunities they participated in during the past 12 months. The teacher questionnaire also was based on the principles for web-based survey research (see Student and Parent/Legal Guardian Demographic Questionnaire, section 2.3.2.2.2)

2.3.2.3 LINKING INFORMATION ACROSS DATA SOURCES

Upon entry into the public education system, including interdistrict magnet and charter schools, students are assigned a unique, 10-digit student identifier. Teachers were also assigned a unique, 8-digit identifier that follows them as long as they are employed by a public education agency within the state. Since the Research Team for this study had access to both unique identifiers prior to data collection, they generated a unique, 9-digit parent identifier by and paired with their child's unique 10-digit identifier. While the CSDE does not currently have a mechanism for linking students to their teacher of record, each public school maintains class enrollment information. As part of this evaluation effort, schools were asked to provide the evaluation team with a copy of their official 20th-day of school class enrollments for the 2014-15 and 2015-16 academic years. These course rosters identified the teacher of record associated with each unique 10-digit student identifier, and it is this information that allowed data to be linked across data sources.

2.3.2.4 PREDICTORS

2.3.2.4.1 Forcing Variable

The ability to identify the causal effect of being just eligible for state-funded full-day or school-day prekindergarten in a mature program setting on subsequent outcomes depends on the use of an exogenously chosen (i.e., externally determined) date-of-birth cut-off as the means for determining eligibility. Critical to the ability to identify a causal effect of prekindergarten is the assumption that parents and students cannot manipulate their position relative to this cut-off (i.e., that their birthday is fixed as is the cut-off). A student's position relative to this cut-off definitively identifies them as eligible for the prekindergarten program or not.

Administrative data provided by both the state and the schools was used to identify a student's distance from the cut-off. This relative distance centered on the cut-off date (*CBIRTH*) constitutes the forcing variable, meaning how far their birthday is from the one defined as the cut-off (January 1, 2015). This cut-off is based on Connecticut law (C.G.S.* Sec. 10-15c), which states that a child who is 4 years old on or before the first of January of any year is eligible to attend prekindergarten. Being born at the cut-off for eligibility translates to a value of zero. Being born before the cut-off is associated with negative values and being born on or after the cut-off is associated with zero and positive values. As a result, this study is measuring the difference in current knowledge between those who were in prekindergarten last year (based on the birthdate eligibility) relative to those who were not eligible last year but who enrolled in 4-year-old prekindergarten this year.

While students in the control group may self-select into other prekindergarten settings (including the possibility of other state- or federally funded programs, e.g., partial day, Head Start, etc.), their relative age makes them ineligible for the state-funded program, and likewise means that these students will begin kindergarten a year later than their peers who just made the cut-off. Thus, the study is estimating the effect of just making the cut-off and experiencing full-day or school-day, state-funded prekindergarten compared to just missing the cut-off and having a different, but not clearly defined, set of experiences.

2.3.2.4.2 Treatment Indicator

A dichotomous indicator has been defined for eligibility for full-day or school-day, state-funded prekindergarten. This variable (*ELIG*) is equal to one for non-negative values of the forcing variable (i.e., having birthdate at or before the cut-off) and equal to zero otherwise (i.e., having a birthdate after the cut-off). As a result, the study treats all those who qualify for full-day or school-day, state-funded prekindergarten as treated, and all those who do not qualify for this specific program as in the control group.

2.3.2.4.3 Student-level Covariates

A set of student-level covariates available to the Research Team was added through the state-level administrative data to improve the precision of the estimates. These covariates include indicators for gender, self-identified race and ethnicity, English-language learning status, disability status, primary language, and free- or reduced-price lunch eligibility. To complement the data available in the state administrative data, information was collected from students' parents, specifically information regarding prior and supplementary early education experiences in which their child participated.

2.3.2.4.4 Teacher-level Covariates

A set of teacher-level covariates was included to account for possible differences in the experiences of both treated and control students that may be attributed to the characteristics of their specific teacher. These covariates are available from the state's Certified Staff files and include indicators for teacher race, sex, highest degree earned, years of experience, and whether a teacher is a long-term substitute, as well as a continuous measure of age.

2.3.2.4.5 Parent-level Covariates

Specific information from a student's parent/legal guardian was collected to complement the available state administrative data. Specifically, data from a student's parent/legal guardian was collected regarding their respective demographic and educational background, including: highest degree earned, total years of education, an indicator for whether their first language is English, employment status, and approximate income.

2.3.3 Proposed Analyses

2.3.3.1 GENERAL OVERVIEW OF RD METHODOLOGY

The RD design was first pioneered by Thistlethwaite and Campbell (1960), but has only been used on a large scale in applied social science over the last 15 years. This analytic approach relies on a scenario where subjects are effectively assigned to treatment or control on the basis of a value on an available covariate. The subject's position relative to a predetermined cut-off value of that covariate was then sorted into treatment or control solely based on the position of their covariate value relative to the cut-off. This methodology relies upon the assumption that neither the person doing the assignment, nor the person being assigned, can manipulate their position relative to the covariate cut-off, making the assignment – though based on a covariate – effectively random. Provided that a small set of assumptions are met (in addition to those listed just above, see Schochet, 2008), the average outcomes of the *de facto* treatment and control groups can then be compared to estimate the causal effect of treatment on a subject's outcomes for those subjects close to the cut-off used to assign treatment.

2.3.3.2 SPECIFIC OVERVIEW OF RD ANALYSIS FOR CONNECTICUT PREKINDERGARTEN EVALUATION

Comparing the average outcomes of students who are just eligible for full-day or school-day, state-funded prekindergarten to those who are just ineligible allows for an estimate of the causal effect of eligibility for this program. These effects will be unbiased and reliable provided that it can be established that the data are consistent with two fundamental identifying assumptions. First, it must not be possible for students or their families to manipulate the student's position relative to the cut-off date used to define eligibility. The timing of the introduction of this program relative to the age of the students it serves, as well as the difficulty of timing exactly a child's date of birth, makes this first assumption reasonable. The second assumption is that students on either side of the discontinuity are similar on both observed and unobserved characteristics. The plan is to establish, using available administrative data, that the students on either side of the cut-off are similar on observable characteristics and therefore plausibly similar on unobservable characteristics. In addition, it will be established that no other discontinuities exist in exposure to full-day or school-day, state-funded prekindergarten.

The following statistical model is *proposed* to estimate the causal impact of eligibility on student outcomes:

$$Y_{ijk} = \beta_0 + \beta_1 \text{ELIG}_{ijk} + \beta_2 \text{CBIRTH}_{ijk} + \beta_3 \text{ELIG} * \text{CBIRTH}_{ijk} + X_i' \gamma + \pi_k + \epsilon_{ijk}$$

In this specification, Y_{ijk} represents a generic outcome for student i , in school j , and district k . The variable $ELIG$ represents the dichotomous indicator of eligibility, $CBIRTH$ is a student's position relative to the cut-off date for eligibility, and $ELIG*CBIRTH$ is an interaction term that allows the relationship between the outcome and relative age to differ for those who are and are not eligible. In some specifications the vector X of student covariates is included to improve the precision of our estimates, π_k is a vector of school fixed effects to restrict all comparisons to within site, and ϵ_{ijk} is a mean zero student-level error term, clustered on values of the discrete forcing variable (Card & Lee, 2008).

2.3.3.3 SENSITIVITY ANALYSES

2.3.3.3.1 *Linearity Statements*

In the specification provided, the assumption made is that the relationship between student-level outcomes and their relative age is linear. While this assumption that the relationship is locally linear is arguably appropriate (Lee & Lemieux, 2010), the plan is to test the sensitivity of this study's findings to using fully flexible higher-order terms in polynomial specifications, or by exploring polynomial splines of this relationship.

2.3.3.3.2 *Bandwidth Statements*

In prior evaluations of prekindergarten that rely on a date-of-birth eligibility cut-off to identify the causal effect of an analytic window, or bandwidth, six months on either side of the cut-off was used to determine which students should be included in the analysis. In this study, the primary specification adopts the same bandwidth but, as is consistent within the literature that relies on RD designs, the sensitivity of this study's findings was tested using multiple choices of bandwidth including an optimal bandwidth chosen according to the suggestion of Imbens and Kalyanaraman (2013) (see also, Lee & Lemieux, 2010).

2.3.3.3.3 *Concerns about Selection*

Students who are eligible for full-day or school-day, state-funded prekindergarten and who chose to participate may differ on important observable and unobservable characteristics from those students who are eligible and choose not to participate. To account for some of this potential selection problem, the primary RD design was complemented with a modified RD design that uses the observable student- and parent-level data described above to generate a student's propensity to participate or not participate in full-day or school-day, state-funded prekindergarten conditional on eligibility. Similar to the approach used by Weiland & Yoshikawa (2013), the primary model was then re-estimated using weights consisting of the inverse of a student's estimated propensity.

2.3.4 *Study Procedures and Timeline*

An overview of the major evaluation milestones is included in Table 2.2.

EARLY CHILDHOOD REGRESSION DISCONTINUITY STUDY
INTRODUCTION AND STUDY METHODS

TABLE 2.2: TIMELINE OF PROPOSED STUDY MILESTONES

Month	Major Milestones
2014	
May	<ul style="list-style-type: none"> ✓ CASE under contract with the CGA to conduct the RD Study ✓ Research Team (UConn) selected, study advisors identified, and study committee established ✓ Initial meetings with OEC and CSDE to provide a study overview and review study tasks involving each agency, including negotiation of CSDE MOU for access to administrative data
June	<ul style="list-style-type: none"> ✓ CSDE/UConn MOU for access to administrative data executed ✓ First transfer of administrative data to the Research Team ✓ Letter to principals, directors, superintendents, and priority and competitive School Readiness Liaisons sent introducing the study, including participation requirements, and a joint letter from the OEC and CSDE Commissioners requesting cooperation ✓ First study committee meeting with presentations by NIEER on RD studies and the Research Team on the proposed methodological approach for the study ✓ Scope of Work Task 1: Research Plan section of study report completed. The plan identifies RD methodology, summarizes results of representative RD studies, identifies proposed methods/tools, identifies planned number of subjects and study sites, and identifies plan for selection of participants
July	<ul style="list-style-type: none"> ✓ Research Team submitted application to UConn's Institutional Review Board (IRB) ✓ UConn IRB approval procured for feeder analyses ✓ Feeder analyses initiated ✓ Met with priority School Readiness Liaisons to provide study overview and Q&A ✓ Met with CGA to discuss study schedule and consideration of additional policy questions. Data collection moved from fall 2014 to fall 2015 ✓ IRB application approved (See Appendix D)
August	<ul style="list-style-type: none"> ✓ Notification received that Prekindergarten Information System (PKIS) eliminated ✓ Research Team/CASE works with OEC on a solution to obtain "PKIS-like" prekindergarten data ✓ Feeder analyses on hold pending consideration of additional policy questions ✓ Second study committee meeting held to update on progress and discuss potential additional policy questions
September	<ul style="list-style-type: none"> ✓ Met with CGA to finalize policy questions to be included in the study. CASE submitted an amendment to study scope of work incorporating additional policy question to CGA for approval ✓ OEC and the Research Team/CASE work out solution for obtaining "PKIS-like" prekindergarten data ✓ Research Team submitted amendment #1 to UConn's IRB ✓ UConn assists OEC with implementation of proposed solution ✓ Met with competitive School Readiness Liaisons to provide study overview and Q&A ✓ IRB amendment #1 approved (See Appendix D)

EARLY CHILDHOOD REGRESSION DISCONTINUITY STUDY
INTRODUCTION AND STUDY METHODS

TABLE 2.2: (CONTINUED)

2015	
February	✓ Transfer of administrative data from OEC to UConn (September 2014 through December 2014) completed
April	<ul style="list-style-type: none"> ✓ Determined additional policy questions would not be included in the study; work continued on policy questions in original scope of work ✓ Feeder analyses restarted ✓ Preparations for fall 2015 data collection initiated
May	✓ Updated letter on study progress sent to directors, superintendents, and priority and competitive School Readiness Liaisons in preparation for fall 2015 data collection, including an updated joint letter from the OEC and CSDE Commissioners requesting cooperation
June	<ul style="list-style-type: none"> ✓ Research Team submitted continuation amendment #2 to UConn's IRB ✓ Third study committee meeting: Research Team provided project update, initial feeder analyses results, and preparations for fall data collection ✓ Teacher and parent/guardian questionnaires drafted for study committee review ✓ Research team initiated process to renew MOU with CSDE for administrative data ✓ Update on fall data collection to directors, superintendents, Priority and Competitive School Readiness Liaisons ✓ Legislation adopted providing additional funding for the study
July	<ul style="list-style-type: none"> ✓ Preparation for fall data collection continues ✓ IRB continuation amendment #2 approved (See Appendix D)
August	<ul style="list-style-type: none"> ✓ Preparation for fall data collection continues ✓ Research Team submitted amendment #3 to UConn's IRB
September	<ul style="list-style-type: none"> ✓ Ordered data collection assessment materials ✓ Finalized teacher and parent/guardian questionnaires ✓ Assessor recruitment initiated and completed ✓ Scope of Work Task 2: Implementation section of study report complete. Includes solution and transfer of OEC "PKIS-like" data and transfer of CSDE data to the Research Team, feeder analyses results, sites and students for data collection confirmed, assessor recruitment and training summary, data collection schedule, and IRB approval secured – including any required amendments. ✓ IRB amendment #3 approved ✓ Established data collection schedule
October	<ul style="list-style-type: none"> ✓ Transfer of administrative data from OEC to UConn completed ✓ Data cleaning initiated ✓ Data collection begins
November	✓ Data collection ends
December	✓ Potential data collection extension month, if data collection cannot be completed by the end of November.

2.3.5 Proposed Assessor Recruitment, Training, and Data Collection Schedule

A recruitment email will be developed and distributed to public and private universities throughout Connecticut. This recruitment flyer will target junior and senior undergraduate students in education preparation programs and graduate students in Schools/Colleges of Education, as these students are required to complete assessment coursework as part of their formal university training. The recruitment email will direct those interested in serving as assessors on the project to complete an online assessor application form. Based on the assumption of each assessor conducting three assessments per day, a target of 20 assessors has been set.

Assessor training will be conducted in September 2015. Further, once treatment and control students are identified, a schedule will be established for conducting subject assessments. Google maps will be used during this process to help ensure that none of the assessors have to drive more than one hour from their home to/from a school. A number of progress monitoring systems will be put in place, including the creation of an electronic system to organize deployment (e.g., school, when, student name, assessor home address, material pickup, material drop off, assessment scoring), the creation of an organizational chart of responsibilities including whom to contact and when, the creation of an assessment material check-in/check-out process, the creation of a storage site for assessment materials. To ensure that data collection will occur and is on track, an electronic assessment implementation monitoring system will be developed and used by the Research Team.

Assessor training materials will be developed or obtained, as follows

- Training session protocols
- An assessor handbook and assessment materials for training purposes
- Student identification for assessors to provide to teachers and parent/guardians
- Observation guidelines for adults attending assessment sessions to ensure no interference with the data collection that could invalidate the result

Scheduled training sessions, totaling five hours of training

- First session: focused on filling out required paperwork, a brief overview of the purpose of the study, and the assessments
- Second session: focused on assessment process (before, during, and after), as well as confidentiality issues.

2.3.6 Evaluation Approaches Guiding Study Efforts

A number of evaluation orientations guide study process and practices. The preference for rigorous methodological work in this particular evaluation is based on the work of Campbell and colleagues (Campbell & Stanley, 1963, 1966; Cook & Campbell, 1979; Shadish, Cook, & Campbell, 2001; Thistlethwaite & Campbell, 1960). Further, the belief that the primary purpose of evaluation should be as a mechanism for providing information for decision-making efforts

is rooted in Cronbach's recognition that evaluation does and should play a role in these deliberations (Cronbach & Associates, 1980). However, like Weiss (Weiss, 1979, 1997; Weiss & Bucuvalas, 1980), it is also recognized that use of evaluative information need not be confined to immediate use. Rather, it can occur gradually, over a period of time, as the result of "the build-up of small choices, the closing of small options and the gradual narrowing of available approaches" (Weiss, 1976, p. 226). In line with these orientations, the intent is that the results of this evaluation not only inform conversations in Connecticut concerning whether or not prekindergarten makes a difference, but also serve as a first step in better understanding for whom and under what conditions it matters most. Both of these questions are of paramount importance as Connecticut expands state-funded prekindergarten programming.

2.3.7 Ethical Practices and Principles Guiding Evaluation Efforts

To ensure that evaluation procedures adhere to ethical principles and legal requirements, approval from the UConn-Storrs IRB was procured prior to conducting research involving human subjects. See Appendix D for detail on IRB approvals, including amendments as needed during the conduct of the study. Further, the Research Team ensured compliance with the terms and conditions of the Family Educational Rights and Privacy Act (FERPA) data sharing agreement between the CSDE and UConn throughout the project. Finally, the *Guiding Principles for Evaluators* (American Evaluation Association, 2004) will undergird all of the evaluation work performed in this study.

3.0 IMPLEMENTATION PLAN

The following is a discussion of study implementation as conceptualized at the beginning of the study, and as enacted between June and August 2015.

3.2 METHODS

3.2.1 Study Timeline

As planned, major timeline milestones between June and August 2015 were achieved (see Table 3.1).

TABLE 3.1: TIMELINE OF STUDY MILESTONES

Month	Major Milestones
2015	
June	<ul style="list-style-type: none">✓ Research Team submitted continuation amendment #2 to UConn's IRB✓ Third study committee meeting: Research Team provided project update, initial feeder analyses results, and preparations for fall data collection✓ Teacher and parent/guardian questionnaires drafted for study committee review✓ Research team initiated process to renew MOU with CSDE for administrative data✓ Update on fall data collection to directors, superintendents, Priority and Competitive School Readiness Liaisons✓ Legislation adopted providing additional funding for the study
July	<ul style="list-style-type: none">✓ Preparation for fall data collection continues✓ IRB continuation amendment #2 approved (See Appendix D)
August	<ul style="list-style-type: none">✓ Preparation for fall data collection continues✓ Research Team submitted amendment #3 to UConn's IRB

3.2.1 Study Participants

3.2.1.1 RANDOM SELECTION OF SITES

Random selection of sites was completed in August 2015. Superintendents and School Readiness Liaisons were notified via email if their town or district was randomly selected for inclusion in the study. If their town or district was randomly selected, they were also informed of how many students in the treatment and control groups would be assessed across the schools and centers in their town or district. For those randomly selected, this notification also stipulated what would be required of their town or district, and a FAQ document prepared by CASE and a letter from the Commissioners of Education and OEC requesting their cooperation on the study was provided.

A separate email notification was sent to principals and center directors informing them if their school or center was randomly selected for inclusion in the study. If their town or district was

randomly selected, they were also informed of how many students in their specific school or center would be assessed. For those randomly selected, this notification also stipulated what would be required of their school or center, and a FAQ document prepared by CASE and a letter from the Commissioners of Education and OEC requesting their cooperation on the study was provided.

3.2.1.2 NOTIFICATION OF SELECTED SITES

Upon notification of selection, sites were reminded that the Connecticut General Assembly requested this evaluation study because of its need for credible evidence from research to provide an indication as to the benefit to children participating state-sponsored prekindergarten. Further, per the Office of Early Childhood's *2015 School Readiness Grant Program: Application for Priority [Competitive] School Readiness District Municipalities Request for Proposal*, "...Grantees are also expected to participate as requested in all state-level evaluation activities" (p. 4). All sites were instructed to refer to their agreements and policies related participation in any state-level evaluation.

Nonetheless, many of the towns and districts have local policies that require study participants to have the choice of opting out. The approved UConn IRB also requires that a notification letter describing the study, including risks and benefits, be sent to each individual study participant that was asked to provide data. This notification letter also describes that participants have the choice to opt out, and the formal procedures for doing so.

3.2.2 Feeder Analysis Implementation for Student Selection

An observation from the literature review of representative studies that utilized an RD research methodology elsewhere in the United States (see Section 2.2) showed RD design around age cut-offs to determine prekindergarten impacts (e.g., Gormley, Gayer, Phillips, & Dawson, 2005; Gormley, Phillips, & Gayer, 2008) was standard practice.

Further, prior research had shown a relationship between selection into prekindergarten and student demographic characteristics, maternal education, and maternal support (Barnett & Carolan, 2013). Other work found that parents with less work flexibility or difficulty accessing public prekindergarten were more likely to opt out of the public prekindergarten system in favor of home-based care (Fuller, Kagan & Loeb, 2002). Based on these findings, some researchers have begun to explore enrollment patterns (e.g., Loeb, Fuller, Kagan, & Carrol, 2003; Weiland & Yoshikawa, 2013) as a means to better understand if patterns of prekindergarten enrollment introduce selection bias. However, there has yet to be work explicitly examining feeder patterns of prekindergarten enrollment to better understand this issue, and strategize for how to address it *a priori*. This is the purpose of the feeder analysis portion of the RD study.

The ultimate goal of the feeder analysis investigation is to have a sample that permits the most credible inference about the effects of *future* prekindergarten programs. This requires comparability between (a) those studied and this future prekindergarten population and (b) the treatment and control groups. There is no way to establish this comparability without a comprehensive feeder analysis.

Thus, the feeder analysis (and data) was conducted to

- Help develop a stratified random sample of programs within the state, classrooms within programs, and children within the classrooms. This is similar to other RD studies of prekindergarten (c.f., Howes, Burchinal, Pianta, Bryant, Early, Clifford, & Barbarin, 2008; Hustedt, Barnett, & Jung, 2007; Wong, Cook, Barnett, & Jung, 2008).
- Ensure that the sites and students randomly sampled are representative and can therefore be generalized for Connecticut's population. This is important for increasing the confidence that any observed effects are due to the effects of prekindergarten participation, and not biased as a function of the sample or prekindergarten enrollment patterns.
- Identify whether and how the characteristics of students or prekindergarten sites have changed as a result of prekindergarten expansion over the past couple of years. If the characteristics of students that participated in prekindergarten prior to expansion, beginning in 2012, differ from those that are current participants, it is important to know on what dimensions these differences exist to determine how they might impact not only the sampling plan, but also the overall RD findings.

For example, if an earlier cohort of students was more likely to come from areas of concentrated poverty, but later cohorts were more likely to come from areas of lower concentrations of poverty, program impacts would be expected to be smaller. This is because, on average, prior prekindergarten RD research has shown larger effects for students from areas of concentrated poverty as opposed to those that come from areas with lower concentrations of poverty. Wilson, Dickinson, and Rowe (2013) showed in their prekindergarten RD evaluation that the addition of two schools from areas of lower concentrations of poverty produced differences on key student demographic variables between the treatment and control groups, which then reduced estimates of program impacts.

To assist in determining comparability, the following set of questions guided the feeder analysis:

- Feeder Analysis Question #1: What are the characteristics of students who have participated in both public prekindergarten and kindergarten programs in Connecticut over time?
- Feeder Analysis Question #2: How much attrition between prekindergarten and kindergarten participation is observed on average?
- Feeder Analysis Question #3: How closely do prekindergarten programs adhere to the birthdate-based rule for eligibility?
- Feeder Analysis Question #4: To what extent are there systematic differences between students who can be tracked and those who cannot be tracked?
- Feeder Analysis Question #5: What observable characteristics are associated with a child repeating prekindergarten?

3.2.2.1 DATA AND SAMPLE

The dataset used for the feeder analysis consists of state administrative data that record enrollment and demographic information for children who were enrolled in state-funded prekindergarten and kindergarten programs in the academic years spanning 2011-2012 through 2013-2014. This span allows for the tracking of the enrollment and promotion patterns of two cohorts of prekindergarten participants into their potential kindergarten settings. Importantly this provides the ability to document — among those who initially enrolled in an eligible prekindergarten program — who remains in prekindergarten for a second year, who is seen enrolling in kindergarten, and who leaves the original sample. The rich set of child information, including race, ethnicity, gender, family income status, home language, English language learner status, disability status, and birthdate, provides important descriptive evidence that allows for policy-relevant analyses of the patterns of participation and supports the ability to answer the research questions. The combined analytic sample includes over 80,000 prekindergarten and kindergarten students.

3.2.2.2 MEASURES AND ANALYTIC APPROACH

The primary outcomes of interest differ by research question but all relate to describing the patterns and characteristics of student participation between prekindergarten and kindergarten.

The following approaches were used to derive answers to the feeder analysis questions:

- Feeder Analysis Question #1
The demographic characteristics of students and their families are presented as important descriptors of the population served (over three years) by the prekindergarten programs in Connecticut.
- Feeder Analysis Question #2
A binary indicator was created representing whether a child attrites (*leaves*) from the sample. This indicator is equal to “1” if the child is not present in the second year sample for their respective cohort, and equal to “0” otherwise.
- Feeder Analysis Question #3
The answer relies on a binary indicator of whether a child is eligible for state-funded prekindergarten (*ELIG*) based on the defined birthdate cut-off. First, the share of students who are enrolled in state-funded prekindergarten who were eligible based on the birthdate cut-off defined by the state is summarized and compared across cohorts. This analysis also summarizes the distribution of birthdates by day among participants. Subsequent analysis explores what observable factors are associated with the decision to enroll in state-funded prekindergarten programs when students are, and are not, eligible based on the state-defined cut-off.
- Feeder Analysis Question #4
Use of a logistic regression model to understand whether there are differences, on average, in the observable student characteristics of students who remain in the state system (either repeating prekindergarten or enrolling in kindergarten) and those who leave. Specifically, a statistical model is fit that allows understanding of the relative contribution of any one student characteristic, controlling for all others, on the odds that a student attrites. Importantly, this approach allows for fixed effects to also be included

for town of residence to reveal whether the descriptive results mask important trends, or whether there are clear differences across municipalities in the state.

- Feeder Analysis Question #5

The primary outcome for this question is an indicator of reenrolling in prekindergarten (*REPEAT*), which is defined as equal to “1” if a child enrolls in prekindergarten in two consecutive years, and equal to “0” otherwise. Again using a logistic regression framework, the log odds ratio is fit for a student enrolling in prekindergarten for a second year as a function of their observable student characteristics, with fixed effects for cohort and town of residence. Fitting unconditional and conditional models allows a determination of the relative importance of observable student characteristics in explaining the outcome, while also permitting the identification of the characteristics with the strongest explanatory power or largest marginal effect on the log odds of repeated enrollment in prekindergarten.

3.2.3 Feeder Analysis Results

3.2.3.1 DESCRIPTIVE RESULTS

Table 3.2 that presents the patterns of enrollment in the state-funded School Readiness program during the 2011-2013 academic years provides an overview of the descriptive results, which help answer the first three feeder analysis questions, as follows:

- Feeder Analysis Question #1

Results in Table 3.2 indicate that prekindergarten programs serve mostly students of color and students from lower-income families. This is consistent with it being

- a. a means-tested program (i.e., delivered to those most in need as opposed to universal access); and
- b. the demographics of the general school-aged elementary school population in Connecticut.

- Feeder Analysis Question #2

Results indicate, on average, that 10% of the prekindergarten sample does not show up in the public kindergarten dataset (Table 3.2). It is hypothesized that this attrition is due to factors such as out-of-state relocation or enrollment in private kindergarten schools.

- Feeder Analysis Question #3

Results indicate that across the 2011-12 and the 2012-13 academic years, approximately 72% of those who advanced to kindergarten from prekindergarten adhered to the birthdate-based rule for eligibility (Table 3.2). This is the group of students of interest because they represent the treatment students. While this percentage is not as high as preferred, the finding is not viewed as a true break (or disruption to) the RD methodology that cannot be accounted for through the sampling and analysis process.

Taken together, the feeder analysis provides a clearer picture of current enrollment of Prekindergarten in the state, as well as how children may or may not progress through the system. Again, this information provides a strong first step to help ensure that the sampling strategy will result in a truly representative sample, and not be biased as a result of potential prekindergarten enrollment patterns changes.

TABLE 3.2: DESCRIPTIVE CHARACTERISTICS FOR TWO COHORTS OF
PREKINDERGARTEN PARTICIPANTS FROM FEEDER ANALYSES

	Advancers ¹		Repeaters ²		Leavers ³		Kindergarten Only ⁴	
	2012	2013	2012	2013	2012	2013	2012	2013
Gender								
Male	52.5%	53.3%	54.5%	52.5%	53.0%	52.0%	51.0%	51.4%
Female	47.5%	46.7%	45.5%	47.5%	47.0%	48.0%	49.0%	48.6%
Race/Ethnicity								
White	40.2%	38.2%	35.4%	36.0%	30.7%	28.9%	62.1%	63.1%
African American/ Black	20.4%	21.4%	25.1%	24.3%	23.5%	23.8%	8.6%	8.0%
Asian	3.7%	4.2%	3.7%	3.6%	5.1%	6.0%	6.4%	5.7%
Hispanic/Latino	32.8%	33.4%	33.8%	33.8%	38.5%	39.5%	19.2%	19.3%
Other	2.9%	2.8%	2.0%	2.3%	2.2%	1.8%	3.7%	3.9%
Lunch Eligibility								
Free/Reduced	53.3%	56.6%	52.3%	55.4%	61.7%	62.0%	28.8%	26.9%
Birthdate Eligibility								
Adhered	72.2%	72.7%	2.2%	3.8%	56.9%	48.3%	23.5%	23.2%
Age								
Mean	4.22	4.22	3.25	3.29	4.03	3.87	5.31	5.31
SD	0.33	0.32	0.43	0.43	0.79	0.68	0.38	0.38
Sample								
N	15,332	15,523	9,517	9,743	3,077	3,267	24,122	22,835
<p><i>Note:</i> These results are representative of all students that attend prekindergarten and kindergarten in Connecticut, not just those enrolled in School Readiness programs. 1. Students in this group advanced from 4-year old Prekindergarten to Kindergarten. 2. Students in this group repeated Prekindergarten, because they began as 3-year olds, and as such, still show up in Prekindergarten as 4-year olds. 3. Students in this group were registered in Prekindergarten, but do not appear in the public school records in Kindergarten. 4. Students in this group did not appear in Prekindergarten, but rather, first appeared in Kindergarten.</p>								

3.2.3.2 LOGISTIC REGRESSION RESULTS

Results relevant to the remaining feeder analysis questions of interest are presented in Table 3.3, as follows:

- **Feeder Analysis Question #4**
Results indicate Asian students were more likely to leave. At the same time, Hispanic/Latino students, students eligible for free or reduced, and female students were less likely to leave (Table 3.3, Leave Column).
- **Feeder Analysis Question #5**
Results in Table 3.3 (Repeat Column) indicate, as expected, repetition of prekindergarten is largely driven by children who were younger than four years of age at their time of enrollment, as these children were not yet eligible to enroll in kindergarten. Further, while females are slightly underrepresented in prekindergarten programs overall, they are less likely to repeat prekindergarten.

TABLE 3.3: LOGISTIC REGRESSION COEFFICIENT ESTIMATES OF THE ODDS OF LEAVING THE SAMPLE OR REPEATING PREKINDERGARTEN FROM FEEDER ANALYSES (N=56,155)

	Leave		Repeat	
	(1)	(2)	(3)	(4)
Female	-0.266 (0.039)	-0.233 (0.04)	-0.222 (0.028)	-0.222 (0.028)
African American/Black	-0.033 (0.067)	-0.305 (0.073)	0.105 (0.068)	0.105 (0.068)
Asian	0.509 (0.084)	0.418 (0.087)	-0.278 (0.068)	-0.278 (0.068)
Hispanic/Latino	0.179 (0.053)	-0.092 (0.057)	0.032 (0.036)	0.032 (0.036)
Other Race	1.713 (0.039)	1.519 (0.047)	1.135 (0.034)	1.135 (0.034)
Free/Reduced Lunch	-1.103 (0.046)	-1.186 (0.048)	-0.997 (0.031)	-0.997 (0.031)
Under Age	-0.778 (0.031)	-0.809 (0.031)	3.619 (0.031)	3.619 (0.031)
Constant	-2.099 (0.039)	-3.836 (1.014)	-2.902 (0.038)	-2.976 (0.340)
Fixed Effect for Town		X		X

Notes. Standard errors are clustered at the town of residence level. Coefficients represent the marginal impact of the identified characteristics on the probability that a student experiences the outcome. Estimates of the intercept are omitted because they do not directly speak to the research questions of interest. Fixed effects for town of residence were included to assess whether the regression estimates mask important trends, or whether there are clear differences across municipalities in the state.

3.2.4 Feeder Analysis Implications

To guide sampling decisions, the Practical Sampling Design (PSD) framework (Henry, 1990) was utilized. This framework was created as a means for researchers to simultaneously attend to issues of validity, credibility, transparency, and precision. The framework makes explicit the "...series of choices that must be made, with each choice having implications for the integrity of the study. The purpose of providing the framework is to help researchers and consumers of research structure their thinking about design choices and the effects of those choices on total error." (Henry, 1990, p. 46). Table 3.4 presents the responses to the 11 questions posed by the PSD framework.

TABLE 3.4: PRACTICAL SAMPLING DESIGN FRAMEWORK PRE-SAMPLING AND
SAMPLING PHASE QUESTIONS RESPONSES (SOURCE: ADAPTED FROM HENRY, G. T. [1990].
PRACTICAL SAMPLING. NEWBURY PARK, CA: SAGE)

Pre-sampling Phase	Response
Q1. What is the nature of the study?	A1. Primarily analytical, with some secondary descriptive elements
Q2. What are the variables of greatest interest?	A2. Outcome variables include language and literacy, mathematics, and social skills
Q3. What is the target population for the study?	A3. Special population: School Readiness programs in Connecticut, specifically full-day and school-day preschool programs.
Q4. Are subpopulations/special groups important for the study?	A4. Yes, Priority and Competitive School Readiness programs
Q5. How will the data be collected?	A5. Administrative records, field work to administer standardized assessments, and developed questionnaires
Q6. Is sampling appropriate?	A6. Yes, as a means to balance study quality with time and financial constraints
Sampling Phase	Response
Q7. What listing of the target population can be used for the sampling frame?	A7. State administrative data for general feeder patterns and monthly data (Sept. – Dec. 2014 & Sept. – Oct. 2015) for specific listings
Q8. What is the tolerable error or estimated effect size?	A8. The estimated effect size is 0.25 standard deviations (assuming a power = 0.8).
Q9. What type of sampling technique will be used?	A9. Probability sampling, specifically stratified sampling to protect from the possibility of a bad sample, to insure representation of groups that have importance for the research/policy context. Proposed strata: Priority/Competitive school status, NAEYC accreditation.
Q10. Will the probability of selection be equal or unequal?	A10. Equal to ensure proportional representation in the final sample.
Q11. How many units will be selected for the sample?	A11. 1,250 total (625 in treatment, 625 in control).
Post Data Collection Phase	Response
Q12. How will nonresponse be evaluated?	A12. Response included in Chapter 4: Data Collection and Analysis
Q13. Is weighting necessary?	A13. Response included in Chapter 4: Data Collection and Analysis
Q14. What are the standard errors and related confidence intervals?	A14. Response included in Chapter 4: Data Collection and Analysis

Based on results presented in Section 3.2.3 and Table 3.4, the target population for this study is a special population, and thus, the results will only be able to generalize to students enrolled in full-day or school-day, state-funded preschool programs. Further, data collection included a mix of administrative records, fieldwork, and questionnaires.

The listing of the target population used for the sampling frame was the monthly enrollment records provided by OEC covering the months of September through December 2014. Probability sampling, specifically stratified sampling, was employed to protect from the possibility of a bad sample, to ensure representation of groups that have importance for the research/policy context. A school's/center's priority or competitive status and NAEYC accreditation were used as strata. The probability of selection will be equal to ensure proportional representation in the final sample.

Lastly, a total sample size of 1,250 students total (625 in treatment, 625 in control) allows for the detection of an effect size equal to 0.25 standard deviations (assuming power equal to 0.8).

3.3 MEASURES AND PROCEDURES

3.3.1 Linking Data Sources

Processes were developed for downloading files from online assessment websites and for linking files from the Teacher and Parent/Legal Guardian Questionnaires, standardized assessments, and OEC and CSDE administrative data.

4.0 DATA COLLECTION AND ANALYSIS

Recognizing that a variety of stakeholders (e.g., representatives from OEC and CSDE, educators, researchers, prekindergarten and kindergarten teachers, staff, and leaders, elected officials, etc.) are interested in, and will read, this report, this chapter balances providing information in easily accessible nomenclature with more nuanced methodological discussions. For this reason, some technical elements, which are important but may not be of interest to the larger readership, have been placed in appendices. This is not meant to convey diminished importance, but rather to balance the interests of multiple stakeholders.

4.1 DATA COLLECTION

4.1.1 *Study Participants*

In total, language, literacy, and mathematics assessments were collected from 529 students (206 kindergarten, 323 prekindergarten). Looking across the data, this sample produced a response rate of approximately 41.8%. In considering this rate relative to other studies analyzing the impact of prekindergarten participation on students, these rates are comparable to a number of other studies deploying similar data collection methods (Peisner-Feinberg, Schaaf, LaForett, Hildebrandt, & Sideris, 2014; Lipsey, Farran, & Hofer, 2015).

It is important to note that, looking across a number of research studies focused on the impact of prekindergarten that employed an RD study design, studies with better response rates tended to use data that were collected by the state (rather than an external entity such as an independent research team) and had mandated participation (Applied Survey Research, 2013; Gormley, Phillips, & Gayer, 2008; Peisner-Feinberg, & Schaaf, 2011). In other words, when data collection was mandated by the state as part of its general reporting requirements (i.e., mandated for all children), responses were, as to be expected, higher. However, when, as was true in this study, researchers were required to collect the data, and provide sites and individuals with the right to refuse participation (i.e., opt out), the rates dropped. For example, in the Georgia study (Peisner-Feinberg, et al., 2014), where the researchers obtained parental consent and conducted the assessments independently from the state, the team was able to achieve a 60% response rate on the consent forms, though were only able to collect data from 53% of those who signed the forms. Therefore, given the commitment of the state to allow for local decision-making regarding participation and ethical concerns regarding parents' ability to determine whether their children participate in the study, the trade-off regarding participation in this study appears to be consistent with other studies with aligned commitments and goals.

Keeping this study's approach and resultant participation rates in mind, the demographic characteristics for the students in this study's sample are included in Table 4.1.

TABLE 4.1. DESCRIPTION OF THE STUDENT SAMPLE

	Prekindergarten (2015-16)	Kindergarten (2015-16)
Number in group	323	206
Gender		
<i>Female</i>	50.31%	54.93%
<i>Male</i>	49.69%	45.07%
Ethnicity		
<i>White</i>	26.02%	36.54%
<i>African American/Black</i>	31.79%	29.33%
<i>Hispanic/Latino</i>	45.03%	42.79%
<i>Asian</i>	3.73%	3.37%
<i>Other</i>	9.09%	6.25%
Lunch		
<i>Free</i>	60.44%	51.94%
<i>Reduced</i>	4.97%	5.34%
Age when assessed (Mean/SD)	4.3 (0.54)	5.4 (0.32)
Average Standard Scores (Mean/SD)		
<i>Basic Reading</i>	91.78 (11.93)	98.24 (10.99)
<i>Broad Math</i>	84.75 (20.22)	91.49 (12.99)
<i>Picture Vocabulary</i>	96.10 (15.44)	101.15 (13.98)
<i>Oral Comprehension</i>	94.16(16.09)	96.70 (14.83)
<i>Note: SD = Standard Deviation. PPVT-4 = Peabody Picture Vocabulary Tests, Fourth Edition. WJ-IV = Woodcock-Johnson, Fourth Edition</i>		

Once the data were collected, an analysis of the characteristics of the students in kindergarten and prekindergarten who were assessed (n=529 students) was conducted and compared to the characteristics of the students in the population and to those in the original sample (i.e., those who the Research Team intended to assess). The purpose of this analysis was to: (1) establish that the students in the final assessed sample were representative of the larger population of interest, and (2) ensure that the birth date cut-off for participation in the state-funded prekindergarten School Readiness program was appropriately distributed in the final sample to allow for use of the RD design. Findings from the analysis indicate that the students from the final sample are generally representative of the population of interest (i.e., prekindergarten students who attend full-day or school-day School Readiness-funded programs). Differences that do exist are present only in prekindergarten with slight, though statistically significant, differences in the percent of White students (population has approximately 37% White students while sample assessed has approximately 26%) and students whose race was identified as “Other” (population has approximately 5% Other students while sample assessed has

approximately 9%) and a greater proportion of students receiving free lunch (approximately 60%) relative to the population (approximately 54%). However, given that the sample still represents the focus group for this study, the sample was deemed appropriate. Additionally, and important for the purposes of the RD design, there are no significant observable differences between the sample prekindergarten and kindergarten students and the distribution of birthdates in the actual sample which is smooth and continuous, and the RD design initially proposed remains appropriate for answering the research questions. A more detailed description of this analysis is included in Appendix E.

A final power analysis was run using the final sample size of 529 students (206 Kindergarten, 323 Prekindergarten). For the language, literacy, and mathematics assessments, power analysis results indicate that the study is adequately powered. Detailed power analysis information is included in Appendix F.

With regard to social skills assessments, these assessments were collected from parents for 85 students and for teachers of 95 students (i.e., one teacher assessment per student). While numbers for the social skills assessments are lower than those associated with the academic assessments, the use of the BASC-3, the social skills assessment, was exploratory. Additionally, few if any other prekindergarten RD design studies have assessed social skills with the BASC-3 or similarly time-intensive measures, a characteristic that may help to explain the low response rates. As a result of the low response rate¹⁷, the examination of social skills within an RD framework cannot be performed. A fuller breakdown of these numbers, and implications for original proposed analyses, are included in Appendix G.

4.1.2 Modifications to Data Collection from Implementation Plan

Overall, the implementation plan was followed with fidelity (Table 4.2). The process of identifying participants was not modified, and the original sample determined using information from the feeder analysis was used throughout data collection. Additionally, no changes were made in the procedures or to the tools used for data collection with students or teachers, even though there were ongoing challenges regarding responsiveness from schools and districts.

However, there were two areas in which modifications were made. The first was in the timeline for data collection. Originally data collection was scheduled to begin in October 2015 with a target completion date of the end of November, extending into December 2015 only if needed. Actual data collection began in late November 2015 and was completed in early March 2016.

¹⁷ It is important to note that this low response rate was despite the promise of a monetary stipend to both teachers and parents per completed assessment.

EARLY CHILDHOOD REGRESSION DISCONTINUITY STUDY
DATA COLLECTION AND ANALYSIS

TABLE 4.2: MODIFIED TIMELINE OF STUDY MILESTONES

Month	Major Milestones
2015	
September	<ul style="list-style-type: none"> ✓ Ordered data collection assessment materials ✓ Finalized teacher and parent/guardian questionnaires ✓ Assessor recruitment initiated and completed ✓ Scope of Work Task 2: Implementation section of study report complete. Includes solution and transfer of OEC “PKIS-like” data and transfer of CSDE data to the Research Team, feeder analyses results, sites and students for data collection confirmed, assessor recruitment and training summary, data collection schedule, and IRB approval secured – including any required amendments. ✓ IRB amendment #3 approved
October	<ul style="list-style-type: none"> ✓ Established data collection schedule
November	<ul style="list-style-type: none"> ✓ Data collection begins ✓ Transfer of administrative data from OEC to UConn completed ✓ Data cleaning initiated
December	<ul style="list-style-type: none"> ✓ Data collection continued ✓ Decision made to extend data collection thru January
2016	
January	<ul style="list-style-type: none"> ✓ Data collection continued ✓ Fourth study committee meeting: Research Team provide data collection update ✓ Decision made to extend data collection thru February
February`	<ul style="list-style-type: none"> ✓ Data collection continued ✓ Decision made to extend data collection thru 1st week of March
March	<ul style="list-style-type: none"> ✓ Data collection completed ✓ Data cleaning completed
April	<ul style="list-style-type: none"> ✓ Data analysis completed ✓ Scope of Work Task 3: Data Collection and Analysis report completed. This section of the report includes confirmation of data transfer from UConn to CSDE/OEC, actual number of sites/students and implications of any variance from proposed plan summarized, and analysis and interpretation of results ✓ Fifth study committee meeting: Research Team presented the report section – Data Collection and Analysis
May	<ul style="list-style-type: none"> ✓ Scope of Work Task 4: Final Report completed. Authorization for public release of study report secured from CASE Governing Council ✓ Sixth study committee meeting: Research Team presented the report section – Discussion and Implications
June	<ul style="list-style-type: none"> ✓ Public release of the final report ✓ Study Briefing to be conducted at the convenience of the CGA’s Education Committee (TBD) ✓ Presentations to stakeholder groups, as determined

The start of data collection was, in large part, changed due to delays in the Research Team's access to data. Moreover, once the data were acquired, it required significant effort and time to prepare the data for use. For example, there were over 200 data files to sort, clean and merge before the data could be used to identify sites and begin the process of collecting student, teacher and parent outcome data.

Once the data were available and ready to be used, the Research Team also faced a number of challenges in completing data collection that further extended the timeline. These challenges included many centers with abbreviated or modified winter schedules, holiday activities, field trips and other seasonal events, and participation in other research studies that precluded access to students for the purposes of assessment. There were also delays regarding timely responses from school-level points of contact, as discussed in greater detail with respect to overall study participation (Appendix E). The point of contact delays often included lags in the time they shared the study information forms with parents to facilitate the ability to opt out. The IRB for the study required a week-long wait time between when parents were first told of the study and student assessment.

As highlighted in the results, the timeline changes appear to have had limited, if any, impact on students' scores. To assess this impact, a method similar to that of Lipsey, Farran, Bilbrey, Hofer, and Dong (2011) was used that allowed for controlling for the date of each child's assessment, which is discussed in greater detail in the Results Section (section 4.2.1). Moreover, as few, if any, RD design studies are able to ensure that all participants are assessed at the same time, lags in assessment will always plague such studies. As such, while the extended data collection was not ideal, the need to control for time assessed is far from unique to this particular study or context.

To ensure assessor capacity throughout data collection, a second assessor training was conducted in January 2016. In total, 58 students and other professionals submitted applications and passed the required background check. Table 4.3 provides information on the assessors utilized for data collection.

TABLE 4.3: ASSESSOR INFORMATION

University Based Assessors	Number of Undergraduate Assessors	Number of Graduate Assessors
UConn	32	12
Southern Connecticut State University	0	3
Fairfield University	0	2
Springfield College	0	1
Other, Non-University Based Assessors	Number of Other Assessors	
Retired School Psychologists	2	
UConn Faculty	6	

The second modification was related to the parent/legal guardian surveys. The initial plan stated that parents would be provided with the option to complete the BASC-3 survey and a Research Team generated background information survey in either an online or paper format. This change was made because the BASC-3 publisher had an online system ready for teachers, but did not yet have it ready for use by parents/legal guardians. This procedure was modified to use only paper surveys to avoid confusion for parents/legal guardians. Therefore, the surveys were provided in paper format to parents with pre-stamped return envelopes.

4.2 DATA ANALYSIS

As noted in section 2.1.3, Purpose of the Evaluation Study, the primary evaluation questions include:

1. Do children who attend full-day or school-day, state-funded preschool programs enter kindergarten with better language and literacy skills than if they had not attended the program?
2. Do children who attend full-day or school-day, state-funded preschool programs enter kindergarten with better mathematics skills than if they had not attended the program?
3. Do children who attend full-day or school-day, state-funded preschool programs enter kindergarten with better social skills than if they had not attended the program?

4.2.1 Answer to Research Question 1

A RD design methodology was applied to the sample of students for whom data were collected. In this analysis, the required assumptions for an RD design were tested, and the estimates were subjected to specification tests to ensure that findings were not driven by model-based decisions, but rather by actual differences in the outcomes of students enrolled in the state-funded prekindergarten School Readiness program as a function of their birthdate eligibility relative to those who just missed this opportunity. Three measures of early literacy skills were used to answer this question. For early language, the Pearson Picture Vocabulary and the Woodcock-Johnson IV oral comprehension subtest were used. For early mathematics skills, five measures were used. For all outcomes, the standard score with a mean of 100 with a standard deviation of 15 was used. Standard scores are those used most regularly in other studies and reports regarding these skills. The preferred specification for all outcomes is the Imbens-Kalyanaraman (IK) optimal bandwidth with a triangular kernel, although the results are not sensitive to varying the bandwidth or using a rectangular kernel (Imbens, & Kalyanaraman, 2012).

In these analytic specifications, only one control variable is used; its purpose is to address the issue of the extended data collection window. Specifically, a linear and quadratic term representing the number of days from when student assessment began that a particular child was assessed was used as a control variable¹⁸. This approach is consistent with what was done

¹⁸ In the implementation plan, the use of other control variables (e.g., student gender, student race, etc.) was highlighted as a possible data analysis strategy to improve precision and statistical power. However, these variables were not used as results show no imbalance at the cut-off on these factors and this approach (i.e., one that does not include covariates) is consistent with the recommendations of a number of experts in the field (e.g., Schochet et al., 2010).

by Lipsey, Farran, Bilbrey, Hofer, and Dong (2011) to address the same concerns, and preserves degrees of freedom relative to other potential methods for addressing the same concern. The elongated window in which students were assessed raised the potential threat that the passage of time and the learning that occurred during this period, if differential in treatment and control groups, might drive the findings. Including these controls allowed the model to account for possible differences in the outcomes that could be driven by when a student was assessed. Using both linear and quadratic terms also ensures that estimates are not driven by potential non-linear relationships between the passage of time and student outcomes. Furthermore, given that a feeder analysis was used to identify the initial sample, the use of covariates regarding demographics was deemed as extraneous – a decision confirmed by robustness checks discussed later and highlighted in Appendix H.

Additionally, to account for nesting (i.e., students nested within schools), cluster standard errors were used, rather than fitting a multilevel model with fixed or random effects, because this strategy does not force the same parametric assumptions about the distribution of the error terms at either the school or student level, or the correlation among them. To assess the robustness of these decisions and the ability of the selected model to best fit the data, a number of other specifications were fitted including those with additional covariates, an alternate clustering approach, and fixed effects for each site. Results of these efforts are described in greater detail in Appendix H and illustrate the overall stability of the findings and the robustness of the selected approach for the outcomes. That said, it is important to note that the findings regarding Basic Reading are no longer significant with the addition of the fixed effects. However, given the size of the final sample and the resultant impact on power to detect effects, the decision to remove the school fixed effects from the initial model is warranted.

Earlier, the following statistical model to estimate the causal impact of eligibility on student outcomes was proposed:

$$Y_{ijk} = \beta_0 + \beta_1 \text{ELIG}_{ijk} + \beta_2 \text{CBIRTH}_{ijk} + \beta_3 \text{ELIG} * \text{CBIRTH}_{ijk} + X_i' \gamma + \pi_k + \epsilon_{ijk}$$

However, as already stated, given the lower than expected sample size and the minimal overlap of students in the control and treatment groups from the same School Readiness programs, the π_k term, representing a vector of school fixed effects, was omitted. Furthermore, a number of robustness checks were performed, and are described in greater detail in Appendix H.

4.2.2 Answer to Research Question 2

The same approach as with question 1 was followed, with the notable difference of focusing on six different math-related outcomes. All math outcomes are subtests of the WJ-IV as follows: broad mathematics, math calculation, applied problems, calculation, and math facts.

4.2.3 Answer to Research Question 3

As noted in section 4.1.1, research question 3 was exploratory. However, given the small numbers of parents and teachers that completed BASC-3 assessments, the social skills within an RD framework could not be examined. Appendix G includes a fuller discussion of these numbers.

4.3 RESULTS

4.3.1. Overview of Findings

Overall, the findings show evidence that attending state-funded prekindergarten in Connecticut, as delivered through the School Readiness program funding stream, is positively impacting students. An overview of results is included in Table 4.4.

TABLE 4.4. OVERVIEW OF IMPACT OF STATE-FUNDED PRESCHOOL PROGRAMS

Claim	Test Evidence	Measures	What Students Do on this Test	Skill Focus
Large, positive and statistically significant effects on a subset of student's early literacy skills (0.69 SD)	Basic reading	WJ-IV: Letter-word identification	Recognizing and naming printed letters and words	Letter/word recognition
		WJ-IV: Word attack	Reading made-up words that conform to conventional spelling rules	Phonemic awareness
Large, positive and statistically significant effects for most student's early numeracy skills (0.48 SD)	Broad math	WJ-IV: Calculations	Arithmetic computation with paper and pencil	Writing numbers to numerical operations
		WJ-IV: Math Fluency	Simple calculations for three minutes	Quickly solving numerical operations
		WJ-IV: Applied Problems	Oral, math "word problems," solved with paper and pencil	Math problem solving
Suggested positive, but non-statistically significant, effects on student's early vocabulary skills	Picture vocabulary	PPVT-IV	Listening to a word describing one of four pictures and then pointing to the picture that the word describes	Picture-to-word recognition
Suggested positive, but non-statistically significant, effects on student's early oral language skills	Oral Comprehension	WJ-IV: Picture Vocabulary	Listening to a word describing one of four pictures and then pointing to the picture that the word describes	Picture-to-word recognition
		WJ-IV: Oral Comprehension	Listening to an oral passage and identifying a missing key word that makes sense	Listening comprehension

TABLE 4.4. (CONTINUED)

Claim	Test Evidence	Measures	What Parents and Teachers Do on this Test	Skill Focus
Unknown effects for student’s early social skills	Social Development	BASC-3	Answer survey questions	Student Externalizing Problems
				Student Internalizing Problems
				Student Behavioral Symptoms Index
				Student Adaptive Skills
Note: Effect sizes are included in the second column of this table in parentheses only for outcomes that are statistically and practically significant. Woodcock-Johnson, Fourth Edition (WJ-IV). Peabody Picture Vocabulary Tests, Fourth Edition (PPVT-IV). Behavior Assessment Scale for Children, Third Edition (BASC-3)				

Looking across results, positive and significant effects on four of the five numeracy measures, and on three of the five measures of early literacy are observed. The effect sizes range from about 0.35 SD to 0.68 SD and are all in the range of what the study is powered to detect. Importantly, while there is some fluctuation in effect sizes across different specifications, the overall conclusions, general magnitude, and statistical significance of the effects are unaffected by these specification checks. Greater detail on each of the findings follows.

4.3.1.1. MATHEMATICS

In Figure 4.1, graphical evidence of the effect of birthdate eligibility for state-funded prekindergarten School Readiness program on mathematics outcomes is presented. Each of the four panels corresponds to the outcome names on the corresponding vertical axis, and in all cases the horizontal axis represents the student birthdate re-centered at the policy relevant January 1 eligibility cut-off. The cut-off is indicated with the dashed vertical line and visually apparent differences in the levels of performance of students immediately on either side of the cut-off represent the effects also presented in table form. What is clear from the visual presentation of these graphs is that the broad math panel shows clear evidence of higher level of performance for students just eligible for prekindergarten relative to those who just missed eligibility. Indeed, the math effects (Table 4.5) are significant in broad math (0.48 SD).

FIGURE 4.1. REDUCED FORM ESTIMATES OF THE EFFECT OF JUST BEING ELIGIBLE TO ENROLL IN THE STATE-FUNDED PREKINDERGARTEN SCHOOL READINESS PROGRAM ON BROAD MATH, ORAL LANGUAGE, PICTURE VOCABULARY, AND BASIC READING OUTCOMES

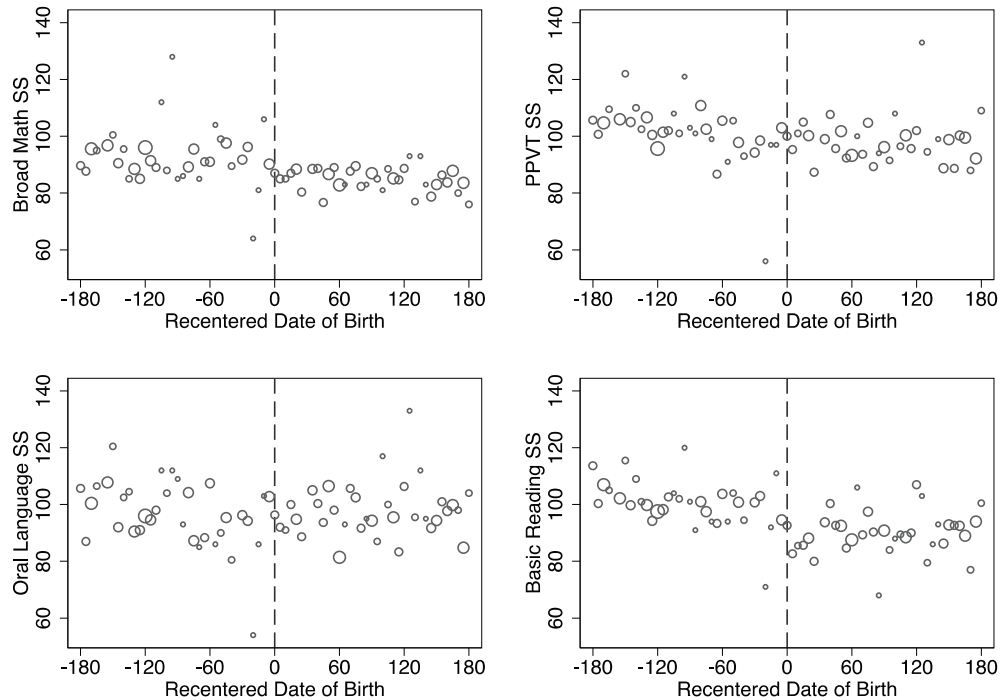


TABLE 4.5. REDUCED-FORM ESTIMATES OF THE EFFECT OF THE STATE-FUNDED PREKINDERGARTEN SCHOOL READINESS PROGRAM

	(1) Broad Math	(2) Oral Language	(3) Basic Reading	(4) PPVT
IK Bandwidth	7.198** (2.865)	3.000 (6.005)	13.028*** (4.189)	1.216 (4.817)
N	299	236	113	262
Bandwidth = 150	7.889** (3.489)	5.151 (6.924)	11.692*** (3.455)	1.993 (6.218)
N	176	176	176	173
Bandwidth = 180	7.458** (3.270)	3.442 (6.228)	10.608*** (3.176)	0.990 (5.568)
N	219	219	219	217
Bandwidth = 210	7.099** (3.083)	2.522 (5.696)	10.361*** (2.979)	1.081 (4.999)
N	255	255	255	253

Notes: Heteroskedasticity robust standard errors clustered by date of birth are in parentheses (* $p < .10$ ** $p < .05$ *** $p < .01$). The coefficients shown are generated using OLS with the stated bandwidth and a triangular kernel.

4.3.1.2. ORAL COMPREHENSION, BASIC READING, AND PICTURE VOCABULARY

The Figure 4.1 graphs, analogous to the math outcomes, present early oral comprehension, basic reading, and picture vocabulary outcomes. The figures are laid out in the same manner and visual evidence of effects should be interpreted the same way. This visual evidence suggests that the basic reading measure shows clear evidence of higher level of performance for students just eligible for prekindergarten relative to those who just missed eligibility. There is suggestive, though noisy, evidence that this may be true for oral comprehension and picture vocabulary.

Fitting statistical models further corroborates visual evidence of the effects of being eligible for prekindergarten. Impact estimates on both language and literacy outcomes are also presented in Table 4.5. In literacy, estimates on all five outcomes for the optimal bandwidth choice (as determined using the IK procedure), as well as several other choices of bandwidth, are presented. Across specifications there is clear and statistically significant evidence of impact on student's basic reading (0.68 SD).

Similar estimations were conducted for oral comprehension and picture vocabulary. Across specifications there is suggestive, but not statistically significant, evidence of impact on picture vocabulary (0.05 SD) and oral comprehension (0.31 SD) scores. The point estimates vary slightly across specifications; however, the general magnitude is consistent.

4.3.1.3. SUMMARY OF RESULTS

In the sample with preferred specification, the study is adequately powered to detect effect as small as 0.5 SD when no covariates are included, which is the specification reported and which is consistent with generally accepted practice in RD analyses. Analyses are sufficiently powered even in the preferred bandwidth to detect the effects reported for five of the eight statistically significant estimates. Using a larger bandwidth and a flexible polynomial specification of the running variable, the study is adequately powered to detect effects as small as 0.39 SD, and in this specification all impact estimates exceed this minimum detectable effect size (MDES). Further information on specification checks is included in Appendix H.

Taken as a whole, for the measures used, there is a clear positive impact on literacy and mathematics outcomes for students just eligible for the state-funded prekindergarten School Readiness program relative to students who just missed being eligible and who did not participate.

5.0 DISCUSSION AND IMPLICATIONS

5.1 SUMMARY OF FINDINGS

Study findings indicate that when children are enrolled in Connecticut's state-funded School Readiness full-day or school-day prekindergarten program there are positive impacts on a number of their kindergarten readiness skills. Specifically, the analyses conducted indicates that students enrolled in these prekindergarten programs display more proficient (1) early literacy and (2) early numeracy skills as measured by the WJ-IV test than students not enrolled in the program.

Alternatively, being enrolled in these programs does not seem to have an impact on these students' oral comprehension or picture vocabulary skills. In this section, these findings are explored in greater detail, including a discussion of how these findings compare to findings from other states in their statewide evaluations of prekindergarten programs using the same type of methodology.

However, prior to comparing these findings compare to other studies, it is useful to review the conclusions that can be made given the analytic approach utilized – that of a RD design. The positive aspect of this design is, when successfully completed, it provides opportunities to make causal claims regarding the impact of an intervention. As such, a RD design provides a powerful tool to assess the impact of interventions for which it is not logistically possible or ethical to randomly assign to the population – for instance, the ability to enroll in, or attend, state-funded School Readiness full-day or school-day prekindergarten programs.

However, as is typical with any study design, there are limitations regarding the scope of the conclusions that can be drawn when using a RD design. For this study, what can be concluded given these statistically significant positive results regarding early literacy and early numeracy? Table 5.1 highlights what can and cannot be concluded for the present RD study.

TABLE 5.1: WHAT CAN AND CANNOT BE CONCLUDED FROM THIS RD STUDY

Can Conclude from RD	Cannot Conclude from RD
<ul style="list-style-type: none"> On average, the School Readiness full-day or school-day prekindergarten programs makes a positive difference in the areas identified as statistically significant. Specifically, on average, prekindergarten students who attend School Readiness full-day or school-day programs do better, on average, in early literacy and early numeracy. 	<ul style="list-style-type: none"> <i>What about the School Readiness program makes a positive impact?</i> The RD design does not provide information about the quality of instruction, the curriculum resources, or other factors that might vary across the sample. <i>What is the best of all possible School Readiness programs?</i> This study cannot tell us whether another model of delivery might be better, nor can it tell us about comparative differences in delivery between full-day, school-day, extended-day, and half-day programming. <i>Which aspects of the School Readiness program generated the most valuable outcomes?</i> This study cannot tell us about the cost-benefit associated with different funding configurations.

As shown in Table 5.1, there are a number of questions of great interest that are beyond the scope of the present study and suggest a need for future research to explore these issues in more detail. However, there are a number of meaningful conclusions that can be drawn from these findings that are discussed in greater detail in the rest of this chapter.

For now; however, it is important to note that the findings indicate that being enrolled in the School Readiness full-day or school-day program produces positive results in early literacy and early numeracy skills, on average.

5.2 STUDY LIMITATIONS

Similar to other studies, there are a number of limitations associated with this study. Two key limitations of this study and efforts to address them are as follows.

1. **Representativeness:** Though a number of efforts were taken to gather a highly representative sample (e.g., the feeder analysis, updates from CASE throughout the planning phase and up to the implementation phase of the study provided to state and district personnel, meetings with School Readiness liaisons, the donation of additional administrative support from UConn to help with communication and assessment coordination, etc.), given the lower than anticipated participation rates (40.7%) questions regarding the true representativeness of the programs and children in the sample are reasonable. That said, it is important to note that the participation rate in this study reflects that of other studies in which an external entity, such as an independent research team rather than a state were the primary data collectors (Peisner-Feinberg, et al., 2014; Lipsey, et al., 2015). Additionally, though there were some statistical differences between the final sample and the population, these differences

avored the targeted group (e.g., slightly more students of color and students living in poverty were randomly sampled). However, even with these points, a larger sample would have been preferable, and recommendations for how this might be accomplished are provided in the Commissioning Future Statewide Prekindergarten Studies subsection (5.4.2) of this report.

2. The Data Collection Window: For a number of reasons beyond the control of the Research Team, the data collection for this study was prolonged. This extended timeframe suggests that it may be the case that the differences found among students may be attributable to their experience in kindergarten, rather than to skills with which they entered these programs as a result of prekindergarten enrollment. In other words, since the Research Team was not able to collect data from students directly upon enrollment in kindergarten or prekindergarten, students were exposed to new learning that might have equally or even more heavily impacted their performance on the assessments than prior year learning. As noted in section 4.2.1, steps were taken to address this issue statistically. Specifically, and consistent with the work of other researchers (e.g., Lipsey, Farran, Bilbrey, Hofer, and Dong, 2011), a linear and quadratic term representing the number of days from when student assessment began to when a particular child was assessed was used as a control variable. Results indicated that there was no statistical effect of time of testing on the results, and suggest that despite assessment going past students' initial entry into school, prekindergarten attendance still played a positive role.

5.3 HOW IMPORTANT ARE CONNECTICUT'S EFFECTS?

Standard practice for researchers, policymakers, educators, program staff, and other key stakeholder groups has been to use Cohen's (1988) benchmarks to draw inferences about whether the size of an effect is substantively important. This study follows that of methodological innovators (Cooper, Hedges, & Valentine, 2009; Hill, Bloom, Black, & Lipsey, 2008) who argue that other benchmarks should be used to allow for more appropriate inferences to be drawn. For example, a framework provided by Bloom, et al. (2008) suggest that more appropriate inferences can be drawn using one of several strategies:

- a. Create benchmarks related to average student growth over the course of a year;
- b. Create benchmarks that contextualize student group membership achievement (or other) gaps;
- c. Create benchmarks that contextualize prior interventions against the current intervention; or
- d. Create benchmarks related to cost-benefit or cost-efficiency analysis (p.6-7)

Also see Hill, Bloom, Black, & Lipsey (2008) or Lipsey, Puzio, Yun, Herbert, Steinka-Fry, Cole, & Busick (2012) for further discussion of the four approaches of comparisons. Given the scope and focus of this particular study, the benchmarks considered are those relating to approach C. This approach was used in three different ways.

First, to identify which effects are most notable, effect size benchmarks calculated by Hattie (2009) were used. According to Hattie (2009), and based on synthesis of over 800 meta-analyses, “the effect size of 0.40 sets a level where the effects of innovation enhance achievement in such a way that we can notice real-world differences, and this should be a benchmark of such real-world change” (p. 17). Furthermore, according to his synthesis of findings, the average effect of prekindergarten, across multiple contexts and assessments is 0.45. As illustrated in Figure 5.1, both of these benchmarks further support the claims that large, positive, and statistically significant effects on student’s early literacy and numeracy skills were detected and are noteworthy, with early literacy and numeracy skills effect sizes both above these benchmarks.

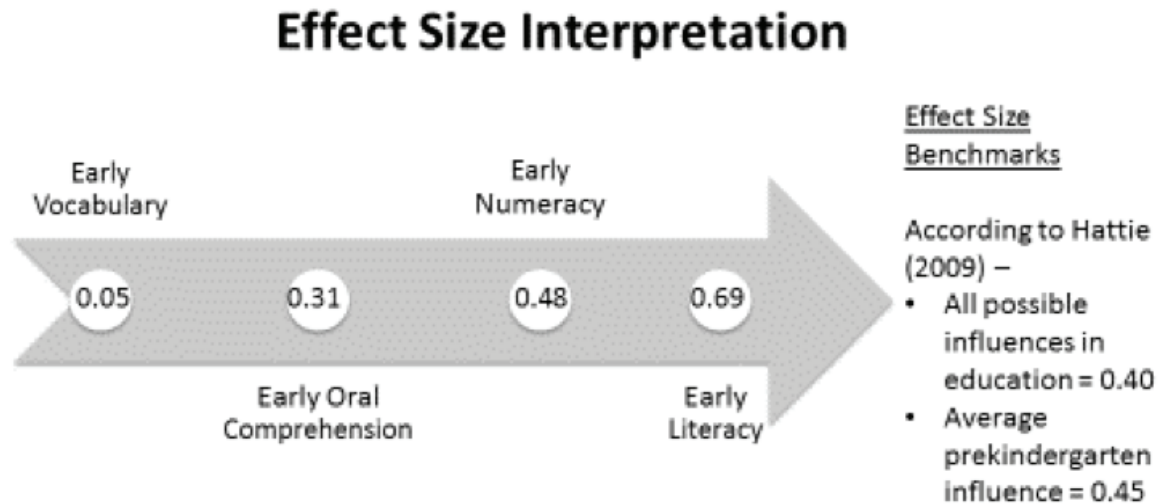


FIGURE 5.1: INTERPRETING CONNECTICUT EFFECTS IN RELATION TO BENCHMARKS

Second, effect sizes were descriptively contextualized in relation to those found in other statewide prekindergarten evaluations (Figure 5.2). Importantly, only other statewide studies that used the same assessments were used. In examining results from Figure 5.2, it should be noted that this study was the first to use composite outcome measures. This is a strength of this study because it represents an outcome that assesses a wider content area than has been assessed in prior prekindergarten, statewide impact studies that used an RD design. At the same time, this creates a situation where one-to-one effect size comparisons are unavailable, as other states only looked at one sub-test (e.g., applied problems). Effect sizes included in Figure 5.2 for early vocabulary are a one-to-one match. However, effect sizes from other states for early numeracy are those that came only from the applied problems sub-test, as opposed to the Broad Math outcome used in this study. Similarly, as opposed to the Basic Reading outcome used in this study, effect sizes from other states for early literacy are those that came only from the letter-word sub-test. Finally, effect sizes from other states for early oral comprehension are those that came only from the oral comprehension sub-test, not from the composite oral comprehension outcome, which is what was used in this study.

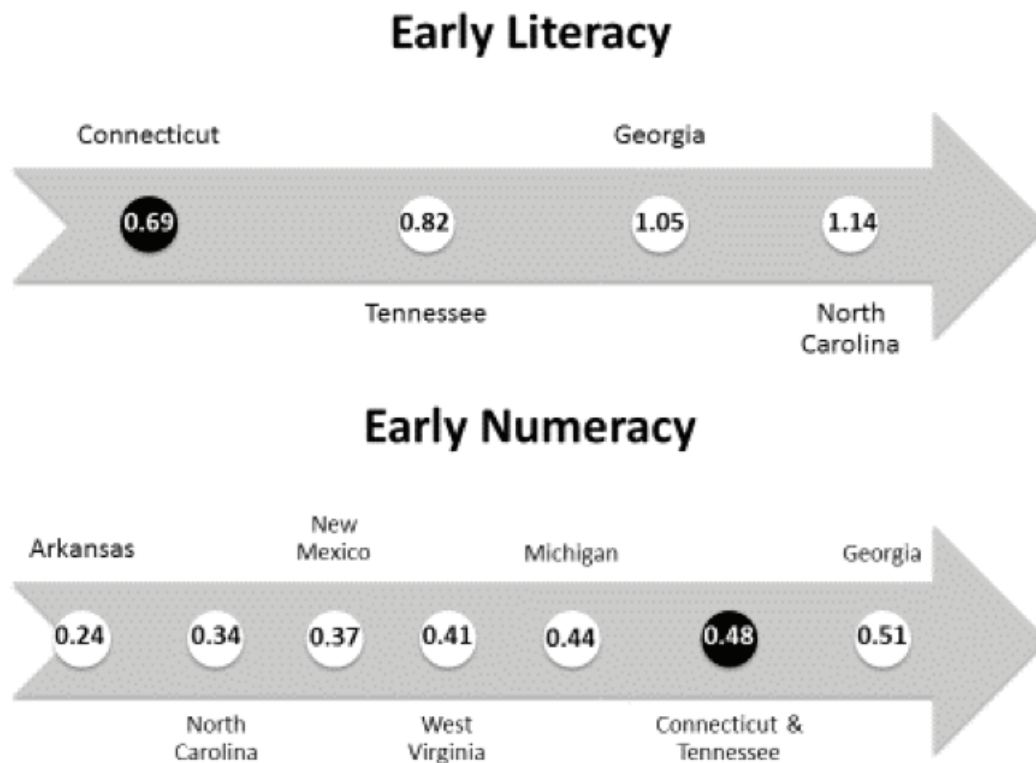


FIGURE 5.2: INTERPRETING CONNECTICUT EFFECTS IN RELATION TO PREKINDERGARTEN EFFECTS FOUND IN OTHER STATES

Lastly, effect sizes reported in Figures 5.1 and 5.2 can be descriptively compared to other prior research studies. For example, effect sizes reported for other state-funded prekindergarten programs range from .23–.53 (Gilliam & Zigler, 2001), and prekindergarten programs generally from .10 to .13 (Magnuson, Ruhm, & Waldfogel, 2004). Those reported for high-quality childcare programs seldom exceed .10 (NICHD Early Child Care Research Network & Duncan, 2003; Peisner-Feinberg et al., 2011). The Abecedarian project, widely acknowledged as a highly successful early intervention program, reported effect sizes of .73 and .79 for children ages 4 and 5 years old (Ramey, Campbell, Burchinal, Skinner, Gardner, & Ramey, 2000), and the highly praised Perry Preschool program reported effect sizes of .60 (Ramey, Bryant, & Suarez, 1985).

5.4 RECOMMENDATIONS FOR FUTURE STATEWIDE PREKINDERGARTEN STUDIES

5.4.1 Future Evaluation Questions

The positive findings of this study related to the relationship between being enrolled in the state-funded School Readiness prekindergarten program and students' early literacy and early numeracy skills bolsters prior research in this area (e.g., Hustedt, Barnett, Jung, & Thomas, 2007; Lipsey, Farran, Bilbrey, Hofer, & Dong, 2011; Peisner-Feinberg, & Schaaf, 2011; Peisner-

Feinberg, Schaaf, LaForett, Hildebrandt, Sideris, 2014), making an even stronger case that such prekindergarten programs can and do positively impact student learning. These findings also suggest the need for further studies regarding some of the mechanisms that helped to produce these results, as well as the non-findings (e.g., Why is there little difference in the oral language skills between children enrolled in the state-funded School Readiness prekindergarten program and kindergarten programs?). Table 5.2 includes recommendations for further exploration regarding these findings that may be of value.

TABLE 5.2: FUTURE EVALUATION QUESTIONS FOR CONNECTICUT BASED ON STUDY FINDINGS

Question Category	Potential Evaluation Questions
What works?	<ul style="list-style-type: none"> • Do replication studies support impact study findings across different cohorts of students? • Do longitudinal replication studies support impact study findings long-term? • Do children who attend full-day or school-day, state-funded preschool programs enter kindergarten with better social skills than if they had not attended the program?
What works for whom?	<ul style="list-style-type: none"> • Do results vary by state-funded preschool program type? • Do results vary by student characteristics (e.g., gender, race/ethnicity, income)? • Do results vary by student skill level (e.g., English proficiency)?
What works, for whom, and under what conditions?	<ul style="list-style-type: none"> • Do results vary by program quality? • Do results vary by the amount of school/system instructional support?
Which aspects are valuable?	<ul style="list-style-type: none"> • What is the relationship between program costs and outcomes observed? • Which aspects of the School Readiness program generated the most valuable outcomes?

While the present study provided insights regarding students' reading, math, oral comprehension and vocabulary skills, there are a number of other outcomes that might be useful to explore. In particular, the researchers intended to focus on student social skills, but were unable to conduct these analyses due to small sample sizes. Therefore, it is suggested that research be pursued to better understand the varied impact of prekindergarten on students for a variety of outcomes, including non-cognitive skills (e.g. social skills). Additionally, it would be useful to know whether and to what degree such results persist over time as well as if different cohorts of students have different results. This information could lead to further research regarding whether other external factors such as those related to the economy, the degree of teacher turnover or school culture may more broadly impact students' success.

Another area that warrants further exploration is related to variability in either the type of programs being offered or the groups of students enrolled in them. For example, are there certain groups of students by demographic features (race/ethnicity, socio-economic status, gender, etc.) or particular skills levels (e.g., English proficiency, etc.) who benefit more or

less from these programs? Alternatively, how might extending the school-day or the type of program (e.g., School Readiness versus non-school readiness) differently impact student success?

Perhaps of most significant interest, is research that provides a better understanding of the specific site-based mechanisms that helped to produce the outcomes from this study. For example, research that examines the relative content and quality of instruction at each prekindergarten site would help to better understand the average and variability across students' outcomes. It should be noted that, though the findings regarding students' early literacy and early numeracy skills are positive, they do not explain why these particular skills and not others (e.g., oral comprehension and picture vocabulary) proved to result from prekindergarten enrollment. For example, perhaps one reason the math effects are particularly robust ($ES=0.48$) is due the general under-emphasis given to children's early mathematical skill development relative to that of early literacy. Indeed, research suggests that many parents do not view early math as an important area in which to focus attention (Blevens-Knabe, Austin, Musun, Eddy, & Jones, 2000). Consequently, it can be inferred that math is left more to the school to provide instruction and therefore greater differences in the math skills are indicated among students who did or did not attend prekindergarten.

Similarly, the lack of findings for oral comprehension and vocabulary may be interpreted in a variety of ways. One might hypothesize that the null finding represents a need for the additional infusion of these topics into the classroom. Alternatively, the lack of difference could be related to alternative care providers doing a fairly good job of enhancing children's skills in these arenas. This is in contrast to building children's basic reading skills that may require greater training and may help to explain the positive findings from the study. Whatever the reasons for these findings might be, it is clear that additional research is needed to begin to unpack the reasons underlying the causal relationship. For example, observations of early learning centers, including a deep analysis of the curriculum and instruction, would be useful to better understand these findings, as would visits to children in alternative settings (i.e., with a family member, in a home-based daycare, etc.) to evaluate whether and to what degree the aforementioned skills were being cultivated.

5.4.2 Commissioning Future Statewide Prekindergarten Studies

This study represented the first statewide study of prekindergarten in Connecticut. The following lessons learned from this study will be useful in conducting future evaluation studies.

Low Participation

Historically, the state has been committed to a governance model that allows for local decision-making regarding participation, and ethical concerns regarding the ability of a parent/guardian to determine the participation of their child in this type of study. As this study demonstrated, this model contributed to a reduction in participation rates of districts, schools, centers, and students. Options for strategies to be considered to increase participation rates include:

- Mandate that schools and centers receiving state funding for prekindergarten programs and kindergarten participate in state-mandated studies. This change would require schools and centers to notify parents and legal guardians of study testing consistent

with how parents are notified of annual state assessment testing periods; and a change in Connecticut's governance model, in that decision-making would be shifted back to the state.

- Require schools and centers to file a letter of cooperation with OEC and CSDE indicating their willingness to participate in state-mandated studies. Researchers can then plan to include only those who have agreed to participate, although results would only be applicable to those that filed a letter and would not address the issue of parental opt outs.
- Incorporate low participation rates in the planning process to ensure the evaluation study is adequately powered to detect effects.
- Consider alternative levels of stipends or alternative stipend disbursement methods to incentivize a greater proportion of parents and teachers to respond to requests for data. Extant research on the general population suggests that providing a cash incentive when inviting participants to complete a survey, rather than paying participants after they complete the survey, increases response rates by as much as 17%, to achieve final response rates of between 35.8% and 50.3% (Millar, & Dillman, 2011). However, there is not yet any research to indicate what the right incentive amount is, nor research that has been conducted specifically on improving response rates for teachers or parents (c.f., Goritz, 2006; Hopkins, & Gullickson, 1992; Millar & Dillman, 2011).
- Incorporate alternative parent/guardian data collection strategies, such as a telephone administration of surveys, as was done for the Lipsey et al. (2011) or conducting the assessments and survey in a student's home. It should be noted that:
 - o Phone surveys require shorter and potentially less informative surveys, and a system in place that provides for the sharing of teacher parent phone numbers with researchers.
 - o Student home-based assessments and surveys require sharing parental contact and address information with researchers. Additionally, this option requires greater investment in study infrastructure to support such efforts.

Student-Level Data

OEC and CSDE eliminated the Prekindergarten Information Management System (PKIS) as of the summer 2014 in anticipation of replacing it with a new data collection system to be administered by OEC. The PKIS information previously collected was paramount to this study, a short term alternative student data collection process was developed by the Research Team/CASE in cooperation with OEC. The following suggestions should be considered to support future evaluation studies:

- Provide user-friendly mechanisms to facilitate the efficient transfer of school and center student data for both prekindergarten and kindergarten in a timely manner. For this study, having information such as classroom rosters (for students who attended prekindergarten last year that are attending kindergarten this year) and student demographic data (e.g. race/ethnicity, socio-economic status, gender, etc.) much sooner would have accelerated the data collection timeline and may have increased participation rates. It is noted that for the analyses conducted for this study, statistical

controls were included to address shifts in the study's timeline, but it would have been preferable to be able to begin data collection in schools and centers much earlier in the school year.

- Include student demographic information (e.g. race/ethnicity, socio-economic status, gender, etc.) necessary to conduct this type of evaluation study in the new student data collection system to be administered by OEC.

5.4.3 Funding Future Research Studies

In addition to state-funded research studies, federal funding sources may be available to support studies to answer the questions cited in Table 5.2. Other states that partnered with researchers have received support for this type of research from the U.S. Department of Education's Institute for Education Sciences (Evaluation of State and Local Education Programs and Policies program, Preschool Curriculum Evaluation Research program, and Early Learning Programs and Policies program). Other state-researcher partnerships have also secured funding to study long-term outcomes, especially those associated with problem behaviors (school disciplinary incidents, delinquent behavior, etc.) from the U.S. Department of Health and Human Service's National Institutes of Health.

It is noted that, Connecticut was successful in securing external funding for this type of work as evidenced by the award of a four-year 2014 Preschool Development Grant from the US Department of Education.¹⁹

¹⁹ <http://www.ct.gov/oec/cwp/view.asp?Q=555208&A=4547>

6.0 REFERENCES

- Abedi, J. (2004). The No Child Left Behind Act and English language learners: Assessment and accountability issues. *Educational Researcher*, 33, 4-14.
- Abedi, J., Hoffstetter, C. H., & Lord, C. (2004). Assessment accommodations for English Language Learners: Implications for policy-based empirical research. *Review of Educational Research*, 74, 1-28.
- American Educational Research Association (1999). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- American Evaluation Association (2004, July). *Guiding Principles for Evaluators*. Retrieved from <http://www.eval.org/p/cm/ld/fid=51>.
- Applied Survey Research (2013, August). *Evaluating Preschool for All Effectiveness*. Research Brief. San Jose, CA: Applied Survey Research.
- Bailey, A. L., & Butler, F. A. (2003). *An evidentiary framework for operationalizing academic language for broad application to K-12 education: A design document* (CSE Tech. Rep. No. 611). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- Baker, E. L., Linn, R. L., Herman, J. L., & Koretz, D. (2002). *Standards for educational accountability systems* (Policy Brief 5, Winter). Los Angeles, CA: University of California, Center for the Study of Evaluation/National Center for Research on Evaluation, Standards, and Student Testing.
- Barnett, W. S., Carolan, M. E., Fitzgerald, J., & Squires, J. H. (2012). *The state of preschool 2012: State preschool yearbook*. New Brunswick, NJ: National Institute for Early Education Research. Retrieved from <http://nieer.org/sites/nieer/files/yearbook2012.pdf>.
- Barnett, S., & Carolan, M. (2013). Trends in State Funded Preschool in the United States: Findings from 10 Years of Policy Surveys. *International Journal of Child Care and Education Policy*, 7, 5-23.
- Barnett, W. S., Carolan, M. E., Squires, J. H., & Clarke Brown, K. (2013). *The state of preschool 2013: State preschool yearbook*. New Brunswick, NJ: National Institute for Early Education Research. Retrieved from <http://nieer.org/sites/nieer/files/yearbook2013.pdf>.
- Bartik, T. J. (2013). *Effects of the pre-k program of Kalamazoo County Ready 4s on kindergarten entry test scores: Estimates based on data from the fall of 2011 and the fall of 2012*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, Upjohn Institute Working Paper 13-198.

- Blevins-Knabe, B., Austin, A. B., Musun, L., Eddy, A., & Jones, R. M. (2000). Family Home Care Providers' and Parents' Beliefs and Practices Concerning Mathematics with Young Children†. *Early Child Development and Care*, 165(1), 41–58.
- Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness*, 1, 289–328.
- Bloom, H. S. (2012). Modern regression discontinuity analysis. *Journal of Research on Educational Effectiveness*, 5, 43–82.
- Bracken, B. A., Keith, L.K., & Walker, K.C. (1994). Assessment of preschool behavior and social-emotional functioning: A review of thirteen third-party instruments. *Assessment in Rehabilitation and Exceptionality*, 1, 331–346.
- Butler, F. A., & Stevens, R. (1997). *Accommodation strategies for English language learners on large-scale assessments: Student characteristics and other considerations* (CSE Tech. Rep. No. 448). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 171–246). Chicago, IL: Rand McNally.
- Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin Company.
- Card, D., & Lee, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, 142, 655–674.
- Coburn, J. L. (2009). *The effect of Tennessee's prekindergarten programs on young children's school readiness skills: A regression discontinuity design*. (Doctoral dissertation). Retrieved from ProQuest, UMI Dissertations Publishing (3397432).
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, NJ: Lawrence Erlbaum.
- Collaborative Center for Literacy Development. (2014). Retrieved from [http://www.kentuckyliteracy.org/~literacy/sites/default/files/resource_tools/Peabody Picture Vocabulary Test.pdf](http://www.kentuckyliteracy.org/~literacy/sites/default/files/resource_tools/Peabody_Picture_Vocabulary_Test.pdf).
- Cook, T. D. (2008). Waiting for life to arrive: A history of the regression-discontinuity design in psychology, statistics and economics. *Journal of Econometrics*, 142, 636–654.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin Company.
- Cook, T. D., Shadish, W. R., & Wong, V. C. (2008). Three conditions under which experiments and observational studies produce comparable causal estimates: New findings from within-study comparisons. *Journal of Policy Analysis and Management*, 27, 724–750.

-
- Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York, NY: Russell Sage.
- Cronbach, L. J., & Associates (1980). *Toward reform of program evaluation: Aims, methods, and institutional arrangements*. San Francisco, CA: Jossey-Bass.
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2008). *Internet, mail, and mixed-mode surveys: The tailored design method* (3rd ed.). New York, NY: Wiley.
- Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M. & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231).
- Flanagan, D. P. (2001). Comparative features of the WJ-III Tests of Cognitive Abilities (Woodcock-Johnson III Assessment Service Bulletin No. 1). Itasca, IL: Riverside Publishing.
- Fuller, B., Kagan, S. L., & Loeb, S. (2002). New lives for poor families: Mothers and young children move through welfare reform. *The Growing Up in Poverty Wave 2 Findings: California, Connecticut, and Florida*. (Technical Report). Stanford, CA: Center for Education Policy Analysis.
- Frede, E. (2005). Assessment in a continuous improvement cycle: New Jersey's Abbott Preschool Program. New Brunswick, NJ: National Institute for Early Education Research.
- Gilliam, W. S., & Zigler, E. F. (2001). A critical meta-analysis of all impact evaluations of state-funded preschool from 1977 to 19987: Implications for policy, service delivery and program evaluation. *Early Childhood Research Quarterly*, 15, 441-473.
- Gormley, W. T. (2008). The effects of Oklahoma's pre-k program on Hispanic students. *Social Science Quarterly*, 89, 916-936.
- Gormley, W. T., & Gayer, T. (2005). Promoting school readiness in Oklahoma: An evaluation of Tulsa's Pre-K program. *Journal of Human Resources*, 40, 533-558.
- Gormley, W. T., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal pre-K on cognitive development. *Developmental Psychology*, 41, 872-884. doi: 10.1037/0012-1649.41.6.872
- Gormley, W. T., Phillips, D. A., & Gayer, T. (2008). Preschool programs can boost school readiness. *Science*, 320, 1723-1724. doi: 10.1126/science.1156019
- Goritz, A. S. (2006). Incentives in web studies: Methodological issues and a review. *International Journal of Internet Science*, 1, 58-70.
- Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses related to achievement*. New York, NY: Routledge.
- Henry, G. T. (1990). *Practical sampling*. Newbury Park, CA: Sage.

- Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, 2, 172–177.
- Hopkins, K. D., & Gullickson, A. R. (1992). *A meta-analysis of the effects of monetary gratuities*, 61, 52–62.
- Howes, C., Burchinal, M., Pianta, R., Bryant, D., Early, D., Clifford R. M., et al. (2008). Ready to learn? Children's pre-academic achievement in pre-kindergarten programs. *Early Childhood Research Quarterly*, 23, 27 – 50.
- Hustedt, J. T., Barnett, W. S., & Jung, K. (2007). *The effects of the New Mexico PreK Initiative on young children's school readiness*. New Brunswick, NJ: National Institute for Early Education Research.
- Hustedt, J. T., Barnett, W. S., Jung, K., & Figueras, A. (2008). *Impacts of New Mexico PreK on children's school readiness at kindergarten entry: Results from the second year of a growing initiative*. New Brunswick, NJ: National Institute for Early Education Research.
- Hustedt, J. T., Barnett, W. S., Jung, K., & Figueras-Daniel, A. (2009). *Continued impacts of New Mexico PreK on children's readiness for kindergarten: Results from the third year of implementation*. New Brunswick, NJ: National Institute for Early Education Research.
- Hustedt, J. T., Barnett, W. S., Jung, K., & Goetze, L. D. (2009). *The New Mexico PreK evaluation: Results from the initial four years of a new state preschool initiative – final report*. New Brunswick, NJ: National Institute for Early Education Research.
- Hustedt, J. T., Barnett, W. S., Jung, K., & Thomas, J. (2007). *The effects of the Arkansas Better Chance Program on young children's school readiness*. New Brunswick, NJ: National Institute for Early Education Research.
- Imbens, G., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies*, 79, 933–958; doi:10.1093/restud/rdr043
- Lamy, C., Barnett, W. S., & Jung, K. (2005a). *The effects of the Michigan School Readiness Program on young children's abilities at kindergarten entry*. New Brunswick, NJ: National Institute for Early Education Research.
- Lamy, C., Barnett, W. S., & Jung, K. (2005b). *The effects of South Carolina's Early Childhood Programs on young children's school readiness*. New Brunswick, NJ: National Institute for Early Education Research.
- Lamy, C., Barnett, W. S., & Jung, K. (2005c). *The effects of West Virginia's Early Education Program on young children's school readiness*. New Brunswick, NJ: National Institute for Early Education Research.
- Lee, D. & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48, 281–355.

- Lipsey, M. W., Farran, D.C., & Hofer, K. G., (2015, September). *A Randomized Control Trial of the Effects of a Statewide Voluntary Prekindergarten Program on Children's Skills and Behaviors through Third Grade* (Research Report). Nashville, TN: Vanderbilt University, Peabody Research Institute.
- Lipsey, M. W., Farran, D. C., Billbrey, C., Hofer, K. G., & Dong, N. (2011, April). *Initial results of the evaluation of the Tennessee voluntary Pre-K Program* (Research Report). Nashville, TN: Vanderbilt University, Peabody Research Institute.
- Lipsey, M. W., Weiland, C., Yoshikawa, H., Wilson, S. J., & Hofer, K. G. (2015). The prekindergarten age-cut-off regression-discontinuity design: Methodological issues and implications for application. *Educational Evaluation and Policy Analysis*, 37, 296–313.
- Loeb, S., Fuller, B., Kagan, S. L., & Carrol, B. (2003). How welfare reform affects young children: Experimental findings from Connecticut – a research note. *Journal of Policy Analysis and Management*, 22, 537–550.
- Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., & Busick, M. D. (2012). *Translating the Statistical Representation of the Effects of Education Interventions into More Readily Interpretable Forms*. National Center for Special Education Research. Retrieved from <http://files.eric.ed.gov/fulltext/ED537446.pdf>.
- Magnuson, K. A., Ruhm, C., & Waldfogel, J. (2007). Does prekindergarten improve school preparation and performance? *Economics of Education Review*, 26, 33–51.
- McClendon, D., Warren, J., Green, K., Burlingam, G., Eggett, D., & McCledon R. (2011). Sensitivity to change of youth treatment outcome measures: A comparison of the CBCL, BASC-2, and Y-OQ. *Journal of Clinical Psychology*, 67, 111–125.
- McGrew, K. S., Schrank, K. A., & Woodcock, R. W. (2007). *Woodcock-Johnson III Normative Update, Technical manual*. Rolling Meadows, IL: Riverside.
- McNamara, J. R., Hollmann, C., & Riegel, T. (1994). A preliminary study of the usefulness of the Behavior Assessment System for Children in the evaluation of mental health needs in a Head Start population. *Psychological Reports*, 75, 1195–1201.
- Millar, M. M. (2011). Improving response to web-based and mixed-mode surveys. *Public Opinion Quarterly*, 75, 249–269.
- NICHD Early Child Care Research Network & Duncan, G. J. (2003). Modeling the impacts of child care quality on children's preschool cognitive development. *Child Development*, 74, 1454–1475.
- Pearson Education, Inc. (2011). Peabody picture vocabulary test (4th ed.) (PPVT™-4). Retrieved from <http://www.pearsonclinical.com/language/products/100000501/peabody-picture-vocabulary-test-fourth-edition-ppv4.html?Pid=PAa30700#tab-scoring>.

- Peisner-Feinberg, E. S., & Schaaf, J. M. (2011). Evaluation of the North Carolina More at Four Pre-Kindergarten Program: Summary of key findings. Chapel Hill, NC: The University of North Carolina, FPG Child Development Institute. Retrieved from <http://fpg.unc.edu/resources/effects-north-carolina-more-four-pre-kindergarten-program-childrens-school-readiness-skill>.
- Peisner-Feinberg, E. S., & Schaaf, J. M. (2011). *Evaluation of the North Carolina More at Four pre-kindergarten program: Summary of key findings*. Chapel Hill, NC: The University of North Carolina, FPG Child Development Institute.
- Peisner-Feinberg, E. S., Schaaf, J. M., LaForett, D. R., Hildebrandt, L. M., Sideris, J. (2014). *Effects of Georgia's pre-k program on children's school readiness skills: Findings from the 2012–2013 evaluation study*. Chapel Hill, NC: The University of North Carolina, FPG Child Development Institute.
- Pianta, R. C., Cox, M. J., & Snow, K. L. (Eds.). (2007). *School readiness and the transition to kindergarten in the era of accountability*. Baltimore, MD: Brookes Publishing.
- Ramey, C.T., Bryant, D.M., & Suarez, T.M. (1985). Preschool compensatory education and the modifiability of intelligence: A critical review. In D. Detterman (Ed.), *Current topics in human intelligence* (pp. 247-296). Norwood, NJ: Ablex.
- Ramey, T. C., Campbell, F. A., Burchinal, M., Skinner, M. L., Gardner, D. M., & Ramey, S. R. (2000). Persistent effects of early childhood education on high-risk children and their mothers. *Applied Developmental Science*, 4, 2–14.
- RFP #DHE0901 - Early Childhood Education Study, for the Connecticut Department of Higher Education; 12/19/08.
- Riley-Ayers, S., Barnet, S. W., & Frede, E. (2008) Prekindergarten Effectiveness Research Designs for the State of Connecticut, *National Institute for Early Education Research (NIEER)*; 03/26/08.
- Sandoval, J., & Echandia, A. (1994). Behavior Assessment System for Children. *Journal of School Psychology*, 32, 419–425.
- Schechter, C. (2008) A Proposal: A Preschool Effectiveness Research Design; 04/23/08.
- Schochet, P. Z. (2008). *Technical methods report: Guidelines for multiple testing in impact evaluations* (NCEE 2008-4018). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Schochet, P., Cook, T. D., Deke, J., Imbens, G., Lockwood, J. R., Porter, J., & Smith, J. (2010). *Standards for Regression Discontinuity Designs*. Retrieved from What Works Clearinghouse Web site: http://ies.ed.gov/ncee/wwc/pdf/wwc_rd.pdf.
- Shadish, W. R., Cook, T. D. & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

-
- Shepard, L. A., Taylor, G., & Betebenner, D. W. (1998). *Inclusion of Limited English Proficient students in Rhode Island's grade 4 mathematics performance assessment* (CSE Tech. Rep. No. 486). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- Thistlethwaite, D., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, 51, 309–317.
- Thurlow, M., Liu, K., Erickson, R., Spicuzza, R., & El Sawaf, H. (1996, August). *Accommodations for students with limited English proficiency: Analyses of guidelines from states with graduation exams* (Minnesota Rep. 6). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Weiland, C. & Yoshikawa, H. (2013). The impacts of an urban public prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills: Evidence from Boston. *Child Development*, 84, 2112–2130.
- Weiss, C. H. (1976). Using research in the policy process: Potential and constraints. *Policy Studies Journal*, 4, 224–228.
- Weiss, C. H. (1979). The many meanings of research utilization. *Public Administration Review*, 39, 426–431.
- Weiss, C. H. (1997). *Evaluation: Methods for studying programs and policies* (2nd Ed.). Englewood Cliffs, NJ: Prentice Hall.
- Weiss, C. H. & Bucuvalas, M. (1980). *Social science research and decision-making*. New York, NY: Columbia University Press.
- Whitehurst, G. J. (2014, February). Does Pre-k work? It depends how picky you are. Retrieved from <http://www.brookings.edu/blogs/brown-center-chalkboard/posts/2014/02/26-does-prek-work-whitehurst>.
- Wilson, S. J., Dickinson, D. K., & Rowe, D. W. (2013). Impact of an early reading first program on the language and literacy achievement of children from diverse language backgrounds. *Early Childhood Research Quarterly*, 28, 578–592.
- Witt, J.C., & Jones, K.M. (1998). Review of the Behavior Assessment System for Children. In J. C. Impara & B. S. Plake (Eds.), *The Thirteenth Mental Measurements Yearbook* (pp. 131–133). Lincoln, NE: Buros Institute of Mental Measurements.
- Wong, V. C., Cook, T. D., Barnett, W. S., & Jung, K. (2008). An effectiveness-based evaluation of five state prekindergarten programs. *Journal of Policy Analysis and Management*, 27, 122–154. doi: 10.1002/pam.20310
- Zehler, A. M., Hopstock, P. J., Fleischman, H. L., & Greniuk, C. (1994). *An examination of assessment of limited English proficient students*. Arlington, VA: Development Associates, Special Issues Analysis Center.

APPENDIX A THE RESEARCH TEAM

Faculty from UConn at the Neag School of Education led the research effort. The six members, all with different areas of expertise relating to educational research and evaluation, and all with rigorous training in methods, are best able to maximize their intellectual capital and utilize an interdisciplinary approach to the evaluation.

Dr. Bianca Montrosse-Moorhead serves as the Study Manager for the project. She is currently an Assistant Professor of Measurement, Evaluation, and Assessment in the Neag School of Education, and Program Coordinator for the Graduate Certificate Program in Program Evaluation at UConn. She earned her PhD in Psychology with an emphasis in Applied Research Methods and Evaluation from Claremont Graduate University in 2009. She has been a primary investigator, co-investigator, or senior personnel on 12 research and evaluation grants awarded for a total of \$10.9 million. Dr. Bianca Montrosse-Moorhead will provide leadership and oversight for all evaluation activities.

Dr. Tamika La Salle serves as the Data Collection Lead for the project. She is currently an Assistant Professor of School Psychology at the Neag School of Education and a Research Scientist for the Collaborative on Strategic Education Reform. Her research efforts have focused on school climate and the effects of culture on students' educational experiences and academic, social, emotional, and behavioral outcomes. She earned her doctorate from Georgia State University's Department of Counseling and Psychological Services. A former educational consultant for the University of Pittsburgh and the Georgia and Mississippi Departments of Education, her work experience includes serving as a psychometrician, school psychologist, project coordinator and survey development and validation consultant. In 2013, she earned a doctorate from Georgia State University's Department of Counseling and Psychological Services.

Dr. Shaun Dougherty serves as the Data Analysis Lead for the project. He is currently an Assistant Professor of Educational Policy and Leadership at the Neag School of Education, an associate at the Center for Education Policy Analysis, and an affiliated faculty member in the Program Evaluation certificate program. In 2013, he earned a doctoral degree in quantitative policy analysis from Harvard University. Dr. Dougherty will provide oversight on data analysis associated with the project.

Dr. Hannah Dostal is an Assistant Professor of Reading Education in the Department of Curriculum and Instruction at UConn. Dr. Dostal is a certified reading specialist and holds a PhD in Education with a concentration in literacy studies and deafness. Dr. Dostal has previously served as a middle school teacher and a writing intervention coordinator for students in grades prekindergarten through 12. She has also worked with school districts and professional associations to build capacity for implementing the English/Language Arts Common Core State Standards across the eastern United States. Dr. Dostal serves as a co-PI on improving teacher quality grants and an Institute of Educational Sciences development grant. As such, she has extensive experience with onsite data collection and project coordination and will assist Dr. La Salle with data collection efforts.

Dr. Jennie Weiner is an Assistant Professor of Educational Leadership at the Neag School of Education, and an affiliated faculty member in the Program Evaluation certificate program. Most recently, she has worked with the Rhode Island Department of Education on issues of school turnaround and capacity building. She also previously served as a senior research associate for the Teacher Advancement Program (TAP) at the Milken Family Foundation. Dr. Weiner earned a doctoral degree with an emphasis in educational policy, leadership, and instructional practice from Harvard University in 2012. She will assist Dr. Montrosse-Moorhead with study report writing efforts.

Dr. Montrosse-Moorhead, Dr. LaSalle, and Dr. Dougherty each had one graduate student assist them with the study efforts.

APPENDIX B

SUMMARY OF AGE CUT-OFF RD STUDIES INVESTIGATING THE EFFECTS OF PUBLICLY FUNDED PREKINDERGARTEN PROGRAMS ON CHILDREN'S ACADEMIC AND SOCIAL-EMOTIONAL SKILLS

STUDY: APPLIED SURVEY RESEARCH (2013); STUDY SITE: SAN FRANCISCO, CA
Data Collection Window: Fall 2012

Study Sample: N=606 included in final analytic sample (n=350 preschool students and n=256 kindergarten students)

Child Outcome Measures

1. Receptive vocabulary via the Receptive One-Word Picture Vocabulary Test-4 (ROWPVT); English and Spanish-Bilingual (SBE) editions.
2. Early literacy via the WJ-III Letter-Word Identification subscale (including Spanish version, Bateria III Woodcock-Muñoz Pruebas de Aprovechamiento).
3. Early math via the WJ-III Applied Problems subscale (including Spanish version, Bateria III Woodcock-Muñoz Pruebas de Aprovechamiento).
4. Self-regulation skills via the Head-Toes-Knees-Shoulders Task (HTKS), including a Spanish and Chinese option;
5. Social skills and problem behavior via the Preschool and Kindergarten Behavior Scales (PKBS).

Control Variables: None reported

Academic and Social/ Emotional Outcomes: No effect was found for receptive vocabulary. However, effects were found for the other 3 outcome measures. A statistically and practically significant effect was found for early literacy; specifically, PK attendance was associated with an average increase of 2.24 points (approximately equal to a 3-month gain). A statistically and practically significant effect was also found for early math skills; specifically, PK attendance was associated with an average increase of 1.59 points (approximately equal to a 3- to 4-month gain). Finally, a statistically significant effect was found for Self-regulation skill; specifically, PK attendance was associated with an average increase of 6.34 points.

STUDY: BARTIK (2013); STUDY SITE: KALAMAZOO COUNTY, MI
Data Collection Window: October and November

Study Sample: N=113 included in final analytic sample (n=62 preschool students and n=51 kindergarten students)

Child Outcome Measures:

1. Vocabulary via the Peabody Picture Vocabulary Test, 4th edition.
2. Early literacy via the WJ-III Letter-Word Identification subscale.
3. Early math via the WJ-III Applied Problems subscale.
4. Behavioral protective factors via positive behavioral items included in the Devereux Early Childhood Assessment (DECA).
5. Behavioral concerns via negative behavioral items included in the DECA.

Control Variables: Child's Age and Other Child/Family Characteristics

Academic and Social/Emotional Outcomes: Statistically significant effects were found for 2 of the 5 outcome measures. Irrespective of the specification, PK attendance was associated with an increase in early math scores (effect sizes ranged from 0.70 to 0.97). In some specifications, but not all, PK attendance was associated with an increase in vocabulary scores (effect sizes ranged from 0.60 to 0.74).

STUDY: Coburn (2009); **STUDY SITE:** Upper Cumberland region of TN

Data Collection Window: Upon PK and K entry during the fall

Study Sample: N=247 included in final analytic sample (n=179 preschool students and n=68 kindergarten students)

Child Outcome Measures:

1. Brigance Screen age equivalent scores, specifically the Preschool Screen II (which measures language comprehension, vocabulary and morphology, and syntax) and K Screen (which measures language comprehension, decoding, cipher knowledge, letter knowledge, vocabulary and morphology, and syntax)

Control Variables: Child's eligibility for free or reduced-priced lunch and gender

Academic and Social/Emotional Outcomes: Results showed a statistically significant effect whereby PK attendance was associated with an increase in Brigance Screen age equivalent scores.

STUDY: Frede (2005); **STUDY SITE:** NJ

Data Collection Window: Not reported, but presumably the Fall

Study Sample: N=more than 3,400. 865 included in final analytic sample (number of preschool and kindergarten students not provided)

Child Outcome Measures:

1. Vocabulary via the Peabody Picture Vocabulary Test, 3rd edition (including Spanish version, Test de Vocabulario en Imagenes Peabody).

2. Early math via the WJ-III Applied Problems subscale (including Spanish version, Bateria III Woodcock-Muñoz Preuebas de Aprovechamiento).
3. Early literacy via the Print Awareness subtest of the Preschool Comprehensive Test of Phonological and Print Processing.
4. State-level classroom quality via the Early Childhood Environment Rating Scale – Revised, the Support for Early Literacy Assessment, and the Preschool Classroom Mathematics Inventory.
5. Child progress via a state-developed benchmark assessment designed to measure children’s oral language and literacy skills through teacher observation and portfolio documentation.
6. Program quality via a state-developed Self-Assessment Validation System.

Control Variables: None Reported

Academic and Social/Emotional Outcomes: Statistically significant effects were found for 2 of the 3 outcome measures. PK attendance was associated with better vocabulary outcomes (4 point difference that is the equivalent to nearly four months in vocabulary development), and better literacy outcomes (76% versus 44% of items answered correctly). No difference in math outcomes was detected.

STUDY: Gormley & Gayer (2005); **STUDY SITE:** Tulsa, OK

Data Collection Window: Upon PK and K entry during the fall

Study Sample: N=3,560 included in final analytic sample (n=1,690 preschool students and n=3,441 kindergarten students).

Child Outcome Measures:

1. Early childhood skills via all items included in the Early Childhood Skills Inventory.
2. Social/ emotional skills via specific items included in the Early Childhood Skills Inventory.
3. Cognitive/ knowledge skills via specific items included in the Early Childhood Skills Inventory.
4. Motor skills via specific items included in the Early Childhood Skills Inventory.
5. Language skills via specific items included in the Early Childhood Skills Inventory.

Control Variables: Child’s Age and Other Child/ Family Characteristics

Academic and Social/Emotional Outcomes:

Statistically significant effects were found for 3 of the 5 outcome measures. PK attendance was associated with an increase in cognitive/knowledge scores (approximately 0.39 standard deviation), an increase in language scores (approximately 0.38 standard deviation), and an increase in motor skills (approximately 0.24 standard deviation). Further, subgroup analyses revealed positive effects for race (i.e., Hispanic students experienced the greatest gains, followed

by Black students, and then White students), and for socio-economic status (i.e., students who qualified for the free-lunch program experienced significant gains).

STUDY: Gormley (2008); Gormley, Gayer, Phillips, & Dawson, (2005); Gormley, Phillips, & Gayer (2008); **STUDY SITE:** Tulsa, OK

Data Collection Window: September (*Collected by teachers.)

Study Sample: N=5,294 included in final analytic sample (n=1,567 preschool students and n=3,727 kindergarten students)

Child Outcome Measures:

1. Early literacy via the WJ-III Letter-Word Identification subscale.
2. Early spelling via the WJ-III Spelling subscale.
3. Early math via the WJ-III Applied Problems subscale.

Control Variables: Child's Age and Other Child/ Family Characteristics

Academic and Social/Emotional Outcomes: Statistically significant effects were found for all outcome measures. PK attendance was associated with an increase in early literacy skills (ES=0.985), an increase in spelling skills (ES=0.743), and an increase in math skills (ES=0.355). Similar to the earlier study, subgroup analyses revealed positive effects across race categories (i.e., White, Black, Hispanic, and Native American), and for children in low-wealth families (i.e., free-and reduced-priced lunch program participation)

STUDY: Hustedt, Barnett, & Jung (2007); Hustedt, Barnett, Jung, & Figueras (2008); Hustedt, Barnett, Jung, & Figueres-Daniel (2009); Hustedt, Barnett, Jung, & Goetze (2009); **STUDY SITE:** NM

Data Collection Window: Fall 2006, Fall 2007, and Fall 2008

Study Sample: N=3,048 included in final analytic sample (number of preschool and kindergarten students not provided)

Child Outcome Measures:

1. Vocabulary via the Peabody Picture Vocabulary Test, 4th edition (including Spanish version, Test de Vocabulario en Imagenes Peabody).
2. Early math via the WJ-III Applied Problems subscale (including Spanish version, Bateria III Woodcock-Muñoz Preuebas de Aprovechamiento).
3. Early literacy via the Print Knowledge subtest of the Test of Preschool Early Literacy (early study efforts used the Pre-CTOPPP to measure this construct for all students and later only for students who were best tested in Spanish).
4. Overall classroom quality via the Early Childhood Environment Rating Scale – Revised.
5. Early language and literacy practices via the Support for Early Literacy Assessment.

6. Early math practices via the Preschool Classroom Mathematics Inventory.

Control Variables: Child's gender, ethnicity, and primary language spoken at home.

Academic and Social/Emotional Outcomes: Statistically significant effects were found for all outcome measures across time. PK attendance was associated with better vocabulary outcomes (average of 5.44 points higher), better math outcomes (average of 1.63 points higher), and better literacy outcomes (average of 24% higher).

STUDY: Hustedt, Barnett, Jung, & Thomas (2007); **STUDY SITE:** AZ

Data Collection Window: Not reported, but presumably the Fall

Study Sample: N=911 included in final analytic sample (n=407 preschool students and n=504 kindergarten students)

Child Outcome Measures:

1. Vocabulary via the Peabody Picture Vocabulary Test, 4th edition (including Spanish version, Test de Vocabulario en Imagenes Peabody).
2. Early math via the WJ-III Applied Problems subscale (including Spanish version, Bateria III Woodcock-Muñoz Preuebas de Aprovechamiento).
3. Early literacy via the Print Awareness subtest of the Preschool Comprehensive Test of Phonological and Print Processing.

Control Variables: None Reported

Academic and Social/Emotional Outcomes: Statistically significant effects were found for all outcome measures across time. PK attendance was associated with better vocabulary outcomes (31% more growth over the year), better math outcomes (37% more growth over the year), and better literacy outcomes (116% more growth over the year).

STUDY: Lamy, Barnett, & Jung (2005a); **STUDY SITE:** MI

Data Collection Window: Not reported, but presumably the Fall

Study Sample: N=865 included in final analytic sample (n=481 preschool students and n=384 kindergarten students)

Child Outcome Measures:

1. Vocabulary via the Peabody Picture Vocabulary Test, 3rd edition (including Spanish version, Test de Vocabulario en Imagenes Peabody).
2. Early math via the WJ-III Applied Problems subscale (including Spanish version, Bateria III Woodcock-Muñoz Preuebas de Aprovechamiento).
3. Early literacy via the Print Awareness subtest of the Preschool Comprehensive Test of Phonological and Print Processing.

Control Variables: None Reported

Academic and Social/Emotional Outcomes: Statistically significant effects were found for all outcome measures. PK attendance was associated with better vocabulary outcomes (24% more growth over the year), better math outcomes (64% more growth over the year), and better literacy outcomes (117% more growth over the year).

STUDY: Lamy, Barnett, & Jung (2005b); **STUDY SITE:** SC
Data Collection Window: Not reported, but presumably the Fall

Study Sample: N=777 included in final analytic sample (n=422 preschool students and n=355 kindergarten students)

Child Outcome Measures:

1. Vocabulary via the Peabody Picture Vocabulary Test, 3rd edition (including Spanish version, Test de Vocabulario en Imagenes Peabody).
2. Early literacy via the Print Awareness subtest of the Preschool Comprehensive Test of Phonological and Print Processing.

Control Variables: None Reported

Academic and Social/Emotional Outcomes: Statistically significant effects were found for all outcome measures. PK attendance was associated with better vocabulary outcomes (42% more growth over the year), and better literacy outcomes (102% more growth over the year).

STUDY: Lamy, Barnett, & Jung (2005c); **STUDY SITE:** WV
Data Collection Window: Not reported, but presumably the Fall

Study Sample: N=720 included in final analytic sample (n=341 preschool students and n=379 kindergarten students)

Child Outcome Measures:

1. Vocabulary via the Peabody Picture Vocabulary Test, 3rd edition (including Spanish version, Test de Vocabulario en Imagenes Peabody).
2. Early math via the WJ-III Applied Problems subscale (including Spanish version, Bateria III Woodcock-Muñoz Preubas de Aprovechamiento).
3. Early literacy via the Print Awareness subtest of the Preschool Comprehensive Test of Phonological and Print Processing.

Control Variables: None Reported

Academic and Social/Emotional Outcomes: Statistically significant effects were found for all outcome measures. PK attendance was associated with better vocabulary outcomes (30% more growth over the year), better math outcomes (63% more growth over the year), and better literacy outcomes (121% more growth over the year).

STUDY: Lipsey, Farran, Bilbrey, Hofer, & Dong (2011); **STUDY SITE:** TN

Data Collection Window: Fall, with an average lag time between the time school commenced and the time the outcome measures were administered equal to 73.9 days (SD=26.8 days).

Study Sample: N=303 included in final analytic sample (n=73 preschool students and n=230 kindergarten students)

Child Outcome Measures:

1. Early literacy via the WJ-III Letter-Word Identification and the Spelling subscales.
2. Language via the WJ-III Oral Comprehension and Picture Vocabulary subscales.
3. Early math via the WJ-III Applied Problems and Quantitative Concepts subscales.
4. Composite Readiness Measure via average of W-scores on all WJ-III subscales.
5. Student behavior via the CooperFarran Behavioral Rating Scale.
6. Kindergarten preparedness via the Academic Classroom and Behavior Record.

Control Variables: Propensity score analysis to adjust for baseline difference, creation of sampling weights whereby the proportion of PK and no PK cases was reflective of the proportions observed in the larger study sample, child demographic characteristics (gender, ethnicity, ELL status, and timing of the outcome assessments), and blocking on classroom.

Academic and Social/Emotional Outcomes: Results showed a statistically significant effect whereby PK attendance was associated with better outcomes on all WJ-III subtests. More specifically, effect size estimates ranged between 0.64 and 0.81 on the Composite measure, between 0.67 and 0.82 on the Letter-Word Identification subscale, between 0.64 and 0.99 on the Spelling subscale, between 0.26 and 0.38 on the Oral Comprehension subscale, between 0.44 and 0.57 on the Picture Vocabulary subscale, between 0.39 and 0.54 on the Applied Problems subscale, and between 0.34 and 0.50 on the Quantitative Concepts subscale.

STUDY: Peisner-Feinberg & Schaaf (2011); Peisner-Feinberg & Schaaf (2010); **STUDY SITE:** NC

Data Collection Window: Not reported, but presumably the Fall

Study Sample: N=1,010 included in final analytic sample (n=501 preschool students and n=509 kindergarten students)

Child Outcome Measures:

1. Early literacy via the Peabody Picture Vocabulary Test (4th edition), the WJ-III Letter-Word Identification subscale, and the Print Knowledge and Phonological Awareness subtests of the Test of Preschool Early Literacy.
2. Early math via the WJ-III Applied Problems subscale and the NCEDL Counting Numbers Task.

Control Variables: None Reported

Academic and Social/Emotional Outcomes: Results showed statistically significant effects on 5 of the 6 measures used. No difference in terms of vocabulary skills was detected (ES=0.06). However, PK attendance was associated with better outcomes letter/word knowledge (ES=1.14), phonological awareness (ES=0.56), print knowledge (ES=1.16), applied math problems (ES=0.34), and counting (ES=0.81).

STUDY: Peisner-Feinberg, Schaff, LaForett, Hildebrand, & Sideris, (2014); **STUDY SITE:** GA
Data Collection Window: Fall 2012

Study Sample: N=1,181 included in final analytic sample (n=570 preschool students and n=611 kindergarten students)

Child Outcome Measures:

1. Early language skills via the WJ-III Letter-Word Identification, Picture Vocabulary, Sound Awareness, and Work Attack subtests and the Naming Letters task.
2. Early math skills via the WJ-III Applied Problems subscale and the NCEDL Counting Numbers Task.
3. General knowledge via the Social Awareness Task.
4. Behavior skills via the Social Skills Improvement System Social Skills and Problem Behaviors subscales.

Control Variables: Child demographic characteristics, family income, and provider type (local school system or private site)

Academic and Social/Emotional Outcomes: Results showed statistically significant effects on 7 of the 10 measures used. No difference in terms of vocabulary skills (ES=0.01), social skills (ES=0.23), or problem behaviors (ES=0.10) was detected. However, children who attended PK had better outcomes in terms of Letter Knowledge (ES=0.89), Letter-Word Identification (ES=1.05), Phonological Awareness (ES=0.59), Phonemic Awareness (ES=1.20), Math Problem-Solving (ES=0.51), Counting (ES=0.86), and Social Awareness (ES=0.43).

STUDY: Weiland & Yoshikawa (2013); **STUDY SITE:** Boston, MA
Data Collection Window: September through December

Study Sample: N=2,018 included in final analytic sample (number of preschool and kindergarten students not provided)

Child Outcome Measures:

1. Receptive vocabulary via the Peabody Picture Vocabulary Test (3rd edition).
2. Pre-reading and reading skills via the WJ-III Letter-Word Identification subscale.
3. Numeracy and early math via the WJ-III Applied Problems subscale and items from the Research-Based Elementary Mathematics Assessment.

4. Executive Functioning skills via the Forward Digit Span, the Backward Digit Span, Dimensional Change Card Sort, Attention items from the Task Orientation Questionnaire, and Pencil Tapping.
5. Emotional development via the Emotion Recognition Questionnaire, Positive Emotion items from the Task Orientation Questionnaire, and Impulse Control items from the Task Orientation Questionnaire.

Control Variables: Child demographic characteristics and primary type of childcare experienced prior to PK entrance.

Academic and Social/Emotional Outcomes: Results showed statistically significant effects for 10 of the 12 measures used. Children who attended PK had better outcomes in terms of receptive vocabulary (ES=0.44), Letter-Word Identification (ES=0.62), Applied Problems (ES=0.59), Numeracy (ES=0.50), Working Memory (ES=0.23), Pencil Tap (ES=0.21), Backward and Forward Digit Span (ES=0.24), Dimensional Change Card (ES=0.28), and Emotion Recognition (0.19). However, no difference in terms of Positive Emotion (ES=0.03) and Impulse Control (ES=0.07) was detected. Subgroup analyses revealed differences for children from low-wealth families and by race/ethnicity.

STUDY: Wong, Cook, Barnett & Jung (2008); STUDY SITE: MI, NJ, OK, SC & WV

Data Collection Window: In this study, researchers combined data from six separate studies already described above. See data collection window details for:

- Frede (2005)
- Gormley & Gayer (2005)
- Gormley (2008); Gormley, Gayer, Phillips, & Dawson, (2005); Gormley, Phillips, & Gayer (2008)
- Lamy, Barnett, & Jung (2005a)
- Lamy, Barnett, & Jung (2005b)
- Lamy, Barnett, & Jung (2005c)

Study Sample: See sample characteristics described in:

- Frede (2005)
- Gormley & Gayer (2005)
- Gormley (2008); Gormley, Gayer, Phillips, & Dawson, (2005); Gormley, Phillips, & Gayer (2008)
- Lamy, Barnett, & Jung (2005a)
- Lamy, Barnett, & Jung (2005b)
- Lamy, Barnett, & Jung (2005c)

Child Outcome Measures:

1. Vocabulary via the Peabody Picture Vocabulary Test, 3rd edition (including Spanish version, Test de Vocabulario en Imagenes Peabody).
2. Early math via the WJ-III Applied Problems subscale (including Spanish version, Bateria III Woodcock-Muñoz Preuebas de Aprovechamiento).
3. Early literacy via the Print Awareness subtest of the Preschool Comprehensive Test of Phonological and Print Processing.

Control Variables: See control variables described in:

- Frede (2005)
- Gormley & Gayer (2005)
- Gormley (2008); Gormley, Gayer, Phillips, & Dawson, (2005); Gormley, Phillips, & Gayer (2008)
- Lamy, Barnett, & Jung (2005a)
- Lamy, Barnett, & Jung (2005b)
- Lamy, Barnett, & Jung (2005c)

Academic and Social/Emotional Outcomes: Results varied by state. In Michigan, PK attendance was not associated with a change in vocabulary, but math and print awareness scores were higher for children that attended PK (0.47 SD and 0.96 SD for math and print awareness, respectively). In New Jersey, improvements were observed for all outcome measures, specifically 0.36 SD for vocabulary, 0.23 SD for math, and 0.50 SD for print awareness. In Oklahoma, improvements were observed for all outcome measures, specifically 0.29 SD for vocabulary, 0.35 SD for math, and 0.43 SD for print awareness. However, the authors expressed concern about the reliability of the math and print awareness measures. In South Carolina, PK attendance was not associated with a change in vocabulary, but print awareness scores were higher for children that attended PK (0.79 SD). In West Virginia, PK attendance was not associated with a change in vocabulary or math, but print awareness scores were higher for children that attended PK (0.83 SD). Examining results across all five states, and weighting them by sample size, yields intent-to-treat estimates of 0.17 for vocabulary, 0.26 for math, and 0.68 for print awareness. Treatment-on-treated estimates across states are comparable. That is, weighted TOT estimates were 0.67 for vocabulary, 0.26 for math, and 0.68 for print awareness.

APPENDIX C

PARENT AND TEACHER QUESTIONNAIRES

The following questionnaires were administered during the data collection process:

1. Parent/Legal Guardian Questionnaire
This questionnaire requested parents/legal guardians to self-report their respective demographic information and was available in both English and Spanish.
2. Teacher Questionnaire
This questionnaire requested teachers to self-report their respective demographic information.

1. Parent/Legal Guardian Questionnaire

PARENT SURVEY

Directions: Please fill in the following information. Your responses to the following questions will be used for research purposes only and will be kept confidential.

Remember that data will be reported in aggregate (e.g., averages). No one will be able to link you to your answers in any published document or report.

Q1. If you consent to participate in this study, select “I consent”

- ☐ I consent
☐ I do not consent

Q2. Please fill in the unique 9-digit ID code you were given:

Section A: Child Demographic Information

Q3. Child’s Name

First: _____

Last: _____

Q4. Child’s birth date

Month: _____

Day: _____

Year: _____

Q5. Child’s gender?

- ☐ Male
☐ Female
☐ Other: _____

Q6. This child is your _____ child.

Q7. This child is the _____.

- ☐ Oldest
☐ Middle
☐ Youngest
☐ Only Child
☐ Other _____

Q8. What is/are the predominantly spoken language(s) at home?

Section B: Child Schooling Information
Q9. What grade is your child in? <input type="checkbox"/> Preschool (<i>Go to Q9b</i>) <input type="checkbox"/> Kindergarten (<i>Continue to next question</i>)
Q9a. Did your child attend preschool? <input type="checkbox"/> Yes (<i>Continue to next question</i>) <input type="checkbox"/> No (<i>Go to Q9c</i>)
Q9b. Some parents choose to use programs for their young children, while others keep their children at home. Prior to your child attending preschool at age 4, did your child participate in any of the following? (Check all that apply.) <input type="checkbox"/> Daycare (<i>Also answer Q9b1</i>) <input type="checkbox"/> Preschool as a 3-year old (ex: Head Start) (<i>Also answer Q9b2</i>) <input type="checkbox"/> Child stayed home with me until they entered preschool at age four (<i>Go to Q10</i>) <input type="checkbox"/> Child stayed home with family member until they entered preschool at age four (<i>Go to Q10</i>) <input type="checkbox"/> Child stayed home with friend until they entered preschool at age four (<i>Go to Q10</i>) <input type="checkbox"/> Child stayed home with paid help until they entered preschool at age four (<i>Go to Q10</i>)
Q9b1. For daycare prior to age 4... How many hours a day did the child attend? _____ How many days a week did the child attend? _____ Over the course of a year, how long was the child enrolled in the program? _____ How many daycare facilities did your child attend? _____
Q9b2. For preschool prior to age 4... How many hours a day did the child attend? _____ How many days a week did the child attend? _____ Over the course of a year, how long was the child enrolled in the program? _____ How many preschool facilities did your child attend? _____ (<i>Go to Q10</i>)
Q9c. Some parents choose to use programs for their young children, while others keep their children at home. Prior to your child attending kindergarten at age 5, did your child participate in any of the following? (Check all that apply.) <input type="checkbox"/> Daycare (<i>Continue to next question</i>) <input type="checkbox"/> Child stayed home with me until they entered kindergarten at age five (<i>Go to Q10</i>) <input type="checkbox"/> Child stayed home with family member until they entered kindergarten at age five (<i>Go to Q10</i>) <input type="checkbox"/> Child stayed home with friend until they entered kindergarten at age five (<i>Go to Q10</i>) <input type="checkbox"/> Child stayed home with paid help until they entered kindergarten at age five (<i>Go to Q10</i>)
Q9c1. For daycare prior to age 5... How many hours a day did the child attend? _____ How many days a week did the child attend? _____ Over the course of a year, how long was the child enrolled in the program? _____ How many daycare facilities did your child attend? _____
Q10. Does your child currently attend an after school program? <input type="checkbox"/> Yes (<i>Continue to next question</i>) <input type="checkbox"/> No (<i>Go to Q10b</i>)

Q10a. What type of after school program? (Check all that apply.)

- ☐ School-based
- ☐ Daycare
- ☐ Sports
- ☐ Arts
- ☐ Other: _____

(Go to Q11)

Q10b. Who typically takes care of your child after school?

- ☐ Myself
- ☐ A Family Member
- ☐ A Friend
- ☐ Paid Help
- ☐ Other: _____

Q11. Does your child have a diagnosed disability?

- ☐ Yes *(Continue to next question)*
- ☐ No *(Go to Section C)*

Q11a. What is the primary diagnosis? (Check all that apply.)

- ☐ Autism
- ☐ Deaf-Blindness
- ☐ Emotional Disturbance
- ☐ Hearing Impairment
- ☐ Intellectual Disability
- ☐ Multiple Disabilities
- ☐ Orthopedic Impairment
- ☐ Specific Learning Disability
- ☐ Speech or Language Impaired
- ☐ Traumatic Brain Injury
- ☐ Visual Impairment
- ☐ Other Health Impaired
- ☐ OHIO - ADD/ADHD
- ☐ Other: _____

Section C: Reasons for Preschool Program Choice

Directions: This next set of questions asks about the reasons you chose the preschool program your child attended last year or is attending this year. If you have other children, please answer the following questions for your preschool or kindergarten child only. If you have multiple children participating in the study, please fill out a separate form for each child.

Remember that data will be reported in aggregate (e.g., averages). No one will be able to link you or your child to your answers in any published document or report. Read each item carefully before responding. Answer as honestly as you can. There are no right or wrong answers.

Q12. To what extent did the following characteristics affect your decision as to where you sent your child to preschool?

	Did not affect my decision at all.	Considered in my decision, but not important.	Significantly affected my decision.
Type of program	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Hours of program	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Willingness to accept subsidies	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Financial aid available	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Fees charged	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Location	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Public transportation accessibility	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Content of program	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Services provided (e.g., transportation, meals, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Program curriculum	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
School/center accreditation status	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Teacher tenure/turnover	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Other (please list those that significantly affected your decision)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Section D: School Climate Survey

Directions: This next set of questions asks about your perceptions of the climate at the preschool program your child attended last year or is attending this year. If you have other children, please answer the following questions for your preschool or kindergarten child only. If you have multiple children participating in the study, please fill out a separate form for each child. Read each item carefully before responding. Answer as honestly as you can. There are no right or wrong answers.

Q13. Indicate how much you agree or disagree with the following statements.

	Strongly Disagree	Disagree	Agree	Strongly Agree
Teachers at my student's school have high standards for achievement.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Teachers at my student's school work hard to make sure that students do well.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Teachers at my student's school promote academic success for all students.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My student's school sets clear rules for behavior.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My student feels safe at school.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My student feels safe going to and from school.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
School rules are consistently enforced at my student's school.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
School rules and procedures at my student's school are fair.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My student feels successful at school.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
My student is frequently recognized for good behavior.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

EARLY CHILDHOOD REGRESSION DISCONTINUITY STUDY
APPENDICES

	Strongly Disagree	Disagree	Agree	Strongly Agree
I feel comfortable talking to teachers at my student's school.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Staff at my student's school communicates well with parents.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel welcome at my student's school.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
All students are treated fairly at my student's school.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Teachers at my student's school treat all students with respect.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My student's school building is well maintained.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My student's textbooks are up to date and in good condition.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Teachers at my student's school keep their classrooms clean and organized.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am involved in the decision making process at my student's school.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am actively involved in activities at my student's school.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I frequently volunteer to help on special projects at my student's school.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Section E: Parent and Child Relationship

Directions: This next set of questions asks about activities you participate in with your child. If you have other children, please answer the following questions for your preschool or kindergarten child only. If you have multiple children participating in the study, please fill out a separate form for each child. Remember that data will be reported in aggregate (e.g., averages). No one will be able to link you or your child to your answers in any published document or report.

Q14. What activities do you and your child participate in together at home? (Check all that apply.)

- ☐ Reading
- ☐ Playing board games
- ☐ Doing arts and crafts
- ☐ Playing sports
- ☐ Watching movies
- ☐ Other: _____

Q15. As a parent, do you struggle with any of the following? (Check all that apply.)

- ☐ Helping with homework
- ☐ Discipline
- ☐ Spending enough time with your child
- ☐ Dealing with your child's behavioral, social, or emotional problems
- ☐ Dealing with your child's academic problems
- ☐ Motivating your child to do well in school
- ☐ Working with the school and teachers
- ☐ Other: _____

Q16. In the past year, how often have you communicated with the school about ways that you can help your child at home?

- ☐ On a weekly basis
- ☐ On a monthly basis
- ☐ Quarterly
- ☐ As requested by the teacher
- ☐ Never

Q17. How much does your current schedule prevent you from becoming involved with your child's

school (e.g., PTA, volunteering, school events)?

- ☐ It does not, I go to all of my child's school events
- ☐ A bit, I attend about half of the school events
- ☐ A lot, I only attend a few events
- ☐ Extremely, I miss all my child's school events

Q18. How much do transportation-related challenges prevent you from becoming involved with your child's current school?

- ☐ Never (*Go to Q19*)
- ☐ Sometimes (*Continue to next question*)
- ☐ Often (*Continue to next question*)
- ☐ Always (*Continue to next question*)

Q18a. If you experience transportation-related challenges, what is the primary reason?

- ☐ I don't have a car
- ☐ I rely on public transportation
- ☐ I have a car, but it is sometimes unreliable
- ☐ Other: _____

Q19. How often do you help your child engage in activities (e.g., homework, educational videos, studying) that are educational outside of school?

- ☐ Never
- ☐ Once a week
- ☐ 2-3 times a week
- ☐ 4 times a week
- ☐ Everyday

Section F: Parent/Guardian Information

Directions: This last set of questions asks about you. Remember that data will be reported in aggregate (e.g., averages). No one will be able to link you or your child to your answers in any published document or report.

Q20. What is your relationship to the child?

- ☐ Mother
- ☐ Father
- ☐ Grandparent
- ☐ Legal Guardian
- ☐ Temporary care taker, (e.g., foster parent)
- ☐ Other: _____

Q21. What is your gender?

- ☐ Male
- ☐ Female
- ☐ Other: _____

Q22. What is your Ethnicity

- ☐ Hispanic or Latino
- ☐ Not Hispanic or Latino

Q23. What is your race? (Check all that apply.)

- ☐ American Indian or Alaska Native
- ☐ Asian
- ☐ Black or African American
- ☐ Native Hawaiian or other Pacific Islander
- ☐ White or Caucasian
- ☐ Other (Please specify): _____

Q24. Please select the cultural groups that you belong to (Check all that apply.)

- | | |
|---|---------------------------------------|
| <input type="checkbox"/> U.S. American | <input type="checkbox"/> Vietnamese |
| <input type="checkbox"/> French Canadian | <input type="checkbox"/> Cape Verdean |
| <input type="checkbox"/> Chinese | <input type="checkbox"/> Jamaican |
| <input type="checkbox"/> Dominican (Dominican Republic) | <input type="checkbox"/> Guatemalan |
| <input type="checkbox"/> Irish | <input type="checkbox"/> Indian |
| <input type="checkbox"/> Italian | <input type="checkbox"/> Portuguese |
| <input type="checkbox"/> German | <input type="checkbox"/> Hungarian |
| <input type="checkbox"/> Haitian | <input type="checkbox"/> Filipino |
| <input type="checkbox"/> Japanese | <input type="checkbox"/> Indonesian |
| <input type="checkbox"/> Mexican | <input type="checkbox"/> Thai |
| <input type="checkbox"/> Puerto Rican | <input type="checkbox"/> Other: _____ |

Q25. What is your marital status?

- ☐ Single (Never married)
- ☐ Married or domestic partnership
- ☐ Divorced
- ☐ Separated
- ☐ Widowed
- ☐ Committed relationship
- ☐ Other: _____

Q26. What is the highest level of school you have completed?

<input type="checkbox"/> Elementary school <input type="checkbox"/> Middle school <input type="checkbox"/> High school, no diploma <input type="checkbox"/> High school graduate (diploma or GED equivalent) <input type="checkbox"/> Technical training (e.g., mechanic, plumber) <input type="checkbox"/> Some college credits, no degree <input type="checkbox"/> Associate's degree <input type="checkbox"/> Bachelor's degree <input type="checkbox"/> Some graduate school, no degree <input type="checkbox"/> Master's degree <input type="checkbox"/> Professional degree <input type="checkbox"/> Doctorate degree <input type="checkbox"/> Other: _____
Q27. What is your work status? <input type="checkbox"/> Full-time <input type="checkbox"/> Part-time <input type="checkbox"/> Seasonal <input type="checkbox"/> Self-employed <input type="checkbox"/> Retired <input type="checkbox"/> Not employed (Go to Q28)
Q27a. What shift(s) do you work? (Check all that apply.) <input type="checkbox"/> Morning <input type="checkbox"/> Afternoon <input type="checkbox"/> Night <input type="checkbox"/> Weekends <input type="checkbox"/> Shifts vary
Q28. What is the combined household income? <input type="checkbox"/> \$5,000 or less <input type="checkbox"/> \$5,001-10,000 <input type="checkbox"/> \$10,001-20,000 <input type="checkbox"/> \$20,001-30,000 <input type="checkbox"/> \$30,001-40,000 <input type="checkbox"/> \$40,001-50,000 <input type="checkbox"/> \$50,001-60,000 <input type="checkbox"/> \$60,001-70,000 <input type="checkbox"/> \$70,001-100,000 <input type="checkbox"/> More than \$100,000
Q29. How many individuals live in the home? <input type="checkbox"/> My child and I live alone <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> 6 or more
Q30. Is this a multi-generational home? A multi-generational home is one that may include grandparents, grandchildren, or other extended family. <input type="checkbox"/> Yes (<i>Continue to next question</i>) <input type="checkbox"/> No (<i>Go to Q31</i>)
Q30a. Who else lives in the home with the family? (Check all that apply.) <input type="checkbox"/> Grandmother

- ☐ Grandfather
- ☐ Aunt
- ☐ Uncle
- ☐ Cousins
- ☐ Friends

Q31. How many school-aged children live in the home?

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6 or more

Thank you for your time

2. Teacher Questionnaire

Online Teacher Survey TEACHER DEMOGRAPHICS

Directions: Please fill in the following information. Your responses to the following questions will be used for research purposes only and will be kept confidential. Remember that data will be reported in aggregate (e.g., averages). No one will be able to link you to your answers in any published document or report.

Q2 Please enter the unique 9-digit code you were given:

Section A: School Information

Q3 School District

Q4 School Name

Q5 How long is the school day?

- ☐ Full day (1)
- ☐ Half day (2)
- ☐ Extended school day (3)

Section B: Teaching Information

Remember that data will be reported in aggregate (e.g., averages). No one will be able to link you to your answers in any published document or report.

Q6 What is your employment status as a teacher?

- ☐ Long-term Substitute (3)
- ☐ Part-time (2)
- ☐ Full-time (1)

Q7 What grade level do you currently teach?

- ☐ Preschool (1)
- ☐ Kindergarten (2)
- ☐ Both (3)

Q7a What kind of preschool program?

- ☐ Head Start (1)
- ☐ State-funded (2)
- ☐ Private (3)
- ☐ Parochial (4)
- ☐ Other (5) _____

Q8a Years of experience teaching preschool

Q8b Years of experience teaching kindergarten

Q8c Years of experience teaching in general

Section C: Classroom Information

Remember that data will be reported in aggregate (e.g., averages). No one will be able to link you to your answers in any published document or report.

Q9 How many students are enrolled in your class?

Q10 Are there other professionals that assist in your classroom?

- ☐ Yes (1)
- ☐ No (2)

Q10a Please indicate the number of assistants you have in each category listed below.

- Paraprofessional (1)
- Assistant Teacher (2)
- Certified Teacher (3)
- Parent (4)
- Volunteer (5)
- Other (6)

Q11 What is the primary curriculum used to guide literacy instruction for your students?

Q12 What is the primary curriculum used to guide math instruction for your students?

Q13 What is the primary curriculum used to guide social skills instruction for your students?

Q14 What is the primary language spoken in the classroom?

- ☐ English (1)
- ☐ Spanish (2)
- ☐ Other (3) _____

Q15 Are other languages spoken in the classroom by students, teachers, or other staff?

- ☐ Yes (1)
- ☐ No (2)

Q15a What languages?

Q16 How do you monitor a student's progress?

Section D: Education

Remember that data will be reported in aggregate (e.g., averages). No one will be able to link you to your answers in any published document or report.

Q17 What is the HIGHEST level of school you have completed? (Check one)

- ☐ High school graduate (diploma or GED equivalent) (1)
- ☐ Some college credits, no degree (2)
- ☐ Associate's degree (3)
- ☐ Bachelor's degree (4)
- ☐ Some graduate school, no degree (5)
- ☐ Master's degree (6)
- ☐ Professional degree (7)
- ☐ Doctorate degree (8)
- ☐ Other (9) _____

Q18 Year that you completed your most recent degree

Q19 Are you a certified teacher?

- ☐ Yes (1)
- ☐ No (2)

Q19a University attended for teaching certification

Q19b What type of teaching certification do you hold? (Check all that apply.)

- ☐ Initial Educator Certificate (1)
- ☐ Interim Educator Certificate (2)
- ☐ Provisional Educator Certificate (3)
- ☐ Professional Educator Certificate (4)
- ☐ Not certified (5)

Q19c Did you complete an alternative certification program?

- ☐ Yes (1)
- ☐ No (2)

Q19d What alternative certification program did you complete?

Q19e Please list additional certifications and/or licenses earned:

- Certification/License 1 (1)
- Certification/License 2 (4)
- Certification/License 3 (5)

Section E: Professional Development

Remember that data will be reported in aggregate (e.g., averages). No one will be able to link you to your answers in any published document or report. Read each item carefully before responding. Answer as honestly as you can. There are no right or wrong answers.

Q20 During the last 12 months, have you participated in any of the following types of professional development? (According to NAEYC, professional development is defined as "a continuum of learning and support activities designed to prepare individuals for work with and on behalf of young children and their families, as well as ongoing experiences to enhance this work.") (Check all that apply.)

- ☐ Informal dialogue with colleagues on how to improve your teaching (1)
- ☐ Course(s) or workshops(s) (3)
- ☐ Reading professional literature (e.g., journals, evidence-based papers) (4)
- ☐ Education conference(s) or seminar(s) (5)
- ☐ Professional development network (6)
- ☐ Individual or collaborative research (7)
- ☐ Coaching/mentoring (8)
- ☐ Peer observation of your classroom (9)
- ☐ Observation visits to other schools (10)
- ☐ Other type of professional development (11) _____
- ☐ Other type of professional development (12) _____
- ☐ Other type of professional development (13) _____

Q21 In the last 12 months, indicate which professional development activities you have participated in based on the list below: (Check all that apply.)

- ☐ Connecticut Preschool Assessment Framework (1)
- ☐ Connecticut Preschool Curriculum Framework (10)
- ☐ Connecticut Early Learning and Development Standards (11)
- ☐ Content covered by standards (Language and Literacy) (12)
- ☐ Content covered by standards (Mathematics) (13)
- ☐ Content covered by standards (Social Skills) (14)
- ☐ Instructional practices (3)
- ☐ Assessment practices (17)
- ☐ Classroom management strategies (18)
- ☐ Technology in the classroom (19)

- ☐ Communicating and partnering with families (20)
- ☐ Developmentally appropriate processes and learning experiences (21)
- ☐ Teaching students with disabilities (5)
- ☐ Teaching dual language learners (6)
- ☐ Multicultural education in early education classrooms (7)
- ☐ Other topic (8) _____
- ☐ Other topic (9) _____
- ☐ Other topic (4) _____

Section F: School Climate

Remember that data will be reported in aggregate (e.g., averages). No one will be able to link you to your answers in any published document or report. Read each item carefully before responding. Answer as honestly as you can. There are no right or wrong answers.

Q22 Indicate how much you agree or disagree with the following statements.

EARLY CHILDHOOD REGRESSION DISCONTINUITY STUDY
APPENDICES

	Strongly Disagree (1)	Disagree (2)	Agree (3)	Strongly Agree (4)
I feel supported by other teachers at my school. (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I get along well with other staff members at my school. (2)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel like I am an important part of my school. (3)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I enjoy working in teams (e.g. grade level, content) at my school. (4)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel like I fit in among other staff members at my school. (5)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel connected to the teachers at my school. (6)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Teachers at my school frequently recognize students for good behavior. (7)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Teachers at my school have high standards for achievement. (8)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My school promotes academic success for all students. (9)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
All students are treated fairly by the adults at my school. (10)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Teachers at my school treat students fairly regardless of race, ethnicity, or culture. (32)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Teachers at my school work hard to make sure that students do well. (13)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel safe at my	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

school. (14)				
I have been concerned about my physical safety at school. (15)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
If I report unsafe or dangerous behaviors, I can be sure the problem will be taken care of. (16)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel safe when entering and leaving my school building. (17)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My school building is well maintained. (19)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Instructional materials are up to date and in good condition. (20)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Teachers at my school keep their classrooms clean and organized. (21)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Teachers make an effort to keep the school building and facilities clean. (22)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Students at my school would help another student who was being bullied. (23)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Students at my school get along well with one another. (24)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Students at my school treat each other with respect. (26)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Students at my school treat other students fairly regardless of race, ethnicity, or culture. (27)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Students at my school show	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

respect to other students regardless of their academic ability. (28)				
Students at my school demonstrate behaviors that allow teachers to teach, and students to learn. (33)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Parents at my school attend PTA meetings or parent/teacher conferences. (31)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
At this school, parents frequently volunteer to help on special projects. (30)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Parents at this school frequently attend school activities. (29)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Section G: Participant Information

Remember that data will be reported in aggregate (e.g., averages). No one will be able to link you to your answers in any published document or report.

Q23 What is your gender?

- ☐ Male (1)
☐ Female (2)
☐ Other (3) _____

Q24 Which of the following categories best describes your Spanish, Hispanic, and/or Latino ethnicity? (Check all that apply.)

- ☐ I am not of Spanish, Hispanic, or Latino origin. (1)
☐ Mexican, Mexican-American, Chicano (3)
☐ Puerto Rican (4)
☐ Cuban (5)
☐ Latin American (6)
☐ Other (Please specify) (7) _____

Q25 What is your race? (Check all that apply.)

- ☐ American Indian or Alaska Native (4)
- ☐ Asian (3)
- ☐ Black or African American (2)
- ☐ Native Hawaiian or other Pacific Islander (5)
- ☐ White (1)
- ☐ Other (6) _____

APPENDIX D

UConn'S INSTITUTIONAL REVIEW BOARD (IRB)

APPROVALS

Protocol #H14-193: "Phase I of the CT Early Childhood Regression Discontinuity Study"

The following approvals were received by the UConn Research Team to conduct this study from the UConn IRB.

1. Initial Application: Provided initial approval for the study; Approved: July 2014.
2. Amendment #1: Provided for transmission and analysis of the "PKIS-like" data provided by OEC, and added a graduate student to assist the Research Team; September 2014.
3. Amendment #2: Provided for continued analysis of data transferred by OEC and CSDE to the Research Team, and added graduate students to assist the Research Team; Approved: July 2015
4. Amendment #3: Provided for Fall 2015 data collection; Approved: September 2015
5. Amendment #4: Provided for data collection into February 2016 and changes to key personnel; Approved: January 2016

1. INITIAL APPLICATION



DATE: July 31, 2014

TO: Bianca Montrosse-Moorhead, Ph.D.
Educational Psychology, Unit 3064

FROM: Pamela I. Erickson, Ph.D.
Institutional Review Board Member
FWA# 00007125

RE: Protocol #H14-193: "Phase I of the CT Early Childhood Regression Discontinuity Study"
Please refer to the Protocol# in all future correspondence with the IRB.
Funding Source: Connecticut Academy of Science and Engineering
Approval Period: From: July 31, 2014 Valid Through: July 31, 2015
"Expiration Date"

The Institutional Review Board (IRB) approved this protocol on July 31, 2014. The research presents no more than minimal risk to human subjects and qualifies for expedited approval under category #5: Research involving materials (data, documents, records, or specimens) that have been collected, or will be collected solely for nonresearch purposes (such as medical treatment or diagnosis).

The IRB has determined that the study referenced above meets the criteria for Waiver of Informed Consent stated in 45 CFR 46.116(d) as follows:

- The research involves no more than minimal risk to the subjects;
- The waiver or alteration will not adversely affect the rights and welfare of the subjects;
- The research could not be practicably carried out without the waiver or alteration; and
- Whenever appropriate, the subjects will be provided with additional pertinent information after participation in the study.

The principal investigator must notify the IRB immediately of any changes that may affect the status of the research study referenced above.

All investigators at the University of Connecticut are responsible for complying with the attached IRB "Responsibilities of Research Investigators."

Re-approval: It is the investigator's responsibility to apply for re-approval of ongoing research at least once yearly, or more often if specified by the IRB. The Re-approval/Completion Form (IRB-

Office of the Vice President for Research
Research Compliance Services
438 WHITNEY ROAD EXTENSION, UNIT 1346
STORRS, CT 06269-1246
PHONE 860.486.8902
FAX 860.486.1044
compliance.uconn.edu

An Equal Opportunity Employer

2) and other applicable re-approval materials must be submitted **one month** prior to the expiration date noted above.

Modifications: If you wish to change any aspect of this study, such as the procedures, the consent forms, the investigators, or funding source, please submit the changes in writing to the IRB using the Amendment Review Form (IRB-3). All modifications must be reviewed and approved by the IRB prior to initiation.

Audit: All protocols approved by the IRB may be audited by the Post Approval Monitor.

Please keep this letter with your copy of the approved protocol.

Attachments:

1. Validated Appendix A
2. Validated IRB-1
3. "Responsibilities of Research Investigators"

2. AMENDMENT #1



DATE: September 26, 2014

TO: Bianca Montrosse Moorhead, Ph.D.
Educational Psychology, Unit 3064

FROM: Jaci L. VanHeest, Ph.D.
Chair, Institutional Review Board
FWA# 00007125

RE: Protocol #H14-193: "Phase I of the CT Early Childhood Regression Discontinuity Study"
Please refer to the Protocol# in all future correspondence with the IRB.

The request for approval of an amendment received September 20, 2014 for the above-referenced protocol was approved by the Institutional Review Board (IRB) on September 26, 2014. This amendment is eligible for expedited review under 45 CFR 46.110(b)(2): minor changes in previously approved research during the period (of one year or less) for which approval is authorized. The amendment includes:

1. Changes to key personnel: Yujia (Eva) Li is being added as a graduate research assistant.

Amendment Approval Date: September 26, 2014
Approval is Valid Until: July 31, 2015

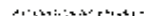
Please keep this Amendment Approval letter with your copy of the approved protocol.

Attachments:

1. Validated IRB-3 Amendment Review Form
2. Validated Revised Appendix A

Office of the Vice President for Research
Research Compliance Services
438 WHITNEY ROAD EXTENSION, UNIT 1246
STORRS, CT 06269-1246
PHONE 860.486.8802
FAX 860.486.1044
compliance.uconn.edu

An Equal Opportunity Employer



Re-approval: It is the investigator's responsibility to apply for re-approval of ongoing research at least once yearly, or more often if specified by the IRB. The Re-approval/Termination Form (IRB-2) and other applicable re-approval materials must be submitted one month prior to the expiration date noted above.

Modifications: If you wish to change any aspect of this study, such as the procedures, the consent forms, the investigators, or funding source, please submit the changes in writing to the IRB using the Amendment Review Form (IRB-3). All modifications must be reviewed and approved by the IRB prior to initiation.

Audit: All protocols approved by the IRB may be audited by the Post Approval Monitor.

Please keep this letter with your copy of the approved protocol.

Attachments:

1. Validated IRB-2 Re-approval Form
2. Validated IRB-3 Amendment Form
3. Validated Revised Appendix A
4. "Responsibilities of Research Investigators"

4. AMENDMENT #3



DATE: September 30, 2015

TO: Bianca Montrosse-Moorhead, Ph.D.
EPSY

FROM: Brandi Simonsen, Ph.D. *BS/nw*
Vice-Chair, Institutional Review Board
FWA #00007125

RE: Protocol #H15-247: "Phase 2 of the CT Early Childhood Regression Discontinuity Study"
Please refer to the Protocol# in all future correspondence with the IRB.
Funding Source: CT Academy of Science and Engineering (on behalf of the CT General Assembly)
Approval Period: From: September 30, 2015 Valid Through: September 30, 2016
"Expiration Date"

On September 3, 2015, the Institutional Review Board (IRB) reviewed the above-referenced research study by expedited review and determined that modifications were required to secure approval. Those requirements have been met, and the IRB granted approval of the study on September 30, 2015. The research presents no more than minimal risk to human subjects and qualifies for expedited approval under category #7: Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies. Enclosed are the validated consent and notification forms, which are valid through September 30, 2016. **A copy of the approved, validated consent form (with the IRB's stamp) must be used to consent each subject.**

The IRB found that the protocol meets the criteria for approval stated in 45 CFR Part 46, Subpart D, Section 404: The research presents no greater than minimal risk to the minor subjects and adequate provisions have been made to solicit verbal assent of the minor subject. The IRB has determined that the study referenced above meets the criteria for Waiver of Informed Consent for waiving parental permission as stated in 45 CFR 46.116(d) as follows:

- The research involves no more than minimal risk to the subjects;
- The waiver or alteration will not adversely affect the rights and welfare of the subjects;
- The research could not be practicably carried out without the waiver or alteration; and
- Whenever appropriate, the subjects will be provided with additional pertinent information after participation in the study.

The principal investigator must notify the IRB immediately of any changes that may affect the status of the research study referenced above.

Additionally, per 45 CFR 46.117(c)(2), the IRB waived the requirement for the investigator to obtain a signed consent form for the parent participants because it found that the research presents no more than minimal risk of harm to subjects and involves no procedures for which written consent is normally required outside of the research context.

Office of the Vice President for Research
Research Compliance Services
438 WHITNEY ROAD EXTENSION, UNIT 1246
STORRS, CT 06269-1246
PHONE 860.486.8802
FAX 860.486.1044
compliance.uconn.edu

An Equal Opportunity Employer

All investigators at the University of Connecticut are responsible for complying with the attached IRB “Responsibilities of Research Investigators.”

Re-approval: It is the investigator's responsibility to apply for re-approval of ongoing research at **least once yearly**, or more often if specified by the IRB. The Re-approval/Completion Form (IRB-2) and other applicable re-approval materials must be submitted **one month** prior to the expiration date noted above.

Modifications: If you wish to change any aspect of this study, such as the procedures, the consent forms, the investigators, or funding source, please submit the changes in writing to the IRB using the Amendment Review Form (IRB-3). All modifications must be reviewed and approved by the IRB prior to initiation.

Audit: All protocols approved by the IRB may be audited by the Research Compliance Monitor.

Please keep this letter with your copy of the approved protocol.

Attachments:

1. Validated Teacher Consent Form and Parental Notification Form/Information Sheet
2. Validated Appendix A
3. Validated IRB-1
4. “Responsibilities of Research Investigators”

5. AMENDMENT #4



Corrected Letter

DATE: January 26, 2016

TO: Bianca Montrosse-Moorhead, Ph.D.
EPSY

FROM: Brandi Simonsen, Ph.D. *BSI/nw*
Vice-Chair, Institutional Review Board
FWA #00007125

RE: Protocol #H15-247: "Phase 2 of the CT Early Childhood Regression Discontinuity Study"
Please refer to the Protocol# in all future correspondence with the IRB.

On January 7, 2016, the Institutional Review Board (IRB) reviewed a request for approval of an amendment to the above-referenced research study by expedited review, and determined that modifications were required to secure approval. Those requirements have been met, and the IRB granted approval of the study on January 26, 2016. This amendment is eligible for expedited review under 45 CFR 46.110(b)(2): minor changes in previously approved research during the period (of one year or less) for which approval is authorized.

The amendment includes:

1. Changes to key personnel: four temporary graduate student research assistants have been hired: Laura Kerns, Jesslynn Neves, Tiffany Polk, and Monique Golden.
2. Data collection will be extended into February 2016.

Amendment Approval Date: January 26, 2016

Approval is Valid Until: September 30, 2016

Please keep this Amendment Approval letter with your copy of the approved protocol.

Attachments:

1. Validated IRB-3 Amendment Review Form
2. Validated Revised Consent Forms
3. Validated Revised IRB-1 Protocol Form
4. Validated Revised Appendix A
5. Validated School/Center Frequently Asked Questions

Office of the Vice President for Research
Research Compliance Services
438 WHITNEY ROAD EXTENSION, UNIT 1248
STORRS, CT 06269-1248
PHONE 860.486.8802
FAX 860.486.1044
compliance.uconn.edu

An Equal Opportunity Employer

APPENDIX E

ANALYSIS OF POPULATION AND SAMPLES, AND OF TREATMENT AND CONTROL GROUPS

E.1. DESCRIPTIVE ANALYSIS OF SAMPLE SIZE MONITORING DURING DATA COLLECTION

Progress on data collection is documented in Tables E.1 and Table E.2. The top portion of each table (lines 2 through 5) documents reasons for attrition from the original, randomly drawn sample. These results highlight that 432 students across 191 schools were opted out of the study by superintendents, principals/center directors, or parents (line 2; column 3/11/16-Final).

In regards to language, literacy, and mathematics assessments (Table E.1 and Table E.2), the final sample for both the kindergarten group and the prekindergarten group is included in line 6 on both tables.

TABLE E.1. KINDERGARTEN SAMPLE (Language, Literacy, and Mathematics)									
	12/31/15	1/15/16	1/29/16	2/5/16	2/12/16	2/19/16	2/26/16	3/4/16	3/11/16 FINAL
(1) Goal - Student Sample Size Goal	650	650	650	650	650	650	650	650	650
(2) Opt outs (Superintendent, School/Center, Parents)	364	167	176	176	196	211	211	211	211 (n=101 schools)
(3) Group A non-responders (i.e., students with no Point-of-Contact identified at the school/center level)		32	16	13	13	0	0	0	0 (n=0 schools)
(4) Other (e.g., student moved, student can't be located, student held back, etc.)		124	128	139	127	123	125	149	149 (n=75 schools)
<i>Sub-total - Potential Student Sample Size [(1) - (2) - (3) - (4)]</i>	286	327	330	322	314	316	314	290	290
(5) Group B non-responders (i.e., students with a non-responsive Point-of-Contact; students with a Point-of-Contact that refuse to give dates)	N/A	N/A	85	81	63	79	71	71	79 (n=30 schools)
<i>Revised Sub-total - Potential Student Sample Size [(1) - (2) - (3) - (4) - (5)]</i>	286	327	245	241	251	237	243	219	211
(6) ACTUAL Total Data Collection Completed As of Date at Top of the Column	124	158	174	182	185	186	193	206	211 (n=98 schools)
<i>Note: For the 12/31/15 update, line items (2), (3), and (4) were reported as one number. Beginning with the 01/05/16 update, these line items were separated. Also, beginning with the 01/29/16 update, line item (5) was added to the weekly update table. Line items (2), (3), (4), (5), and (6) change weekly, and as a result, can influence sub-total numbers, data collection balance numbers, and weekly estimate numbers.</i>									

TABLE E.2. PREKINDERGARTEN SAMPLE (Language, Literacy, and Mathematics)									
	12/31/15	1/15/16	1/29/16	2/5/16	2/12/16	2/19/16	2/26/16	3/4/16	3/11/16 FINAL
(1) Goal - Student Sample Size Goal	650	650	650	650	650	650	650	650	650
(2) Opt outs (Superintendent, School/Center, Parents)	310	108	141	135	177	182	197	208	221 (n=90 schools)
(3) Group A non-responders (i.e., students with no Point-of-Contact identified at the school/center level)		259	171	112	1	0	0	0	0 (n=0 schools)
(4) Other (e.g., student moved, student can't be located, student held back, etc.)		18	24	30	78	46	44	66	68 (n=40 schools)
<i>Sub-total - Potential Student Sample Size</i> [(1) - (2) - (3) - (4)]	340	265	314	373	394	422	409	376	361
(5) Group B non-responders (i.e., students with a non-responsive Point-of-Contact; students with a Point-of-Contact that refuse to give dates)	N/A	N/A	96	71	35	102	98	44	28 (n=20 schools)
<i>Revised Sub-total - Potential Student Sample Size</i> [(1) - (2) - (3) - (4) - (5)]	340	265	218	302	359	320	311	332	333
(6) ACTUAL Total Data Collection Completed As of Date at Top of the Column	67	139	141	156	176	224	280	310	333 (n=78 schools)
<i>Note: For the 12/31/15 update, line items (2), (3), and (4) were reported as one number. Beginning with the 01/05/16 update, these line items were separated. Also, beginning with the 01/29/16 update, line item (5) was added to the weekly update table. Line items (2), (3), (4), (5), and (6) change weekly, and as a result, can influence sub-total numbers, data collection balance numbers, and weekly estimate numbers.</i>									

In looking at Tables E.1 and E.2, one item of note is that the decision was made at the end of February to extend data collection into March. By 2/26/16, data had been collected from a total of 473 children. At the same time, power analyses indicated that a sample size of 500 would be adequately powered to detect effects for language, literacy, and mathematics scores. Based on this information, coupled with 71 assessments already scheduled during the first two weeks of March, it was decided to move forward with collecting this data. As a result, the Research Team was able to collect data for all of the 71 students and hence has a sample size large enough to appropriately detect effects, according to the power analysis, regarding the outcomes of interest.

Another item of note is that during the final effort to complete data collection, two groups of “non-responders” emerged.

1. (Group A, line 3): Sites that refused to respond to communication, both multiple email and phone calls, from the Research Team regarding identifying a school/center Point-of-Contact (PoC). In other words, the research team was never able, despite multiple attempts in various modalities (e.g., phone, email, formal letters) to gain information regarding the PoC at these sites. Further, during the second week in February, the decision was made to move to consider school principals and center directors as the PoC, rather than relying on them to identify someone. This shift accounted for the drop in Group A (line 3) numbers during the week ending on 2/19/2016.
2. (Group B, line 5): Points-of-Contact for “non-responders” who (i) once identified, refused to respond to communication, both email and phone calls, from the Research

Team to schedule assessments, and/or (ii) responded to communication, but refused to provide dates due to continued concerns, questions, or general apprehension about the study, etc. As such, the Research Team made initial contact but was unable to complete the assessments at these sites.

During January and February, the Research Team accelerated their efforts to connect with these groups often checking with principals about the PoC. This work resulted in the drop in Group B (line 5; column 3/11/16-Final) by the end of data collection.

E.2. STATISTICAL ANALYSIS COMPARING ORIGINAL RANDOMLY DRAWN SAMPLE TO SAMPLE OF STUDENTS ASSESSED

E.2.1. Quantitative Analysis Comparing the Population, Random Sample and Sample Assessed

The Research Team conducted an analysis of the characteristics of the students in kindergarten and prekindergarten who were actually assessed and compared them to the characteristics of the students in the original sample. The purpose of this analysis was twofold:

1. To establish that the students in the actual final sample were representative of the original sample and the larger population of interest - students who are eligible for and participate in the state-funded prekindergarten School Readiness program.
2. To ensure that the birth date cut-off for participation in the state-funded prekindergarten School Readiness program was appropriately distributed in the sample to allow for use of the RD design.

Findings from the analysis indicate:

1. The students from the actual sample are generally and practically representative of the original, randomly drawn sample.
2. The distribution of birthdates in the actual sample is smooth and continuous.
3. The RD design initially proposed is still appropriate for answering the research questions.

E.2.1.1. FINDING 1 -- ANALYSES INDICATE THAT THE FINAL ASSESSED SAMPLE SUFFICIENTLY REPRESENTS THE POPULATION OF INTEREST.

In the first set of analyses, focused on comparing the final sample of students to the intended sample and overarching population of interest, the records of kindergarten students who were assessed were merged with the original sample of randomly selected kindergarten students. The merged records were then tested statistically to determine whether there were any statistically significant differences in the observable characteristics of kindergarten students from the actual sample compared to the population as a whole (Table E.3). Specifically, measures of gender, race, and free- or reduced-price lunch eligibility were tested. Disability status and ELL status could not be used due to significant amounts of missing data on these

variables. Further, age was explored separately and is included in Finding 2. Across these tests, no evidence of statistically significant differences on any of these characteristics was found when using an alpha of 0.05.

TABLE E.3. COMPARISONS BETWEEN POPULATION AND SAMPLE ASSESSED

	Prekindergarten Population (2015-16)	Prekindergarten Sample Assessed (2015-16)	Kindergarten Population (2015-16)	Kindergarten Sample Assessed (2015-16)
Gender				
<i>Male</i>	50.7%	49.7%	51.5%	45.1%
<i>Female</i>	49.3%	50.3%	48.7%	54.9%
Race/Ethnicity				
<i>White</i>	36.5%	26.0%**	34.9%	36.5%
<i>African American/ Black</i>	29.0%	31.8%	30.2%	29.3%
<i>Hispanic/Latino</i>	43.2%	45.0%	44.2%	42.8%
<i>Asian</i>	4.1%	3.7%	3.7%	3.4%
<i>Other</i>	5.2%	9.1%**	3.9%	6.3%
Lunch				
<i>Free</i>	54.2%	60.4%*	52.7%	51.9%
<i>Reduced</i>	5.7%	5.0%	5.1%	5.3%
Notes: * denotes that differences detected between population and sample assessed at * $p < 0.05$. ** denotes that differences detected between population and sample assessed at * $p < 0.01$. Sample assessed only includes students with PPVT-4 and WJ-IV scores.				

This same process was then repeated for the prekindergarten students and there were a few small, statistically significant differences between the sample assessed and the population. Differences that do exist are present only in prekindergarten with slight, though statistically significant, differences in the percent of White students (population has approximately 37% White students, while the sample assessed has approximately 26%) and students whose race was identified as "Other" (population has approximately 5% Other students, while the sample assessed has approximately 9%) and a greater proportion of students receiving free lunch (approximately 60%) relative to the population (approximately 54%). However, given the goals of the study, the sample is still considered representative of the School Readiness program population. Therefore, it was concluded that the data from the actual sample is representative of the population of interest, which is necessary to make an inference about the impact of the state-funded prekindergarten School Readiness program.

E.2.1.2. Finding 2 -- Analyses indicate that the distribution of birthdates is smooth and continuous.

Satisfied that the students in the actual sample represent the population of interest, the distribution of student birthdates from the actual sample was then assessed, as RD design relies on a relatively smooth distribution of data that constitute the forcing variable, or the variable that determines eligibility for treatment. The RD design for this study required that the student

birthdates from the actual sample be distributed smoothly around the January 1 date used to determine eligibility for prekindergarten in a given year. To test this assumption of smoothness and continuity of birthdates around the cut-off, the kindergarten and prekindergarten student records from the actual sample were combined and it was found that the distribution of birthdates is smooth and continuous (Figure E.1). This finding provides confidence that there will be a sufficient number of students near the birthdate cut-off to ensure the ability to assess the impact of being just eligible and enrolling in prekindergarten versus just missing eligibility and not enrolling in the state-funded prekindergarten School Readiness program.

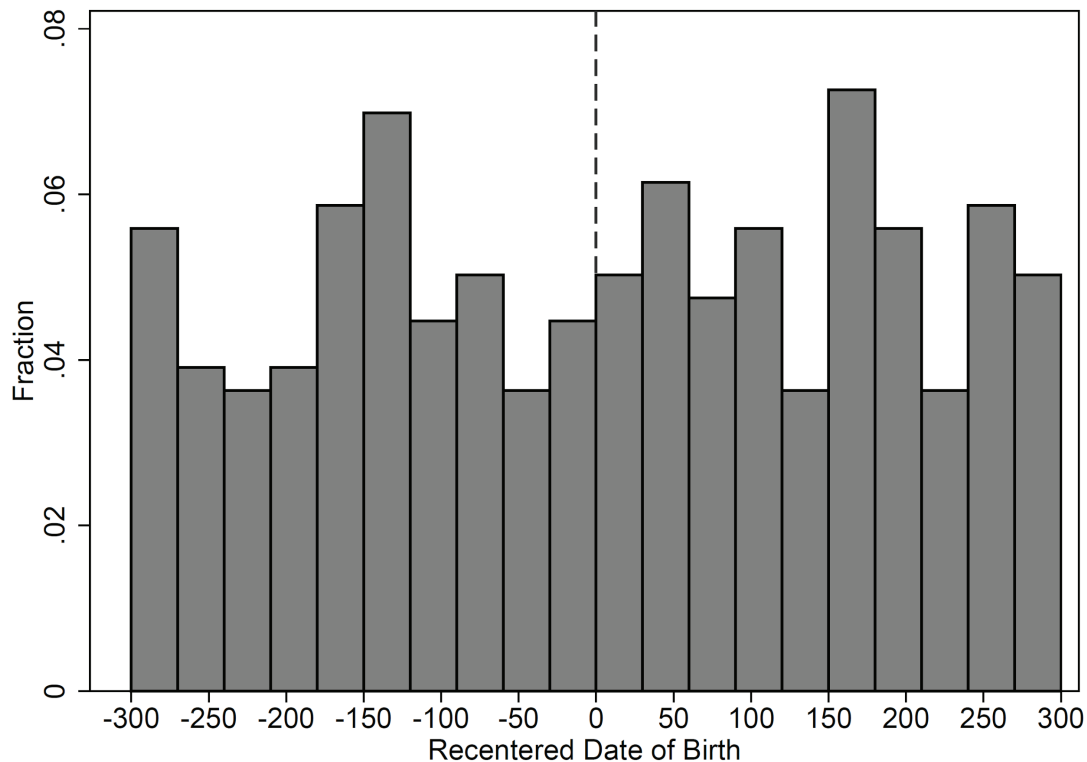


FIGURE E.1. DISTRIBUTION OF DATE OF BIRTH AROUND THE JANUARY 1 DATE USED TO DETERMINE ELIGIBILITY FOR STATE-FUNDED PREKINDERGARTEN

E.2.1.3. FINDING 3 -- ANALYSES INDICATE THAT THE TREATMENT AND COMPARISON SAMPLES ARE COMPARABLE.

Table E.4 presents results analyzing whether the treatment and comparison samples are comparable across different bandwidth specifications. The conclusion of the statistical tests that were applied determined that the assessed sample and specifically the two groups within it (i.e., prekindergarten and kindergarten students) are statistically equivalent. This is important as an assumption in the RD design is that the only true difference between the groups on either side near to the eligibility threshold, in this case students' birthdates, is whether they met the threshold or not. The lack of differences on observable demographic characteristics in the assessed sample is clearly illustrated by the lack of statistical differences on observable characteristics between treatment (experienced prekindergarten) and control (did not experience prekindergarten) near the threshold used to determine eligibility. Additionally, this lack of difference holds true as the analysis bandwidth grows larger to include more of the sample, as it did when assumptions of linearity were relaxed in further testing.

TABLE E.4. COVARIATE BALANCE AT BIRTHDAY CUT-OFF

	(1) Male	(2) White	(3) Black	(4) Latino	(5) Asian	(6) Other race	(7) Low Income
IK bandwidth	0.066 (0.137)	0.062 (0.115)	0.075 (0.145)	-0.197 (0.141)	0.028 (0.027)	0.039 (0.058)	0.016 (0.121)
N	263	285	244	266	286	285	304
Bandwidth = 120	0.106 (0.194)	0.157 (0.167)	-0.022 (0.195)	-0.164 (0.198)	0.024 (0.021)	0.122 (0.093)	0.110 (0.161)
N	139	139	139	139	139	139	139
Bandwidth = 150	0.063 (0.168)	0.123 (0.145)	0.035 (0.170)	-0.181 (0.174)	0.023 (0.028)	0.098 (0.075)	0.041 (0.150)
N	178	178	178	178	178	178	178
Bandwidth = 180	0.079 (0.151)	0.098 (0.131)	0.071 (0.152)	-0.209 (0.157)	0.022 (0.029)	0.070 (0.067)	0.020 (0.139)
N	223	223	223	223	223	223	223
Bandwidth = 210	0.066 (0.138)	0.066 (0.121)	0.081 (0.137)	-0.201 (0.143)	0.025 (0.028)	0.047 (0.061)	0.026 (0.130)
N	259	259	259	259	259	259	259
mu	0.529	0.235	0.382	0.441	0.000	0.029	0.735

Notes: Heteroskedasticity robust standard errors clustered by application score are in parentheses (* $p < .10$ ** $p < .05$ *** $p < .01$). Each coefficient is the reduced form estimate of the relationship between being eligible for state funded prek based on birthdate and the listed covariate. The coefficients shown are generated by local linear regression using a triangular kernel and specified bandwidth. Also listed is the mean of the covariate for students just before the threshold for qualifying for prekindergarten.

E.2.1.4. FINDING 4 -- ANALYSES INDICATE THE REGRESSION DISCONTINUITY (RD) DESIGN INITIALLY PROPOSED TO ANSWER THE RESEARCH QUESTIONS IS STILL APPROPRIATE.

Based on all of the robustness checks contained herein, the collected data can be analyzed in the manner originally proposed. Furthermore, the smoothness and continuity of the birthdate distribution also ensure that the RD design proposed to answer the study's research questions will be utilized as proposed.

APPENDIX F

FINAL POWER ANALYSIS

F.1. LANGUAGE, LITERACY, AND MATHEMATICS ASSESSMENTS

With a final sample size of 544 students (211 kindergarten, 333 prekindergarten), assuming a level of power of at least 0.8, and accounting for treatment and control imbalance, the study is able to detect an effect as small as 0.36 assuming inclusion of up to four covariates in the analysis that themselves could account for about 25% of the variation in the outcome. Without any covariates the MDES is 0.42.

F.2. SOCIAL SKILLS ASSESSMENTS

With a sample size of BASC parental assessments for 85 students (27 kindergarten, 58 prekindergarten), assuming a level of power of at least 0.8, the study is able to detect an effect of 1.0 SD assuming inclusion of up to four covariates in the analysis that themselves could account for about 20% of the variation in the outcome. Without any covariates the MDES is 1.1 SD. Similarly, with a sample size of BASC teacher assessments for 95 students (53 kindergarten, 42 prekindergarten), assuming a level of power of at least 0.8, the study is able to detect an effect of 0.9 SD assuming inclusion of up to four covariates in the analysis that themselves could account for about 20% of the variation in the outcome. Without any covariates the MDES is 1.0 SD. Collectively, this suggests that this study is underpowered to detect effects for the social skills outcome. However, previously stated, this aspect of the study was exploratory only and can therefore provide insights regarding potential areas of future study and exploration.

APPENDIX G

ANALYSIS OF BASC-3 RESPONSES

G.1. DESCRIPTIVE COMPARISON

Table G.1 presents descriptive information on the sample of students assessed with the PPVT-4 and the WJ-IV, and the sample of students with BASC-3 assessments completed by their teachers and parents/legal guardians. Descriptive information on student gender and age when assessed is fairly comparable. However, there are noticeable descriptive differences on student ethnicity and whether the student receives free or reduced priced lunch.

TABLE G.1. COMPARISON OF THE ASSESSED SAMPLE FOR THE WJ-IV AND PPVT-4 TO THE BASC-3

	Sample Assessed (PPVT-4 or WJ-IV only)		Sample Assessed (BASC Teacher only)		Sample Assessed (BASC Parent only)	
	PreK (2015-16)	K (2015-16)	PreK (2015-16)	K (2015-16)	PreK (2015-16)	K (2015-16)
Number in group	333	211	42	53	58	27
Gender						
<i>Female</i>	50.31%	54.93%	52.63%	40.48%	52.63%	46.15%
<i>Male</i>	49.69%	45.07%	47.37%	59.52%	47.37%	53.85%
Ethnicity						
<i>White</i>	26.02%	36.54%	21.05%	23.08%	42.86%	46.15%
<i>African American/Black</i>	31.79%	29.33%	28.95%	33.33%	17.54%	23.08%
<i>Hispanic/Latino</i>	45.03%	42.79%	55.26%	51.28%	45.61%	34.62%
<i>Asian</i>	3.73%	3.37%	0.00%	10.26%	4.44%	7.69%
<i>Other</i>	9.09%	6.25%	6.67%	2.56%	6.67%	3.85%
Lunch						
<i>Free</i>	60.44%	51.94%	65.79%	51.35%	47.37%	32.00%
<i>Reduced</i>	4.97%	5.34%	7.89%	8.11%	1.75%	0.00%
Age Assessed (M/SD)	4.3 (0.54)	5.4 (0.32)	4.4 (0.50)	5.4 (0.30)	4.5 (0.47)	5.5 (0.24)
<i>Note: PreK = Prekindergarten. K = Kindergarten. M=Mean. SD = Standard Deviation. PPVT-4 = Peabody Picture Vocabulary Tests, Fourth Edition. WJ-IV = Woodcock-Johnson, Fourth Edition. BASC-3 = Behavior Assessment System for Children, Second Edition</i>						

APPENDIX H

ADDITIONAL TESTS OF ESTIMATE ROBUSTNESS

A number of elements have been discussed elsewhere that help inform assessments of the robustness of estimates presented in Chapter 4. These include:

- Use of psychometrically sound instruments to measure outcomes.
- Power analysis revealed study adequately powered.
- No statistical differences were identified between the population and the sample assessed.
- No statistical differences were identified between the treatment and comparison group.
- The distribution of birthdates is smooth and continuous around the cut-off.
- A continuous measure of time was added to the vector X of student covariates as a control variable, in both linear and non-linear form, to account for shifts in the data collection timeline.

Included below is an examination of differences in functional form and bandwidth, as well as models estimated. None of the four outcomes is sensitive to differences in functional form or bandwidth. And, only one out of four outcomes is sensitive to differences in model specification.

These two additional examinations, plus those noted elsewhere, lend credibility to the robustness of estimates presented in Chapter 4.

H.1. TESTING ASSUMPTIONS ABOUT FUNCTIONAL FORM AND BANDWIDTH OF ESTIMATES

As shown in Table H.1, none of the four outcomes is sensitive to differences in functional form or bandwidth.

TABLE H.1. TESTING ASSUMPTIONS ABOUT FUNCTIONAL FORM AND BANDWIDTH
AS THEY PERTAIN TO IMPACT ESTIMATES

	(1) Broad Math	(2) Oral Language	(3) Basic Reading	(4) PPVT
Linear, BW=IK	10.152*** (3.252)	5.459 (7.364)	10.828*** (3.079)	4.425 (4.906)
N	273	199	190	251
Up to quadratic, BW=IK	11.080*** (4.069)	12.954 (9.794)	13.125*** (4.221)	6.224 (7.588)
N	273	199	190	251
Up to quintic, BW=365	10.302** (5.217)	23.600** (11.041)	20.585*** (6.743)	15.858* (9.575)
N	425	425	425	418
Linear, BW=180	10.601*** (3.370)	4.842 (5.750)	10.319*** (3.381)	4.696 (5.111)
N	219	219	219	215
Up to quadratic, BW=180	10.878** (4.418)	12.113 (9.524)	13.432*** (4.014)	8.217 (7.929)
N	219	219	219	215
Linear, BW=210	10.242*** (3.099)	4.028 (5.249)	10.352*** (3.152)	4.426 (4.629)
N	255	255	255	251
Up to quadratic, BW=210	11.346*** (4.167)	10.144 (8.969)	12.568*** (3.853)	6.213 (7.586)
N	255	255	255	251

Notes: Heteroskedasticity robust standard errors clustered by birthdate are in parentheses (* $p < .10$ ** $p < .05$ *** $p < .01$). The reduced form estimates reported here were generated using OLS with an indicator for whether a student was eligible for prekindergarten based on their birthdate. Estimates are reported across multiple bandwidths with both quadratic and linear specifications of the forcing variable included at each bandwidth. All models exclude individual-level covariates, but include linear and quadratic terms to account for differences in outcomes that might otherwise be attributed to when a student was assessed.

H.2. CHECKING ROBUSTNESS OF ESTIMATES USING DIFFERENT MODEL SPECIFICATIONS

As an additional check on the robustness of estimates included in Chapter 4, three additional model specifications were executed: a model with student-level covariates added, a model with standard errors clustered at the school-level (as opposed to being clustered by date of birth), and a model with fixed effects for school included. As shown in Table H.2, none of the results is sensitive to differences in model specification, except for Broad Math.

TABLE H.2. TESTING ROBUSTNESS OF IMPACT ESTIMATES USING DIFFERENT SPECIFICATIONS WITH OPTIMAL BANDWIDTH

	(1) Broad Math	(2) Oral Language	(3) Basic Reading	(4) PPVT
Baseline Model	7.198*** (2.534)	3.000 (5.320)	13.028*** (4.753)	1.216 (4.404)
N	299	236	113	262
Adding Covariates	6.618*** (2.361)	1.349 (4.684)	12.658*** (4.410)	0.130 (3.892)
N	299	236	113	262
Cluster SE at School	6.618** (2.631)	1.349 (5.124)	12.658*** (4.027)	0.130 (4.075)
N	299	236	113	262
Add FE for School	12.133** (5.865)	15.674 (14.081)	8.524 (15.916)	9.038 (6.443)
N	299	236	113	262

Notes: Heteroskedasticity robust standard errors are in parentheses (* $p < .10$ ** $p < .05$ *** $p < .01$). The coefficients shown are generated using OLS with the IK bandwidth and a triangular kernel. Row one estimates are the baseline specification with no covariates, and standard errors clustered by date of birth. Row two estimates adds covariates, and with standard errors still clustered by date of birth. Row three estimates includes covariates, but clusters standard errors at the school level. And row four estimates includes covariates and fixed effects for school, while also clustering standard errors at the school level.

Analytically, the equation to estimate row one is the same equation presented in Chapter 4. This equation was modified for row two; specifically, variables representing student gender, ethnicity, and income were added to the vector X of student covariates. To estimate row three, variables representing student demographics were added to the vector X of student covariates, and at the same time, was modified to represent a mean zero student-level error term, clustered on values at the school-level. Finally, to estimate row four, variables representing student demographics were added to the vector X of student covariates and was modified, such that standard errors were clustered at the school level. In addition, was added to the equation as a vector of school fixed effects, essentially restricting all comparisons within site.

MAJOR STUDIES OF THE ACADEMY

2015

- Connecticut Disparity Study: Phase 3
- Addressing Family Violence in Connecticut: Strategies, Tactics and Policies
- Winter Highway Maintenance Operations: Connecticut
- Shared Clean Energy Facilities
- Methods to Measure Phosphorus and Make Future Predictions

2014

- Energy Efficiency and Reliability Solutions for Rail Operations and Facilities
- Connecticut Biomedical Research Program: Analysis of Key Accomplishments
- Peer Review of a CL&P/UConn Report Concerning Emergency Preparedness and Response at Selective Critical Facilities
- Connecticut Disparity Study: Phase 2

2013

- Analyzing the Economic Impacts of Transportation Projects
- Health Impact Assessments Study
- Connecticut Disparity Study: Phase 1
- Connecticut Stem Cell Research Program Accomplishments

2012

- Strategies for Evaluating the Effectiveness of Programs and Resources for Assuring Connecticut's Skilled Workforce Meets the Needs of Business and Industry Today and in the Future
- Benchmarking Connecticut's Transportation Infrastructure Capital Program with Other States
- Alternative Methods for Safety Analysis and Intervention for Contracting Commercial Vehicles and Drivers in Connecticut

2011

- Advances in Nuclear Power Technology
- Guidelines for the Development of a Strategic Plan for Accessibility to and Adoption of Broadband Services in Connecticut

2010

- Environmental Mitigation Alternatives for Transportation Projects in Connecticut
- The Design-Build Contracting Methodology for Transportation Projects: A Review of Practice and Evaluation for Connecticut Applications
- Peer Review of an Evaluation of the Health and Environmental Impacts Associated with Synthetic Turf Playing Fields

2009

- A Study of the Feasibility of Utilizing Waste Heat from Central Electric Power Generating Stations and Potential Applications
- Independent Monitor Report: Implementation of the UCHC Study Recommendations

2008

- Preparing for Connecticut's Energy Future
- Applying Transportation Asset Management in Connecticut
- A Study of Weigh and Inspection Station Technologies
- A Needs-Based Analysis of the University of Connecticut Health Center Facilities Plan

2007

- A Study of the Feasibility of Utilizing Fuel Cells to Generate Power for the New Haven Rail Line
- Guidelines for Developing a Strategic Plan for Connecticut's Stem Cell Research Program

2006

- Energy Alternatives and Conservation
- Evaluating the Impact of Supplementary Science, Technology, Engineering and Mathematics Educational Programs

CONNECTICUT ACADEMY OF SCIENCE AND ENGINEERING
805 Brook Street, Building 4-CERC, Rocky Hill, CT 06067-3405
Phone: 860-571-7143 • e-mail: acad@ctcase.org
web: www.ctcase.org

CONNECTICUT ACADEMY OF SCIENCE AND ENGINEERING

The Connecticut Academy is a non-profit institution patterned after the National Academy of Sciences to identify and study issues and technological advancements that are or should be of concern to the state of Connecticut. It was founded in 1976 by Special Act of the Connecticut General Assembly.

VISION

The Connecticut Academy will foster an environment in Connecticut where scientific and technological creativity can thrive and contribute to Connecticut becoming a leading place in the country to live, work and produce for all its citizens, who will continue to enjoy economic well-being and a high quality of life.

MISSION STATEMENT

The Connecticut Academy will provide expert guidance on science and technology to the people and to the State of Connecticut, and promote its application to human welfare and economic well-being.

GOALS

- Provide information and advice on science and technology to the government, industry and people of Connecticut.
- Initiate activities that foster science and engineering education of the highest quality, and promote interest in science and engineering on the part of the public, especially young people.
- Provide opportunities for both specialized and interdisciplinary discourse among its own members, members of the broader technical community, and the community at large.

CONNECTICUT ACADEMY OF SCIENCE AND ENGINEERING
805 Brook Street, Building 4-CERC, Rocky Hill, CT 06067-3405
Phone: 860-571-7143 • e-mail: acad@ctcase.org
web: www.ctcase.org