The New Mexico PreK Evaluation: Impacts From the Fourth Year (2008-2009) of New Mexico's State-Funded PreK Program

Jason T. Hustedt, Ph.D.
University of Delaware

W. Steven Barnett, Ph.D.
Kwanghee Jung, Ph.D.
National Institute for Early Education Research
Rutgers University, Graduate School of Education

Allison H. Friedman, Ed.M.
New York University

November 2010

Funding for this project was provided by the Children, Youth and Families Department and the Public Education Department through the Legislature of the State of New Mexico. Additional funding was provided by The Pew Charitable Trusts.

Acknowledgements

This research was carried out with funding from the Children, Youth and Families Department and the Public Education Department through the Legislature of the State of New Mexico. Additional funding was provided by The Pew Charitable Trusts. The conclusions presented in this report are those of the authors and do not necessarily reflect the views of our funders.

We extend special thanks to our research partners at New Mexico State University who carried out the fieldwork for this study, including Dr. Candace Kaye, Susana Rios, Heather Castello, and their dedicated data collection staff. We appreciate assistance from Dr. Scott Hughes, Director, Office of Education Accountability, New Mexico Department of Finance and Administration; Judith Paiz and Dan Haggard at the state Children, Youth and Families Department; Ann Zuni and Brenda Kofahl at the state Public Education Department; Claire Dudley at the New Mexico Children's Cabinet; and the many school administrators and teachers who made this research possible. Thanks also to Dr. Ellen Frede and Alexandra Figueras-Daniel at NIEER and Linda Goetze at Utah State University for their ongoing involvement in this work. Finally, we are particularly grateful to the preschoolers and kindergartners who took time out of their school days to participate in our study. Without their assistance it would not have been possible to conduct this research.

The New Mexico PreK Evaluation: Impacts From the Fourth Year (2008-2009) of New Mexico's State-Funded PreK Program

Executive Summary

The New Mexico PreK initiative has expanded quickly since it began in 2005. During this entire period of rapid growth, New Mexico PreK has participated in an evaluation using a methodologically rigorous design known as the regression-discontinuity approach. This fourth in a series of reports focuses on the impacts of New Mexico PreK on children's vocabulary, math, and literacy skills at the beginning of kindergarten. Children's skills in these key content areas were examined in a total sample of 1,359 children from Public Education Department (PED) and Children, Youth and Families Department (CYFD) PreK sites statewide. Our results from the 2008-2009 school year are consistent with previous findings that New Mexico PreK produces positive impacts for young children that are evident at kindergarten entry.

Specifically, at the beginning of kindergarten:

1. Children's vocabulary scores increased by about 5 raw score points as a result of participating in the New Mexico PreK initiative. These gains represent an improvement of about 24% of the standard deviation for the control group and are attributable to PreK. This finding is important because our vocabulary measure is predictive of children's later success at reading and general cognitive abilities.

2. Children's early math scores increased by about 2 raw score points as a result of participating in the New Mexico PreK initiative. These gains represent an improvement of about 37% of the standard deviation for the control group and are attributable to PreK. This reflects greater success in important skills such as addition, subtraction, and telling time.

3. Children's greatest gains were in the area of early literacy. Their early literacy scores increased by about 23 raw score points as a result of participating in the New Mexico PreK initiative. These gains represent an improvement of about 130% of the standard deviation for the control group and are attributable to PreK. This reflects greater knowledge in areas such as print concepts and phonological awareness.

Each of these findings is both statistically significant and practically meaningful, because these are important content areas related to children's success in kindergarten and beyond. As a result, New Mexico PreK is meeting a common objective of state prekindergarten initiatives nationally, in helping prepare young children for later school success.

Introduction

The New Mexico PreK Evaluation began in 2005 as a state-mandated study examining the impacts of New Mexico's newly established state prekindergarten initiative. This multi-year evaluation by the National Institute for Early Education Research (NIEER) at Rutgers University began the same year New Mexico PreK enrolled its first cohort of children. A major goal of our research is to provide data about child outcomes associated with the statewide PreK program for 4-year-olds during its planned expansion. An additional goal is to provide separate estimates for PreK classrooms operated by the state Children, Youth and Families Department (CYFD) and the state Public Education Department (PED).

This is the fourth in an annual series of reports on the effects of the New Mexico PreK initiative on children's vocabulary, math, and literacy skills at kindergarten entry. Data in the current report are specific to the 2008-2009 school year. The report is organized into four major sections:

1. Information about the national context of state-funded prekindergarten programs across the United States and how New Mexico PreK fits within this context;
2. Details about the methodologically rigorous approach used to examine the impacts of New Mexico PreK on the state's 4-year-olds;
3. Data about the vocabulary, mathematics, and early literacy skills of New Mexico PreK participants who entered kindergarten in fall 2009, including separate estimates for children who attended CYFD and PED PreK sites;
4. A discussion of the results and their implications.

Background

National Context of State-Funded Prekindergarten

State-funded education programs for children prior to kindergarten entry have expanded rapidly during the past two decades (Barnett, Epstein, Friedman, Sansanelli, & Hustedt, 2009; Mitchell, 2001). New Mexico is no exception to this nationwide trend. By the 2008-2009 school year, 38 states operated pre-K programs for children at ages 3 and 4, with some states operating multiple distinct pre-K initiatives. Fully a quarter of all 4-year-olds in the United States – and about 17% of New Mexico's 4-year-olds – were enrolled in state pre-K programs. State pre-K represents a substantial commitment of state funds, with state investments of more than $5 billion nationally in 2008-2009 (Barnett, Epstein, et al., 2009) and an investment that year of up to $19.8 million in total state funds specific to New Mexico.

State pre-K initiatives across the United States share a number of common features – they are administered at the state level, offer group learning experiences for children, and are distinct from states' primary programs to serve children with disabilities or who need child care. A common focus of state pre-K initiatives is to prepare children for success in kindergarten, recognizing that effective early childhood programs help them develop the knowledge, skills, and dispositions needed for school success – including rich vocabulary and complex sentence structure, self-regulation, and cooperative play.

At the same time, state policies regarding the availability and areas of emphasis of prekindergarten initiatives tend to vary widely from state to state. Some states have committed to making state pre-K available to all 4-year-olds whose parents would like them to attend – specifically Florida, Georgia, Illinois, Iowa, New York, Oklahoma, and West Virginia (Barnett, Friedman, Hustedt, & Stevenson Boyd, 2009). To date, Oklahoma has come closest to meeting this goal, with 71% of the state's 4-year-olds currently attending state-funded pre-K programs and an additional 16% enrolled in federally funded Head Start programs (Barnett, Epstein, et al., 2009). During the 2008-2009 school year, 12 other states did not offer any form of statewide pre-K initiative: Alaska, Hawaii, Idaho, Indiana, Mississippi, Montana, New Hampshire, North Dakota, Rhode Island, South Dakota, Utah, and Wyoming (although Alaska and Rhode Island developed pilot programs that started the following school year). A number of states that have not yet developed pre-K programs are in the western United States.

Although state pre-K initiatives have been growing rapidly for the past two decades, the body of research on their effectiveness is still relatively small. Among the state pre-K evaluations that do exist, even fewer are methodologically rigorous (Gilliam & Zigler, 2000, 2004). Previous research with model preschool initiatives shows that high-quality and well-funded programs can make valuable contributions to children's learning and development (Barnett, 2002). Studies of well-known initiatives including the High/Scope Perry Preschool program, the Abecedarian Early Childhood Intervention program, and the Chicago Child-Parent Centers show that these types of programs produce economic benefits that greatly outweigh their costs (Barnett, 1996; Masse & Barnett, 2002; Reynolds, Temple, Robertson, & Mann, 2002). Benefits include higher achievement test scores and lower rates of special education placements and grade repetition, as well as long-term effects such as improved high school graduation rates and reduced crime and delinquency rates.

State pre-K programs typically operate on a larger scale than many of the model programs that have been intensively studied, and are not as well funded. However, a number of states now offer prekindergarten programs that are both high in quality and widely available. Results from a widely cited and methodologically rigorous study of state pre-K participants in Tulsa, Oklahoma (Gormley, Gayer, Phillips, & Dawson, 2005; Gormley, Phillips, & Gayer, 2008) show that high-quality state pre-K programs can generate statistically significant short-term impacts on a range of academic content areas, for children across a range of racial and ethnic groups, and for children from both low- and middle-income backgrounds.

State-Funded Prekindergarten in New Mexico

New Mexico PreK began serving children during the 2005-2006 school year. This statewide initiative offers voluntary center-based prekindergarten to 4-year-olds across New Mexico. Funds to operate PreK classrooms are provided by the state through CYFD and PED; in order to receive these funds, sites submit proposals that are "… evaluated on the percentage and number of public elementary schools in the community that are not meeting the proficiency component required for calculating adequate yearly progress and that are serving children, at least sixty-six percent of whom live within the attendance zone of a Title 1 elementary school" (Pre-Kindergarten Act, NMSA 1978 § 32A-23-6, 2005). Additional criteria used to prioritize

sites for state funding include the number of 4-year-olds in the community who will be served and the adequacy of prekindergarten sites already operating in that community.

Providers that participate in the New Mexico PreK initiative include both public schools and private providers such as Head Start, municipal and community child care facilities, faith-based centers, universities, and family child care homes (Barnett, Epstein, et al., 2009). PreK program sites are administered by either PED or CYFD, depending on which agency provided funding, and overall administrative responsibility for the statewide PreK initiative is shared by PED and CYFD. Key program standards that PreK sites must follow are outlined in Table 1 below.

Table 1. Program Standards for New Mexico PreK During the 2008-2009 School Year

State Pre-K Policy Area	*New Mexico PreK Policy*
Early learning standards	Programs use the full version of the New Mexico Early Learning Outcomes. These are comprehensive standards.
Lead teacher degree	Lead teachers must have a bachelor's degree within five years of being hired. The minimum requirement for a starting teacher is a high school diploma.
Lead teacher specialized training	Lead teachers must have a license in Early Childhood Education (Birth – Grade 3) within five years of being hired.
Assistant teacher degree	Assistant teachers must have an associate's degree in Early Childhood Education within five years of being hired. The minimum requirement for a starting assistant teacher is a high school diploma.
Teacher in-service	In-service training of 45 clock hours/year is required for lead teachers.
Maximum class size	Class sizes are capped at 20 students.
Staff-child ratio	A 1:10 ratio is required.
Required screening/referral and support services	Services include screening and referral in vision, hearing, health, dental, and developmental areas. Four annual parent conferences or home visits are offered, as well as parent education or job training, parenting support or training, parent involvement activities, health services for children, information about nutrition, referral to social services, and transition to kindergarten activities.
Meals	At least one meal (breakfast or lunch) is provided daily.
Monitoring	Site visits and other monitoring activities are conducted.

Source: Barnett, Epstein, et al. (2009)

Enrollment in the New Mexico PreK initiative increased each year between 2005 and 2009. In general, though, states in the Western and Northern Plains regions of the United States have tended to lag behind other parts of the country in making state-funded prekindergarten available to young children (Barnett, Hustedt, Hawkinson, & Robin, 2006). Among the 13 states designated as being in the West by the U.S. Census, only seven offered state prekindergarten during the 2008-2009 school year (Barnett, Epstein et al., 2009). Within these regions, New Mexico is now second only to Colorado (17% vs. 20%) in terms of the percentage of 4-year-olds in each state who are enrolled in state prekindergarten. Table 2 provides more information about annual funding and enrollment levels for the New Mexico PreK initiative.

Table 2. Funding and Enrollment Levels for New Mexico PreK Since 2005

School Year	State Appropriation	Children Budgeted (Number of 4-Year-Olds)
2005-2006	$4,950,000	1,540
2006-2007	$7,990,000	2,194
2007-2008	$13,998,886	3,570
2008-2009	$19,290,300	4,745
2009-2010	$19,842,400	4,963
2010-2011	$15,331,380	4,435

This annual report on the impacts of New Mexico PreK on young children's language, literacy, and mathematics skills builds upon three previous reports using a common methodology (Hustedt, Barnett, & Jung, 2007; Hustedt, Barnett, Jung, & Figueras, 2008; Hustedt, Barnett, Jung, & Figueras-Daniel, 2009). Each of these previous reports presents statistically significant and meaningful impacts of the PreK initiative on skills important to children's school readiness. Again, the current report focuses on the effects of New Mexico PreK programs on children who attended PreK during the 2008-2009 school year, and then entered kindergarten in fall 2009.

Methods

The Regression-Discontinuity Approach

Given the interest in ensuring that scarce state resources are invested effectively, it is important that evaluations of state prekindergarten programs use methodologically rigorous designs. A common evaluation approach involves estimating the effects of an initiative by comparing test scores of children who attended the initiative with the scores of similar children who did not. However, there would be two major problems related to selection bias if this approach was used in the context of evaluating large-scale public prekindergarten initiatives. First, as these types of programs become more widely available, it becomes increasingly difficult to find a comparable group of children who did not attend. Second, even where programs target only a subset of children (such as those from low-income families or Title I schools), the children who attend preschool are different from children who do not. Preschool programs contribute to these differences by targeting different groups of children, but differences also come about because only some eligible families choose to enroll their children. In sum, selection bias can be a problem because programs select children, and also because families select programs. This can lead to differences between preschool participants and non-participants,

ultimately making it difficult to determine the impacts specifically associated with their preschool participation.

The research model for the New Mexico PreK Evaluation employs a sophisticated regression-discontinuity design (RDD), which improves validity by reducing selection bias. More specifically, our solution to the issue of selection bias involves comparing two groups of children, where both groups of children have enrolled in New Mexico PreK. The RDD comparisons rely upon New Mexico's stringent age cut-off for enrollment eligibility (August 31) to define the two groups.

This concept is easier to understand by providing an extreme example: consider two children who differ only in that one was born the day before the age cut-off and the other the day after. When both children are about to reach their fifth birthdays, the slightly younger child will enter PreK. The slightly older child will enter kindergarten having already completed PreK. If both children are tested at the beginning of the school year, the difference in their scores provides an unbiased estimate of effect of PreK. If only children with birthdays one day on either side of the age cut-off were included in a study, the sample size would be unreasonably small. However, the RDD approach can be applied to wider age ranges around the cut-off.

In fact, all children entering kindergarten having just completed New Mexico PreK, and all children beginning New Mexico PreK the same year, can be included in our RDD analyses. The RDD approach has been employed in a growing body of research examining the effects of state-funded pre-K programs in Oklahoma (Gormley et al., 2005), New Jersey (Frede, Jung, Barnett, Lamy, & Figueras, 2007), Arkansas (Hustedt, Barnett, Jung, & Thomas, 2007), and several other states (Wong, Cook, Barnett, & Jung, 2008).

The New Mexico PreK Evaluation has used an RDD approach each fall since the initial cohort of PreK children entered kindergarten in 2006. This report presents results from a fourth set of RDD analyses, and relates to the 2008-2009 school year. Appendix A offers more details about the how the RDD approach has been employed in this study.

Sampling Strategy

In fall 2009, we began our sample selection process by first identifying all sites that had New Mexico PreK programs in place during the 2008-2009 school year and that were continuing operations into the following school year. (In communities offering PreK for the first time in fall 2009, we did not expect to find many kindergartners who had previously completed PreK, and this is a requirement of our analysis.) At each identified PreK site we randomly selected a pre-specified number of children to participate in our research, based on the proportion of New Mexico PreK children statewide who were enrolled at that particular site. PreK children were then selected from class enrollment lists using a procedure to ensure randomness, and were assessed at their PreK site until we had completed the designated number of assessments.

We also chose a corresponding number of kindergartners for each of the selected fall 2009 PreK sites. We identified kindergartners who had participated in New Mexico PreK the previous year using the state's master PreK enrollment list, and randomly selected children from

that list. Children were tracked to their current elementary schools using information about their anticipated kindergarten destinations collected at the end of the previous school year by PED and CYFD, and compiled by the state Office of Education Accountability. Current kindergarten students were then assessed at their elementary schools.

All child assessments were conducted by New Mexico-based data collectors who received training from NIEER and then worked under the day-to-day supervision of researchers from New Mexico State University. These child assessors visited each sampled PreK site as well as kindergarten sites where former PreK participants had enrolled. Research staff conducted child assessments as early as possible during the school year.

The Fall 2009 Sample

The RDD approach allows us to compare two groups of children who participated in New Mexico PreK:

1. The *Preschool* group, or experimental group, is made up of kindergartners who completed New Mexico PreK the previous school year (2008-2009). Our Preschool group does not include children who participated in other forms of early care or education at age 4, rather than enrolling in New Mexico PreK.

2. The *No Preschool* group, or control group, is made up of New Mexico PreK students at the outset of their PreK year. This group is referred to as the No Preschool group even though these children were currently enrolled in PreK during 2009-2010 – at the time of our child assessments they had just begun the school year and had not experienced the preschool "treatment" yet.

By comparing these groups based on data collected in fall 2009, we are able to estimate the effects of attending PreK during the 2008-2009 school year. In fall 2009, the No Preschool group included 706 children from New Mexico PreK classrooms across the state. The Preschool group included 653 children from kindergarten classrooms across the state. The total New Mexico sample size was 1,359 children.

The total fall 2009 sample was 48.7% female. Children's home languages were: English only, or English plus another language, 85.2%; Spanish only, 14.0%; and other languages, 0.8%. The percentage of children in each ethnic category was: Hispanic, 63.8%; White, 20.9%; Native American, 10.7%; Black, 2.2%; Asian, 0.9%; and Other, 1.5%.

Ethnicities of participants in our study generally reflect those of the population of children who attended the New Mexico PreK program. During the 2009-2010 school year, the percentage of all New Mexico PreK children in each ethnic category was: Hispanic, 62.8%; Caucasian, 20.6%; American Indian and Alaska Native, 13.3%; Black, 2.0%; Asian, 1.2%; and Unknown, 0.1%.

Ethnicities of participants in our study also generally reflect those of the population of children in the state of New Mexico. For comparison purposes, New Mexico-specific estimates

from the U.S. Census (Bureau of Business and Economic Research, 2010) show that the percentages of New Mexico children ages birth to 5 in each ethnic category were as follows for 2009: Hispanic, all races, 56.8%; White, 26.4%; American Indian and Alaska Native, 10.3%; two or more races, non-Hispanic, 2.5%; Black, 2.5%; and Asian, 1.4%.

Research Instruments

In the New Mexico PreK Evaluation, children's readiness for kindergarten is measured in three key content areas – vocabulary, mathematics, and early literacy skills. To reflect the diversity of New Mexico PreK participants, all measures selected for the evaluation were chosen so that child assessments could be conducted in either English or Spanish. Details about our standardized child assessment instruments, and protocols used to determine the language of assessment, are provided in the following sections.

Also, recognizing that some New Mexico PreK participants may need assistance with the cultural context of the assessment instruments, we developed an additional protocol so that a "cultural broker" could be present for children who might have difficulties with the instruments due to cultural differences. Teachers were asked to identify children who needed this type of assistance and to identify a cultural broker who could be present during testing. We requested that the cultural broker be someone conversant with the child's own culture. Ideally, this individual would be a school employee familiar to the child.

As fall 2009 marked the beginning of a new evaluation cycle, we made some changes to the child assessment battery. Differences between the battery of instruments used during this data collection cycle and the battery of instruments used previously are highlighted below.

Receptive Vocabulary. As in previous years of the study, children's receptive vocabulary was measured using the Peabody Picture Vocabulary Test, 3rd Edition (PPVT-III; Dunn & Dunn, 1997) and the Test de Vocabulario en Imagenes Peabody (TVIP; Dunn, Padilla, Lugo, & Dunn, 1986) for Spanish-speakers. The PPVT is predictive of general cognitive abilities and is a direct measure of vocabulary size. The rank order of item difficulties is highly correlated with the frequency with which words are used in spoken and written language. This test is adaptive (to avoid floor and ceiling problems), establishing a floor below which the child is assumed to know all the answers and a ceiling above which the child is assumed to know none of the answers. Reliability is good as judged by either split-half or test-retest reliabilities. The TVIP is appropriate for measuring growth in Spanish vocabulary for bilingual students and for monolingual Spanish speakers. The results of these tests are found to be strongly correlated with school success. Raw scores are used in this study.

All children in the New Mexico PreK Evaluation were initially administered the PPVT, regardless of their home language, to get a sense of their receptive vocabulary skills in English. Children who spoke some Spanish were also subsequently administered the TVIP. The testing session was then continued, with all additional measures administered in either English or Spanish, depending upon which language the child's teacher designated as his or her best testing language.

Table 3. Statistical Description of the Sample by Group, for Children Entering PreK and Children Entering Kindergarten from PreK in Fall 2009

	Entering PreK	Entering Kindergarten
Number in group	706	653
Girls (%)	52.0	46.0
Ethnicity (%)		
White	21.5	20.2
Hispanic	61.8	66.0
Native American	10.8	10.7
Other/missing	6.0	3.1
Home Language (%)		
English, or English + another language	84.8	85.6
Spanish only	14.3	13.6
Other	0.8	0.8
Assessment conducted only in English (%)	86.1	89.7
Age (in months) when assessed (Mean/*SD*)	54.86	67.40
	3.47	3.49

Note: *SD* = Standard Deviation

Math Skills. Also consistent with previous years of the study, children's early math skills were measured using the Woodcock-Johnson Tests of Achievement, 3rd Edition (WJ-III; Woodcock, McGrew & Mather, 2001) Applied Problems subtest. For children whose best testing language was Spanish, the Bateria III Woodcock-Munoz (Woodcock, Munoz-Sandoval, McGrew, & Mather, 2005) Problemas Aplicados subtest was used. Subtests of the Woodcock-Johnson are reported to have good reliability. Raw scores are used in this study.

Early Literacy. Children's early literacy skills were measured with the Early Literacy Skills Assessment (ELSA; DeBruin-Parecki, 2005). The ELSA is a child assessment that measures four key principles of early literacy – comprehension, phonological awareness, alphabetic principle, and concepts about print. It has 23 items and appears to be a children's storybook. There are two ELSA protocols that are both available in Spanish and English. An

extensive investigation of the ELSA's psychometric properties conducted by an outside evaluator found good reliability and validity (Cheadle, 2007). The ELSA provides a broader range of information about early literacy than the Test of Preschool Early Literacy (TOPEL; Lonigan, Wagner, Torgesen, & Rashotte, 2007), which was used during the previous evaluation cycle for New Mexico PreK. Also, the ELSA components on Concepts About Print are comparable to the Print Knowledge components on the TOPEL. Raw scores are used in this study.

Results

In Tables 3 and 4, we provide a statistical description of the fall 2009 sample by group (entering preschool and entering kindergarten). As shown in Table 3, the groups are demographically similar. As shown in Table 4, the test scores are different between the two groups, as would be expected, though care must be taken not to interpret the simple differences between the groups' test scores as an estimate of the PreK program's impact.

Data analysis for the New Mexico PreK Evaluation was conducted in STATA (StataCorp, 2005) using raw scores for the child assessment instruments. All standard errors were clustered by classroom. As in previous years of the study, covariates were included in our statistical models to control for gender, ethnicity, and whether assessments were administered in English or Spanish. We also included a new covariate to control for the length of time between the beginning of the school year at each school, and the date the child assessment was conducted.

In our RDD analyses, there is no *a priori* expectation that the estimated relationships between PreK participation and child outcomes should be linear. Therefore, we estimated linear as well as higher order polynomial forms of the regression equation as a check against mis-specifying the functional form of the regression line. We conducted squared and cubic transformations of the selection variable (the difference between birth date and cut-off date) and its interaction with the cut-off variable.

Table 4. Assessment Scores by Group, for Children Entering PreK and Children Entering Kindergarten from PreK in Fall 2009

Measure	Entering PreK		Entering Kindergarten		
	Mean	*SD*	Mean	*SD*	Sample Size
PPVT raw scores	45.64	21.71	64.52	22.08	1,356
WJ-III raw scores	10.02	4.28	15.46	4.40	1,358
ELSA total raw scores	34.44	17.83	71.27	23.86	1,358

Note: *SD* = Standard Deviation

Table 5 shows linear, quadratic, and cubic estimates for each of our child outcome measures, using fall 2009 data. We found that linear models generally provided the best estimates of relationships between participating in the New Mexico PreK initiative and children's scores on each the three measures. The only exception was that the cubic model provided the best estimate for the Comprehension raw score of the ELSA. For the remainder of this report, we will focus on the linear estimates when reporting fall 2009 data for the PPVT-III, WJ-III, and all subscales of the ELSA except Comprehension. For the Comprehension subscale of the ELSA, we will focus only on the cubic estimate.

Also, in presenting RDD findings we prefer to emphasize the results using one year as the margin around the kindergarten cut-off date, as this allows us to use the largest sample size. However, we also conduct separate linear regressions restricting the sample to children born within 3- and 6-month spans before and after the cut-off date. Restricting the sample to observations closest to the cut-point should reduce any potential bias, though the smaller sample sizes also increase the standard errors. As shown in Table 5, linear estimates using 3-, 6-, and 12-month margins are similar for this sample of New Mexico PreK participants.

The primary RDD analyses used in this report are intent to treat (ITT) estimates, which allow us to examine the effect of the PreK and kindergarten entry policy as actually implemented, using our entire sample. However, we also conducted an additional set of analyses with treatment on treated (TOT) estimates. The TOT estimates exclude children whose birthdates are inconsistent with birth-date cut-off requirements for their PreK or kindergarten programs, and are the types of estimates that we emphasized in our previous reports. The two sets of results were very similar. Here, we focus on the ITT results because they are most conservative, and best address policymakers' interest in testing the effects of policies as actually implemented (Wong et al., 2008).

Child Outcomes from the Fourth Year of New Mexico PreK

The effects of New Mexico PreK on children's receptive vocabulary, mathematics, and early literacy skills are summarized below, and are also shown in graphical form in Appendix A.

The estimated effect of state-funded preschool on children's receptive vocabulary was statistically significant ($p < .05$). Attending New Mexico PreK during the 2008-2009 school year was estimated to increase PPVT scores by about 5.24 raw score points at kindergarten entry. This represents an improvement of about 24% of the standard deviation for the control (No Preschool) group.

The estimated effect of New Mexico PreK on children's early math skills was statistically significant for the 2008-2009 school year ($p < .05$). The increase in Woodcock-Johnson-III Applied Problems subtest scores for New Mexico PreK children is about 1.58 raw score points. This represents an improvement of about 37% of the standard deviation for the control group.

Table 5. Estimated Effects Based on Functional Form and Margin Around Kindergarten Cut-Off Date in Fall 2009

	Parametric models used in analysis				
	Linear, 12 months	Quadratic, 12 months	Cubic, 12 months	Linear truncated at 3 months	Linear truncated at 6 months
Receptive Vocabulary	**5.24*** **(2.07)**	6.86* (2.67)	6.61* (2.65)	7.62+ (4.07)	5.41+ (2.98)
Math	**1.58*** **(0.49)**	2.21* (0.59)	2.10* (0.61)	2.76* (0.96)	1.80* (0.66)
Comprehension	0.13 (0.46)	0.12 (0.54)	**0.17** **(0.53)**	0.45 (0.80)	-0.16 (0.62)
Phonological Awareness	**2.77*** **(0.54)**	3.07* (0.65)	2.95* (0.64)	3.99* (0.97)	2.47* (0.74)
Alphabetic Principle	**18.29*** **(1.81)**	20.69* (2.27)	20.01* (2.25)	21.17* (3.67)	18.76* (2.67)
Concepts About Print	**2.01*** **(0.38)**	1.95* (0.48)	1.85* (0.46)	2.22 (0.68)	1.43* (0.50)
ELSA Total	**23.19*** (2.40)	25.83* (3.00)	24.98* (2.91)	27.82* (4.74)	22.50* (3.40)

Note: Receptive vocabulary data represent PPVT raw score point increases. Math data represent WJ-III Applied Problems subtest raw score point increases. Comprehension, Phonological Awareness, Alphabetic Principle, and Concepts About Print represent raw score point increases on the ELSA subtests. Raw score point increases are also used for the ELSA total score.
Robust standard errors are shown in parentheses. Boldfaced terms show the selected estimates.
+ $p < .10$. * $p < .05$.

The estimated effect of the New Mexico PreK initiative on children's overall early literacy scores was also statistically significant for the 2008-2009 school year ($p < .05$). In addition to having statistically significant impacts for the total ELSA score, participation in New Mexico PreK was also associated with statistically significant impacts on three of four ELSA subtests – Phonological Awareness, Alphabetic Principle, and Concepts About Print ($p < .05$). The impact of PreK on children's Comprehension scores was not statistically significant.

The effect of the New Mexico PreK initiative on children's gains in overall early literacy scores is about 23.19 raw score points. This increase represents approximately 130% of the standard deviation for the control group.

Receptive Vocabulary, Math, and Early Literacy Results for CYFD and PED

A parallel set of analyses was conducted to separately examine the impacts of participating in New Mexico PreK sites offered through PED and those offered through CYFD. These analyses were based on the subgroups of 696 children who attended PED sites and 663 children who attended CYFD sites.

Before presenting the results of these analyses, it is important to emphasize that PED and CYFD PreK sites appear to serve different populations of children. There were statistically significant differences between the PED and CYFD subsamples in terms of both children's ethnicities ($p < .001$) and home languages ($p < .01$). These differences were found among both the Preschool and No Preschool groups. Related to the finding for home languages, there were also statistically significant differences between the PED and CYFD subsamples in terms of the language used for child assessments in this study ($p < .001$ for the Preschool Group and $p < .05$ for the No Preschool Group). A number of additional statistically significant differences were found between children tested in the Preschool Group compared to the No Preschool group within CYFD sites. More results from these demographic analyses are provided in Appendix B.

Table 6. Estimated Effects Disaggregated for PED and CYFD Sites and Overall, Fall 2009

Measure	PED	CYFD	All NM PreK
Receptive Vocabulary			
Linear	3.62	6.80*	5.24*
Math			
Linear	1.83*	1.20	1.58*
Comprehension			
Linear	0.38	-0.004	0.13
Phonological Awareness			
Linear	1.92*	3.58*	2.77*
Alphabetic Principle			
Linear	15.53*	21.20*	18.29*
Concepts About Print			
Linear	2.33*	1.59*	2.01*
ELSA Total			
Linear	20.17*	26.37*	23.19*

Note: Receptive vocabulary data represent PPVT raw score point increases. Math data represent WJ-III Applied Problems subtest raw score point increases. Comprehension, Phonological Awareness, Alphabetic Principle, and Concepts About Print represent raw score point increases on the ELSA subtests. Raw score point increases are also used for the ELSA total score.

* $p < .05$.

The results of the additional RDD analyses based on PED and CYFD subgroups, and comparisons with analyses from the overall New Mexico PreK sample, are shown in Table 6. Like the New Mexico PreK initiative as a whole, CYFD and PED programs produced statistically significant impacts on measures of children's early literacy skills for PreK participants who entered kindergarten in fall 2009. The only component of early literacy in which statistically significant results were not found was Comprehension, but this finding was also consistent across PED and CYFD. On our measure of receptive vocabulary, statistically significant results were found for children in CYFD programs and for the overall New Mexico PreK sample, but not for children in PED programs. On our measure of mathematics, statistically significant results were found for children in PED programs and for the overall New Mexico PreK sample, but not for children in CYFD programs. All findings shown in Table 6 are the result of linear regressions with a 12-month margin around the kindergarten cut-off date.

Discussion

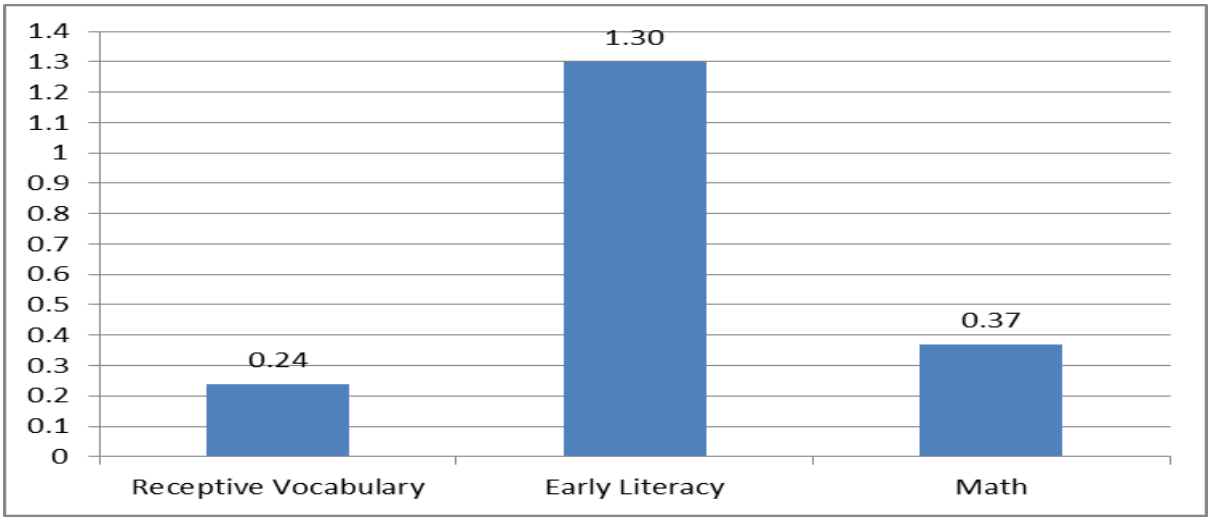
Overall, our results show that participation in New Mexico PreK is associated with a number of benefits for 4-year-old children. Specifically, the PreK initiative produced positive impacts on young children's language, mathematics, and early literacy skills as measured at kindergarten entry. These findings are not just statistically significant, but are also practically meaningful because these are important content areas related to children's success in kindergarten and beyond. As a result, New Mexico PreK is meeting a common objective of state prekindergarten initiatives nationally, in preparing young children for later school success.

Although the data in this report are specific to the New Mexico PreK initiative as implemented during the 2008-2009 school year, they are also consistent with our findings from several years of prior research on the New Mexico PreK initiative. An earlier analysis of data from the first three years of New Mexico PreK (Hustedt, Barnett, Jung, & Goetze, 2009) also showed positive impacts on PreK participants' language, mathematics, and early literacy skills at the beginning of kindergarten.

In Figure 1, we provide information that facilitates the comparison of data across the different academic content areas assessed for the 2008-2009 school year. These findings are presented as effect sizes, which use a common metric to standardize the estimated effects of New Mexico PreK across different types of measures. Overall, the PreK program had its greatest impact on early literacy measures – though, again, we found statistically significant results in each of the three major content areas assessed in this study.

The effect sizes reported below for New Mexico PreK are similar in magnitude to those reported for other well-regarded state prekindergarten initiatives, including the widely researched initiative in Oklahoma (e.g., Gormley et al., 2005). Unlike New Mexico, though, the Oklahoma study and other prekindergarten evaluations conducted by NIEER have focused on well-established prekindergarten initiatives. As we have noted in previous reports, our research in New Mexico offers a relatively unique opportunity to document progress by a prekindergarten initiative during an initial period of large-scale expansion across the state. With this opportunity comes a responsibility to interpret early results cautiously.

Figure 1. The Effect of New Mexico PreK on Children's Scores across Measures



New Mexico PreK is still a new initiative, and was growing at a fast pace during the time period of this study. As the program matures, the statewide context in which it operates may begin to change. For example, educational requirements for lead and assistant teachers in PreK classrooms come with a 5-year phase-in period. When data in this report were collected, New Mexico PreK had not yet existed for five years, so the effects of those requirements remain to be seen. Also, PreK sites were being prioritized for state funding based on proximity to Title I schools as well as schools' annual progress on No Child Left Behind (Barnett, Epstein, et al., 2009; Pre-Kindergarten Act, NMSA 1978 § 32A-23-6, 2005). These priorities have led to a current emphasis on serving disadvantaged children and communities, but with further expansion, New Mexico PreK would be more accessible to populations of children who are less disadvantaged.

As part of our two most recent data collection cycles, we collected additional data in CYFD and PED classrooms in an effort to better understand the separate impacts of PreK classrooms administered in each of these types of settings. Data from the 2007-2008 school year (Hustedt, Barnett, Jung, & Figueras-Daniel, 2009) showed consistent results across CYFD and PED programs, also reflecting the pattern of results for the entire sample. Data from the 2008-2009 school year are more difficult to interpret. Although CYFD and PED programs produced similar results for children's early literacy, only CYFD programs produced statistically significant results for receptive vocabulary, and only PED programs produced statistically significant impacts for mathematics. Again, it is important to point out that PED and CYFD PreK programs appear to serve different populations of children. Those demographic differences may contribute to the differences between impacts produced by PED and CYFD programs during the 2008-2009 school year. Additional data would be needed to clarify the inconsistent findings between the 2007-2008 and 2008-2009 school years.

In our fall 2009 final report from the previous evaluation cycle (Hustedt, Barnett, Jung, & Goetze, 2009), we recommended further expansion of New Mexico PreK on the basis of strong results in terms of children's language, literacy, and early math skills. The child outcomes for

the 2008-2009 school year, presented in the current report, are very similar to those used as the basis for our earlier recommendation of expansion. As a result, the overall body of research supporting this recommendation is further strengthened by our most recent data. During the past five years, New Mexico has begun to assume a role of leadership in the western United States by making steady progress toward a widely available prekindergarten initiative in a region of the country where enrollment in such programs has been low. However, many new opportunities to expand the statewide reach of New Mexico PreK remain, since this initiative still serves fewer than 1 in 5 of the state's 4-year-olds.

References

Barnett, W. S. (1996). *Lives in the balance: Age 27 benefit-cost analysis of the High/Scope Perry Preschool Program*. Ypsilanti, MI: High/Scope Press.

Barnett, W. S. (2002). Early childhood education. In A. Molnar (Ed.), *School reform proposals: The research evidence* (pp. 1-26). Greenwich, CT: Information Age Publishing, Inc.

Barnett, W. S., Epstein, D. J., Friedman, A. H., Sansanelli, R. A., & Hustedt, J. T. (2009). *The state of preschool 2009: State preschool yearbook*. New Brunswick, NJ: National Institute for Early Education Research, Rutgers University.

Barnett, W. S., Friedman, A. H., Hustedt, J. T., & Stevenson Boyd, J. (2009). An overview of prekindergarten policy in the United States: Program governance, eligibility, standards, and finance. In R. C. Pianta & C. Howes (Eds.), *The Promise of Pre-K* (pp. 3-30). Baltimore, MD: Brookes Publishing.

Barnett, W. S., Hustedt, J. T., Hawkinson, L. E., & Robin, K. B. (2006). *The state of preschool 2006: State preschool yearbook*. New Brunswick, NJ: National Institute for Early Education Research.

Bureau of Business and Economic Research, University of New Mexico. (2010). *State and county population estimates by age, sex, race and Hispanic origin from the Census Bureau*. Retrieved October 20, 2010, from http://www.unm.edu/~bber/demo/coestchar.htm

Cheadle, J. E. (2007). *The Early Literacy Skills Assessment (ELSA) psychometric report for both English and Spanish versions*. Ypsilanti, MI: High/Scope Early Childhood Reading Institute.

Cleveland, W. S., & Devlin, S. J. (1988). Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association, 83,* 596-610.

DeBruin-Parecki, A. (2005). *Early Literacy Skills Assessment (ELSA).* Ypsilanti, MI: High/Scope Educational Research Foundation.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test-Third Edition (PPVT-III).* Circle Pines, MN: AGS Publishing.

Dunn L. M., Padilla, E. R., Lugo, D. E., & Dunn, L. M. (1986). *Test de Vocabulario en Imágenes Peabody (TVIP)*. Circle Pines, MN: AGS Publishing.

Frede, E., Jung, K., Barnett, W. S., Lamy, C. E., & Figueras, A. (2007). *The Abbott Preschool Program Longitudinal Effects Study (APPLES)*. New Brunswick, NJ: National Institute for Early Education Research.

Gilliam W. S., & Zigler, E. F. (2000). A critical meta-analysis of all evaluations of state-funded preschool from 1977 to 1998: Implications for policy, service delivery and program evaluation. *Early Childhood Research Quarterly, 15,* 441-473.

Gilliam, W., & Zigler, E. (2004). *State efforts to evaluate the effects of prekindergarten: 1977 to 2003*. Unpublished manuscript: Yale University Child Study Center.

Gormley, W.T., Jr., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal pre-K on cognitive development. *Developmental Psychology, 41*, 872-884.

Gormley, W.T., Jr., Phillips, D., & Gayer, T. (2008). Preschool programs can boost school readiness. *Science, 320,* 1723-1724.

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, Rutgers University.

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, Rutgers University

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.

Lonigan, C. J., Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (2007). *Test of Preschool Early Literacy (TOPEL)*. Austin, TX: Pro-Ed.

Masse, L. N., & Barnett, W. S. (2002). *A benefit-cost analysis of the Abecedarian Early Childhood Intervention*. New Brunswick, NJ: National Institute for Early Education Research.

Mitchell, A.W. (2001). *Education for all young children: The role of states and the federal government in promoting prekindergarten and kindergarten*. FCD Working Paper. New York: Foundation for Child Development.

Pre-Kindergarten Act, NMSA 1978 § 32A-23 (2005).

Reynolds, A. J., Temple, J.A., Robertson, D.L., & Mann, E.A. (2002). *Age 21 cost-benefit analysis of the Title I Chicago Child-Parent Centers.* (Discussion Paper no. 1245-02). Madison, WI: Institute for Research on Poverty. Available at http://www.irp.wisc.edu/publications/dps/pdfs/dp124502.pdf.

StataCorp. (2005). *Stata Statistical Software: Release 9*. College Station, TX: StataCorp LP.

Trochim, W. M. K. (1984). *Research design for program evaluation*. Beverly Hills, CA: Sage Publications.

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*(1), 122-154.

Woodcock, R. W., Munoz-Sandoval, A. F., McGrew, K. S., & Mather, N. (2005). *Bateria III Woodcock-Munoz.* Itasca, IL: Riverside Publishing.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson Tests of Achievement.* Itasca, IL: Riverside Publishing.

Appendix A: Applying the RDD Approach to the New Mexico PreK Study

The regression-discontinuity design (RDD) takes advantage of a strict kindergarten enrollment policy that determines enrollment using the child's date of birth to define the groups. By relying on this assignment rule, one that is unlikely to be related to child and family characteristics, the RDD seeks to reduce the likelihood of selection bias. Thus, rather than comparing children who attended and did not attend the New Mexico PreK initiative (raising concerns that the same child and family factors that led program eligibility or a family seeking to enroll a child in the program also contribute to differences in learning and development), the RDD approach compares two groups of children who enroll in New Mexico PreK. One group has completed PreK and the other is just entering.

One way to interpret the RDD approach is to view it as similar to a randomized trial for children near the age cut-off. The RDD creates groups that *at the margin* differ only in that some were born a few days before the age cut-off and others a few days after the cut-off. When these children are about to turn 5 years old, the slightly younger children will enter PreK and the slightly older children will enter kindergarten having already attended PreK. If all of the children are tested at that time, the difference in their scores can provide an unbiased estimate of the effect of the preschool initiative under reasonable circumstances. However, if the sample was restricted to children with birthdays near the age eligibility cut-off, the total sample size would be too small.

Alternatively, the RDD can be viewed as modeling the relationship between an assignment variable (age) and measures of children's learning and development. The pre-cut-off sample is used to model the relationship prior to treatment. The post-cut-off sample is used to model the relationship after the treatment. This approach can be applied to wider age ranges around the cut-off, though its validity depends on correctly modeling the relationship. As there is always some uncertainty about what this looks like (is it linear, and if not what does the curve look like?), we test a variety of models (different functional forms for the equation) to see which model best fits the data, in addition to conducting other tests of the RDD assumptions. Under either view, it is important that there is minimal misallocation (exceptions to the rule) around the age cut-off.

To identify the proper functional form for our RDD analyses, we conducted a graphical analysis and a series of parametric regressions using alternate specifications. We begin with graphs for each child outcome measure, shown in Figures A1-A3. Two types of lines are fitted onto scatterplots on each side of the cut-offs. The first plot in each figure depicts a linear regression line, and the second shows a non-parametric regression line based on locally weighted smoothing, called Lowess. This strategy can be useful for data exploration because it relaxes assumptions about the form of the relationship between the assignment and outcome (Cleveland & Devlin, 1988). For each y_i, we obtain a smoothed value through weighted regressions involving only observations within a local interval, with observations closer to y_i weighted most heavily.

Each plot in Figures A1- A3 shows an estimated regression line for children's predicted test scores by age, measured by the number of days their birth date is from the program enrollment cut-off date. The section of the line to the left of the enrollment cut-off date (shown as "0" in the figures) represents scores of children beginning the state pre-K program, while the section of line to the right of the enrollment cut-off date represents scores for children entering

kindergarten. The discontinuity in the regression line at the cut-off date shows the estimated effect of New Mexico PreK.

Next, we run a series of regressions to obtain parametric estimates of the treatment effect. In order to describe the effect of PreK participation on child outcomes, we model children's vocabulary, math, and early literacy scores. For the *i*th child in classroom *j*, the relevant equation is:

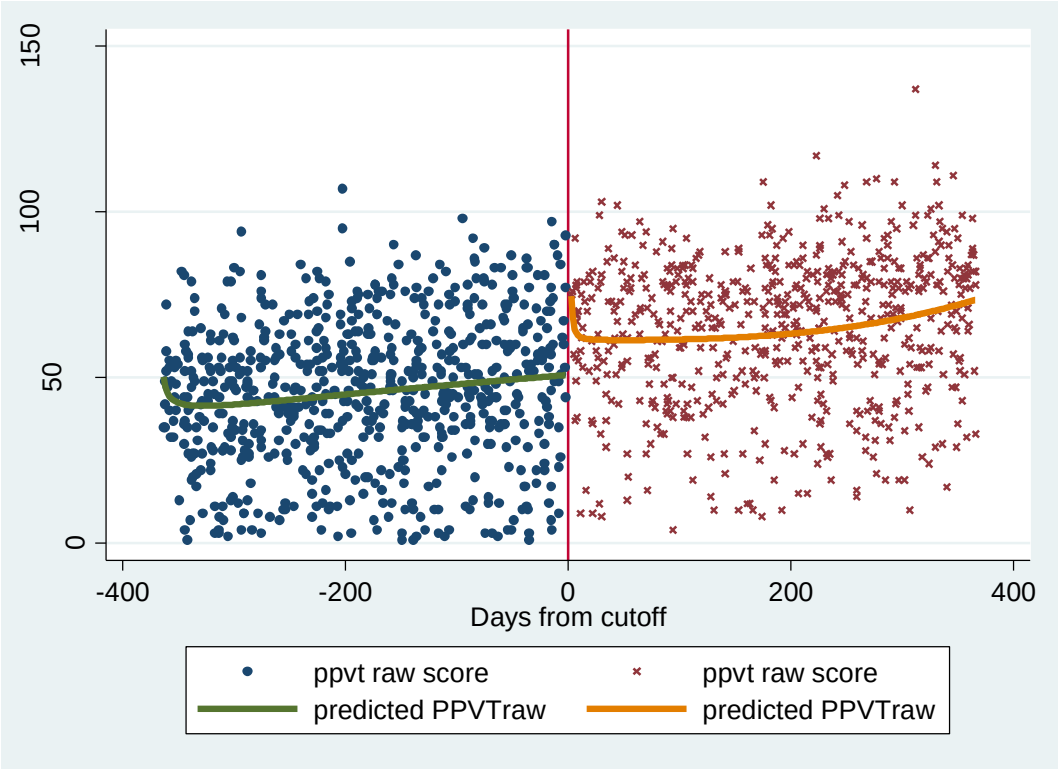
$$Y_{ij} = a + BX_{ij} + \beta_1(\text{Pre-K})_{ij} + g(AV)_{ij} + \varepsilon_i$$

In this equation, Y_{ij} is child *i*'s outcome, X_{ij} is a vector of child characteristics, Pre-K_{ij} is a dichotomous indicator variable such that T=1 for the PreK "treatment" and T=0 for no treatment, and $g(AV)_{ij}$ is a smooth function of the continuous assignment variable. We check the robustness of our estimates by considering alternative specifications for $g(AV)_{ij}$, including polynomials and interaction terms. We determine the order of the polynomial approximation to the $g(AV)_{ij}$ function by examining the statistical significance of the higher order and interaction terms.

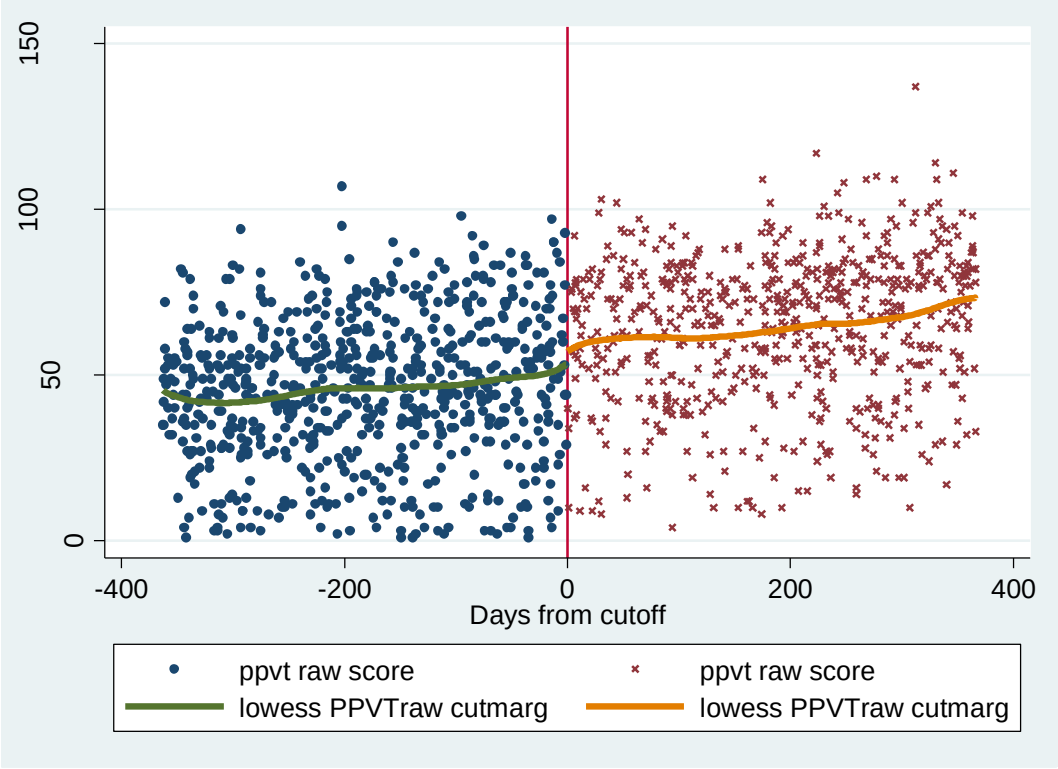
Following Trochim (1984), when the functional form of the regression model is ambiguous, we overfit the model by including more polynomial and interaction terms than needed, yielding unbiased but less efficient estimates. In all the parametric analyses we use Huber-White standard errors adjusted for clustered data at the classroom level. As a final parametric check on functional form, we truncate the dataset to include only observations near the cut-off. In placing greater weight on these observations, we eliminate the influence of extreme assignment variable values that often play a disproportionate role in mis-specifying functional form. We rerun the parametric analyses including only children who have birthdates within 6 and then 3 months on either side of the enrollment cut-off.

Graphical and parametric analyses provide evidence that the response function was linear for receptive vocabulary, quadratic for early literacy outcomes, and cubic for mathematics outcomes. Estimates were robust with respect to narrowing the time window around the birthdate cut-off to 6 and 3 months.

Figure A1. Linear and Lowess Plots of PPVT Receptive Vocabulary Results

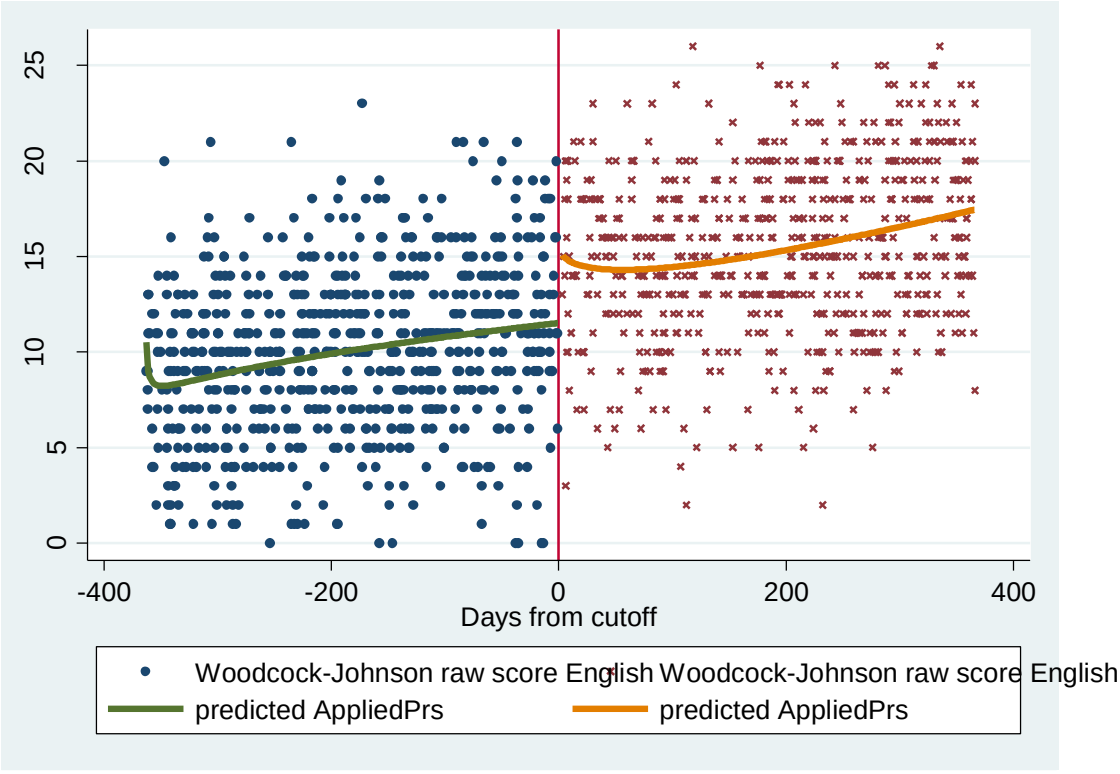


(1) Linear Plot

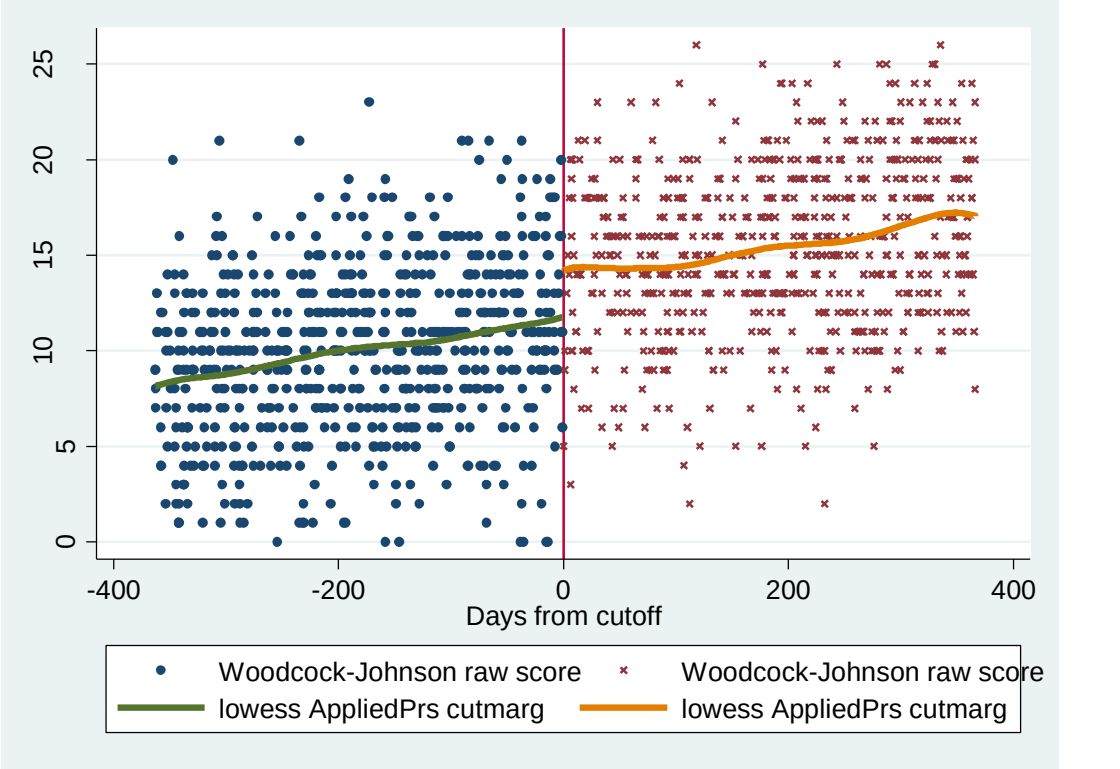


(2) Lowess Plot

Figure A2. Linear and Lowess Plots of WJ-III Applied Problems Results

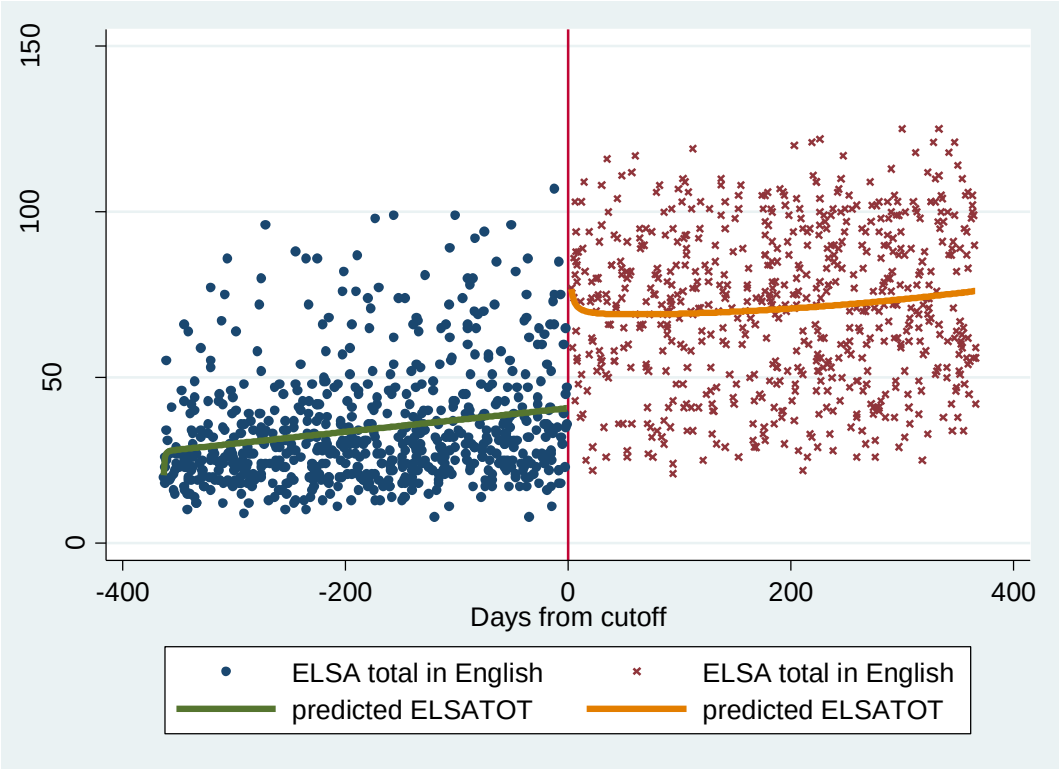


(1) Linear Plot

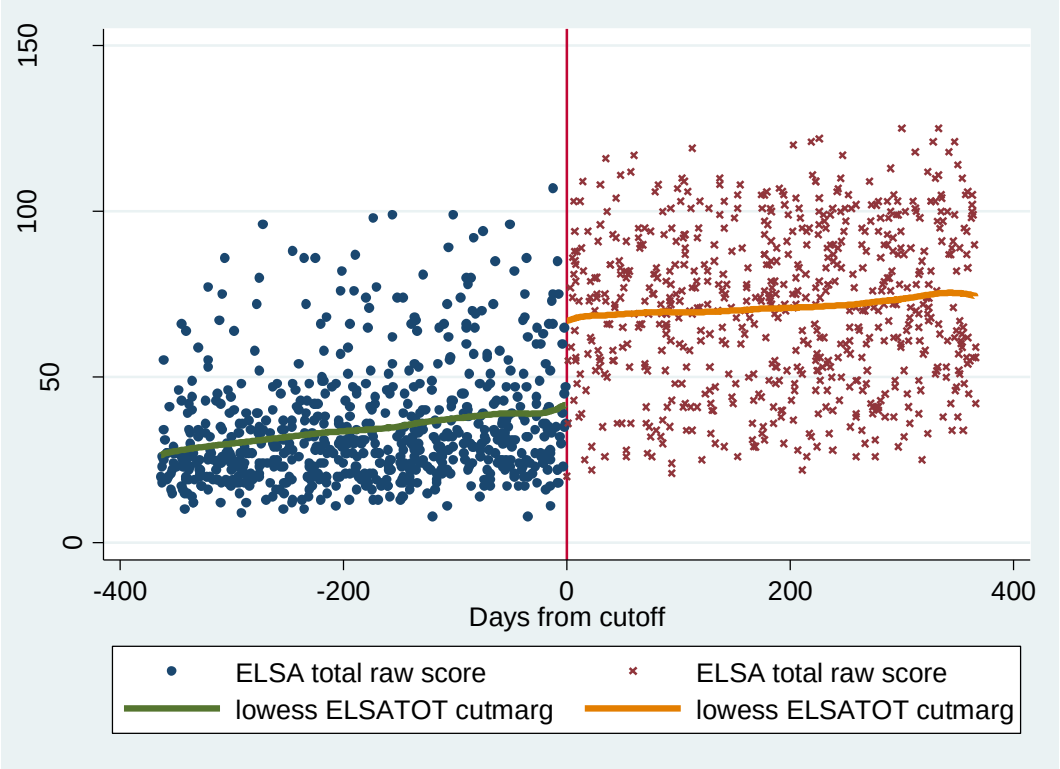


(2) Lowess Plot

Figure A3. Linear and Lowess Plots for ELSA Results



(1) Linear Plot



(2) Lowess Plot

Appendix B: Demographic Comparisons for PED and CYFD Children

Below we present demographic comparisons of children attending PED PreK sites and CYFD PreK sites. Table B1 focuses on comparing PED with CYFD children at PreK entry and again at kindergarten entry. Statistically significant differences between PED and CYFD are noted. Table B2 focuses on comparisons of children at PreK entry and kindergarten entry *within* the individual PED and CYFD subgroups. Statistically significant differences between the groups of children at PreK entry and the groups of children at kindergarten entry (within either the PED or CYFD subgroup) are noted.

Table B1. Comparisons between PED and CYFD at PreK Entry and Again at Kindergarten Entry

	Entering PreK		Entering Kindergarten	
	PED	CYFD	PED	CYFD
Number in group	359	347	337	316
Girls (%)	48.7	54.5	49.0	42.1
Ethnicity (%)***				
White	17.5	25.6	11.9	29.1
Hispanic	65.5	57.9	70.0	61.7
Native American	15.3	6.1	16.3	4.7
Other/missing	1.7	10.4	1.8	4.4
Home Language (%)**				
English, or English + another language	83.0	87.0	80.4	90.8
Spanish only	17.0	11.2	18.7	8.5
Other	0	1.7	0.9	0.6
Assessment conducted only in English (%)	83.0	89.6*	83.1	96.5***
Age (in months) when assessed (Mean/*SD*)	54.88	54.89	67.49	67.23
	3.45	3.57	3.62	3.49

Note: *SD* = Standard Deviation
*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

Table B2. Comparisons Between PreK and K Entry, Separately for PED and CYFD

	Entering PreK		Entering Kindergarten	
	PED	CYFD	PED	CYFD
Number in group	359	347	337	316
Girls (%)	48.7	54.5[a]	49.0	42.1[a]
Ethnicity (%)				
White	17.5	25.6	11.9	29.1
Hispanic	65.5	57.9	70.0	61.7
Native American	15.3	6.1	16.3	4.7
Other/missing	1.7	10.4	1.8	4.4
Home Language (%)				
English, or English + another language	83.0	87.0[a]	80.4	90.8[a]
Spanish only	17.0	11.2[a]	18.7	8.5[a]
Other	0	1.7[a]	0.9	0.6[a]
Assessment conducted only in English (%)	83.0	89.6[a]	83.1	96.5[a]
Age (in months) when assessed (Mean/*SD*)	54.88[a]	54.89[b]	67.49[a]	67.23[b]
	3.45	3.57	3.62	3.49

Note: *SD* = Standard Deviation
Comparisons with the same superscript are significantly different from each other ($p < .05$).