

**Impacts of New Mexico PreK
on Children's School Readiness
at Kindergarten Entry:**

**Results from the Second Year
of a Growing Initiative**

**Jason T. Hustedt, Ph.D.
W. Steven Barnett, Ph.D.
Kwanghee Jung, Ph.D.
Alexandra Figueras, M.A.**

**The National Institute for Early Education Research
Rutgers University**

June 2008

Funding for this project was provided by the Department of Finance and Administration through the Legislature of the State of New Mexico, and The Pew Charitable Trusts.

Many thanks to our colleagues at Utah State University's Early Intervention Research Institute, who partnered with us to carry out this study, including Dr. Linda Goetze; Albuquerque-based logistics coordinators Marissa Nordstrom Perez and CJ Ondek; and their field staff. We would also like to thank Dr. Scott Hughes and Dr. Peter Winograd at the Office of Education Accountability, New Mexico Department of Finance and Administration; Judith Paiz at the state Children, Youth and Families Department; Ann Zuni and Richard LaPan at the state Public Education Department; Julienne Smrcka at the New Mexico Children's Cabinet; Dr. Kurt Steinhaus, formerly of the Office of the Governor; and the many school administrators and teachers who made this research possible. Finally, we are especially grateful to the preschoolers and kindergartners who took time out of their school days to participate in our study.

**Impacts of New Mexico PreK on Children's
School Readiness at Kindergarten Entry:
Results from the Second Year of a Growing Initiative**

Executive Summary

This is the second in a series of reports that estimates of impacts of participating in the New Mexico PreK initiative at age 4. We used a rigorous research design, the regression-discontinuity approach, to assess the academic skills of kindergartners who had enrolled in New Mexico PreK during the 2006-2007 school year. Children's receptive vocabulary, math, and early literacy skills were examined in a sample of 924 PreK participants from across the state. Findings show that New Mexico PreK continued to produce positive effects on children's learning during its second year of operations, consistent with findings from a similar study of the initiative's first year.

Specifically, by the beginning of kindergarten:

1. Children's vocabulary scores increased by about 6 raw score points due to their participation in the New Mexico PreK initiative. This represents an improvement of about 25% of the standard deviation of the control group, attributable to PreK. As our measure of vocabulary is strongly predictive of general cognitive abilities and later reading success, this finding is particularly important.
2. Children's early math scores increased by more than 2 raw score points due to their participation in New Mexico PreK. This represents an improvement of about 50% of the standard deviation of the control group, attributable to PreK. Early math skills assessed include simple addition and subtraction, basic number concepts, telling time, and counting money.
3. Children's scores on our measure of early literacy rose by about 14 percentage points. This represents an improvement of about 59% of the standard deviation of the control group, attributable to PreK. Children who attended New Mexico PreK know more letters, more letter-sound associations, and are more familiar with words and book concepts by the time they enter kindergarten.

Findings on these measures of vocabulary, math, and early literacy were both statistically significant and meaningful. Separate analyses were conducted to examine the impacts of attending PreK programs offered by the Public Education Department and those offered by the Children, Youth and Families Department. Although these analyses using disaggregated samples did not yield statistically significant results in all cases, this is likely due to the smaller sample sizes involved. In order to better examine the individual impacts of PED and CYFD PreK programs, and to further document the effects of New Mexico PreK as it grows, we will gather additional data in Fall 2008.

Introduction

In recent years, state prekindergarten initiatives have expanded rapidly, with total enrollments now topping one million children across the United States (Barnett, Hustedt, Friedman, Stevenson Boyd, & Ainsworth, 2007). A common goal of state pre-K programs is to prepare young children for kindergarten, recognizing that effective preschool programs help children to develop the knowledge, skills, and dispositions they need for success in school, such as rich vocabulary and complex sentence structure, an understanding of story structure, self-regulation, cooperative play, and abstract thinking. However, policies, areas of emphasis, and spending levels of state-funded pre-K programs vary widely from state to state. Some states offer pre-K to all children whose parents wish them to enroll, while others focus on offering pre-K to children meeting specific criteria, such as family incomes below a certain level. Also, some state pre-K initiatives require teachers to have training comparable to teachers in the K-12 system, while others require less in the way of formal training. Given the tremendous growth in state-funded pre-K programs, and the high degree of variability across different states, it is important to examine the effectiveness of these types of programs in improving children's potential for school success.

While state-funded pre-K has not been extensively studied, previous research with model initiatives establishes that high-quality and well-funded preschool programs can make valuable contributions to improving 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 scores on achievement tests 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. In general, though, state-financed preschool programs are not as well funded as many model programs that have been intensively studied, and the state programs are larger and serve more diverse populations.

New Mexico is one of 38 states currently offering state-funded pre-K (Barnett et al., 2007). Established in 2005, New Mexico PreK is a statewide initiative offering voluntary center-based prekindergarten to 4-year-old children. Potential New Mexico PreK sites submit proposals for funding 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). In prioritizing sites for funding, additional criteria specified in state statute include the adequacy of prekindergarten sites that already exist in a community, and the number of 4-year-olds in a community to be served by the proposed New Mexico PreK site.

Administrative responsibility for New Mexico PreK is shared by the state Children, Youth and Families Department (CYFD) and the state Public Education Department (PED). As a result, funding and enrollment for New Mexico PreK are split across programs administered by CYFD and programs administered by PED. Participating providers include public schools as well as private centers including Head Start, child care facilities, faith-based centers, and tribal programs. New Mexico PreK classrooms feature maximum class sizes of 20 with staff-child ratios of 1:10, and offer a variety of comprehensive and family support services in addition to their emphasis on early childhood education. Standards requiring all lead teachers to have bachelor's degrees and licensure in early childhood education are being phased in over time, such that teachers must meet this requirement within 5 years of their New Mexico PreK site being established (Barnett et al., 2007). Thus, lead teachers at sites that began operating during the initial year of the New Mexico PreK initiative must have bachelor's degrees and early childhood licensure by Fall 2010.

Each year since New Mexico PreK began, the state has increased both appropriation levels and total enrollments for state-funded prekindergarten. During its initial year of funding, the 2005-2006 school year, the appropriation was \$4.95 million. A total of 1,538 4-year-olds were enrolled. The appropriation grew to \$7.99 million for the 2006-2007 school year, with 2,195 4-year-olds enrolled. By the 2007-2008 school year, the appropriation of state funds grew to \$14 million, with an enrollment of 3,570 4-year-olds. The budget for the 2008-09 school year calls for a further increase in New Mexico PreK funding, to \$19.39 million, with an anticipated enrollment level of 4,867 4-year-olds. This trend of expansion is expected to continue in the upcoming years.

The National Institute for Early Education Research began an in-depth 4-year study of the New Mexico PreK initiative during the 2005-2006 school year, with funding from the state Office of Education Accountability. A primary goal of this research is to examine the impacts of New Mexico PreK on young children's receptive vocabulary, early math, and early literacy skills, as the initiative carries out a planned expansion. In an earlier report (Hustedt, Barnett, & Jung, 2007), we presented results from the first year of the new initiative. These findings are extremely positive, with statistically significant and meaningful impacts in each of the measured content areas by the time 4-year-olds from the initial year of New Mexico PreK entered kindergarten in Fall 2006. The current report uses a similar methodology to examine the impacts of participating in the second year of the initiative, among children who attended PreK as 4-year-olds during the 2006-2007 school year and entered kindergarten in Fall 2007.

Methods

The Research Model

Our ongoing evaluation of the New Mexico PreK initiative is based upon a regression-discontinuity design (RDD), which seeks to reduce selection bias. In state prekindergarten evaluations, the effects of an initiative are often estimated by comparing

test scores of children who attended the pre-K program with the scores of similar children who did not. However, as programs become more widely available, it is more difficult to find a comparable group of children who did not attend. Even where programs target only a subset of children (such as those from low-income families), a problem remains: children who attend preschool are different from children who do not. Preschool programs that target specific groups of children create these differences, but differences also come about because only some parents choose to enroll their children. In sum, children who attend state prekindergarten programs differ from those who do not because programs select children and families select programs.

Our solution to the problem of selection bias is to compare two groups of children who enrolled in New Mexico PreK. The RDD comparisons rely upon the state'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 are about to turn 5 years old the slightly younger child will enter PreK and the slightly older child will enter kindergarten having already completed PreK. If both are tested at that time, 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 approach can be applied to wider age ranges around the cut-off. In fact, all children entering kindergarten having completed the New Mexico PreK initiative, and all children beginning New Mexico PreK the same year, can be included in RDD analyses. The RDD approach has been used recently in a growing body of research examining the effects of state-funded pre-K programs in Oklahoma (Gormley, Gayer, Phillips, & Dawson, 2005), New Jersey (Frede, Jung, Barnett, Lamy, & Figueras, 2007), Arkansas (Hustedt, Barnett, Jung, & Thomas, 2007), and a number of other states (Wong, Cook, Barnett, & Jung, 2008). Additional information about the how the RDD approach was used in New Mexico is provided in Appendix A.

In Fall 2006, we implemented the RDD approach in New Mexico for the first time, addressing the research question of whether participating in the New Mexico PreK initiative at age 4 has an impact on children's academic skills at kindergarten entry (Hustedt, Barnett, & Jung, 2007). We conducted an additional RDD study in Fall 2007, and describe the results in the current report. The RDD methodology will be used for a third and final time as part of this project in Fall 2008. By conducting separate RDD studies during three consecutive school years, we will be able to document the estimated effects of the New Mexico PreK initiative as it begins to mature.

Sampling Strategy

When selecting our sample for Fall 2007, we first identified all CYFD and PED sites that were operating New Mexico PreK programs by the beginning of the 2006-2007 school year and that continued operations during the 2007-2008 school year. (We would not expect to locate sufficient numbers of kindergartners who had already completed New Mexico PreK in communities offering the initiative for the first time in 2007-2008.)

At each identified PreK site we randomly selected a pre-specified number of children to participate in our study, based on the proportion of New Mexico PreK enrollees statewide enrolled at that particular site. Individuals were next selected from PreK 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.

In Fall 2007, we also chose a corresponding number of kindergartners for each prekindergarten site. We identified kindergartners who had participated in New Mexico PreK the previous year using the state's 2006-2007 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.

New Mexico-based research staff—trained by NIEER and working under the supervision of an Albuquerque-based coordinator and Utah State University—visited each sampled New Mexico PreK site as well as the identified kindergarten sites in the same communities. Research staff conducted child assessments as early as possible in the school year.

The Sample

Our RDD methodology relies upon two groups of children, as mentioned previously. The group of kindergartners who attended the New Mexico PreK initiative the previous year (2006-2007) is called the *Preschool* group, or the experimental group. Children who received some form of early care or education other than the New Mexico PreK initiative at age 4 were not included in this group. The second group of children is called the *No Preschool* group, or the control group. This group is referred to as the No Preschool group despite the fact that these children were enrolled in PreK during the 2007-2008 school year, because at the time of the assessments they were just beginning their preschool year and had not had the preschool “treatment” yet.

In New Mexico, the No Preschool group included 519 children enrolled in 131 New Mexico PreK classrooms across the state. The Preschool group included 405 children enrolled in 214 kindergarten classrooms across the state. The total New Mexico sample size for Fall 2007 was 924 children.

Our sample for the 2007-2008 school year is 53.8% female. Children's home languages are: English, 69.6%; Spanish, 13.8%; both English and Spanish, 7.6%; both English and Navajo, 5.4%; and Other, 3.6%. The percentage of children in each ethnic category is as follows: Latino, 57.3%; Native American, 19.5%; White, 18.6%; African American, 1.5%; Asian, 1.5%; and Other, 0.2%.

Ethnicities of participants in our study generally reflect those of the population of children participating in the New Mexico PreK program. For the 2007-2008 school year,

the percentage of all New Mexico PreK children in each ethnic category is as follows: Hispanic, 63.0%; American Indian and Alaska Native, 15.7%; Caucasian, 18.0%; Black, 2.2%; and Asian, 1.1%.

For purposes of comparison, New Mexico-specific estimates from the U.S. Census (Bureau of Business and Economic Research, 2007) show that the percentage of New Mexico children ages birth to 5 in each ethnic category is as follows for 2006: Hispanic, all races, 54.3%; White, 30.3%; American Indian and Alaska Native, 10.6%; Two or more races, Non-Hispanic, 2.0%; Black, 1.7%; and Asian, 1.1%.

Instrumentation

Child outcome measures in the New Mexico PreK study focused on receptive vocabulary, mathematics, and early literacy skills. Each of the measures selected for this study allowed child assessments to be conducted in either English or Spanish. Specific details about the measures and our protocols for determining the language used for each child assessment are provided below. Also, recognizing that some children in New Mexico may need assistance with the cultural context of the standardized assessment instruments, we developed an additional protocol so that a “cultural broker” could be present for children who might have difficulties with the instruments as a result of 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 familiar with the child’s own culture, and ideally a school employee familiar to the child.

Receptive Vocabulary. Once again during the 2007-2008 school year, children’s receptive vocabulary was measured using the Peabody Picture Vocabulary Test, 3rd Edition (PPVT-III; Dunn & Dunn, 1997) and for Spanish-speakers, the *Test de Vocabulario en Imagenes Peabody* (TVIP; Dunn, Padilla, Lugo, & Dunn, 1986). 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 reported here.

All children in this study 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.

Mathematical Skills. Children's early mathematical skills were again measured with the Woodcock-Johnson Tests of Achievement, 3rd Edition (WJ-III; Woodcock, McGrew & Mather, 2001) Subtest 10 Applied Problems. For children whose best testing language was Spanish, the *Bateria Woodcock-Munoz Pruebas de Aprovechamiento – Revisado* (Woodcock & Munoz, 1990) *Prueba 25 Problemas Aplicados* was used. Subtests of the Woodcock-Johnson are reported to have good reliability. Raw scores are reported here.

Early Literacy. Print knowledge was measured using the Print Knowledge subtest of the Test of Preschool Early Literacy (TOPEL; Lonigan, Wagner, Torgesen, & Rashotte, 2007) during the 2007-2008 school year. The TOPEL is the published, normed version of the unpublished Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP; Lonigan, Wagner, Torgesen & Rashotte, 2002), which was used in the New Mexico PreK study during the 2006-2007 school year. As the TOPEL has not yet been published in Spanish, in 2007-2008 we continued using the Spanish version of the Pre-CTOPPP with children whose best testing language was Spanish.

The TOPEL has been used with both middle-class and low-income samples, and subtests are reported to have good to excellent reliability. Print Knowledge items measure whether children recognize individual letters and letter-sound correspondences, and whether they differentiate words in print from pictures and other symbols. Percentages of items answered correctly out of the total 36 Print Knowledge subtest items are reported here.

Results

A statistical description of the sample by group (entering preschool and entering kindergarten) is provided in Tables 1 and 2. As can be seen in Table 1, the groups are generally similar in their demographic characteristics, indicating that the sampling approach was successful in obtaining comparable groups. The only statistically significant difference between the groups was that cultural brokers were more frequently present for children entering PreK than for children entering kindergarten ($p < .05$). Differences in children's gender and home languages approached but did not reach statistical significance. All subsequent analyses were statistically controlled for whether a cultural broker was present. Additional covariates were used in our statistical models to control for time of assessment, gender, ethnicity, and whether the assessment instruments were administered in English or Spanish. The test scores are different between the two groups, as would be expected (see Table 2). However, care must be taken not to interpret the simple differences between the groups' test scores as an estimate of the PreK program's impact. Differences in standard scores are age-adjusted, and their relationship to our RDD estimates of program effects is discussed below.

Our RDD analyses were conducted in STATA (StataCorp, 2005) using raw scores, and all standard errors were clustered by classroom. In these RDD analyses, there is no *a priori* expectation that the estimated relationship between PreK participation and child outcomes should be linear. Therefore, in addition to using linear models, we also

estimated higher order polynomial forms of the regression equation. This provides 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. We found that the linear model provided the best estimate of the relationship between participating in New Mexico PreK and children’s scores on the PPVT, our measure of vocabulary skills. For our measure of early literacy, the TOPEL, the second-order (quadratic) model provided the best estimate. For our measure of math skills, the WJ-III, the third-order (cubic) model provided the best estimate. Linear, quadratic, and cubic estimates for each measure are shown in Table 3. For the remainder of this report, we will focus on the linear estimate for the PPVT-III, the quadratic estimate for the TOPEL, and the cubic estimate for the WJ-III.

Table 1. Statistical Description of the Sample by Group, for Children Entering PreK and Children Entering Kindergarten from PreK

	Entering	
	PreK	Kindergarten
Number in group	519	405
Girls (%)	56.5	50.4
Ethnicity (%)		
White	19.8	17.0
Latino	55.5	59.5
Native American	20.4	18.3
Other/missing	3.8	5.2
Cultural broker present (%)	3.6	2.2
Home Language		
English only	73	67.9
Spanish only	11.4	17.0
English/Spanish	7.9	7.4
English/Native American Language	5.8	6.9
Assessment conducted only in English (%)	82.3	80.2
Age (in months) when assessed (Mean/ <i>SD</i>)		
	55.18	67.68
	4.34	3.78

Note. *SD* = Standard Deviation

Table 2. Assessment Scores by Group, for Children Entering PreK and Children Entering Kindergarten from PreK

Measure	PreK		Entering Kindergarten		Sample Size
	Mean	<i>SD</i>	Mean	<i>SD</i>	
PPVT raw scores	46.11	22.43	61.03	22.63	921
PPVT standard scores	87.77	21.01	88.80	19.82	920
WJ-III raw scores	10.20	4.56	14.88	4.71	923
TOPEL % correct	32.88	23.59	69.98	25.74	920

Note: *SD* = Standard Deviation

In reporting the RDD findings, we prefer to emphasize the results using one year as the margin around the kindergarten cut-off date, because this allows for the largest sample size. However, we also conducted 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 increase the standard errors. As shown in Table 3, linear estimates using 3-, 6-, and 12-month margins are similar. The primary analyses presented in the remainder of this report include children with birthdays up to 12 months before and 12 months after the kindergarten cut-off date.

Our primary analyses were “sharp” regression-discontinuity models that included 893 children in our sample, dropping 31 children (3.4% of the total) whose birth-date information appears to be inconsistent with the birth-date cut-off requirement for their PreK or kindergarten programs. When less than 5% of the sample is dropped in this way, exclusion of such cases is thought to have little effect on the result (Judd & Kenny, 1981; Shadish, Cook, & Campbell, 2002; Trochim, 1984). As an additional check to confirm that the findings were similar, we also conducted instrumental variable analyses with all 924 children. Results of these analyses were indeed similar, as shown in Table 4.

Table 3. Estimated Effects Based on Functional Form and Margin Around Kindergarten Cut-Off Date

	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	5.55* (2.35)	8.26* (3.30)	13.19* (3.83)	12.08* (4.14)	10.26* (3.04)
Math	1.88* (0.56)	1.99* (0.81)	2.26* (0.94)	2.61* (1.07)	2.34* (0.78)
Early Literacy	21.36* (3.32)	13.66* (4.40)	13.36* (5.55)	15.15* (6.05)	17.39* (4.47)

Note: Receptive vocabulary data represent PPVT raw score point increases. Math data represent WJ-III Applied Problems subtest raw score point increases. Early literacy data represent increases in percentage correct on the TOPEL Print Knowledge subtest.

Robust standard errors are shown in parentheses. Boldfaced terms show the selected estimates.

* Significant at $p < .05$.

Child Outcomes at Kindergarten Entry

Estimated effects of the New Mexico PreK initiative on children’s receptive vocabulary, mathematics, and early literacy skills are summarized in this section using results from the “sharp” RDD models (see Table 3). The effects of the New Mexico PreK program on children’s receptive vocabulary, mathematics, and early literacy are shown in graphical form in Appendix A.

The estimated effect of state-funded preschool on children’s receptive vocabulary as measured by the PPVT is statistically significant ($p < .05$). Attending the New Mexico PreK initiative at age 4 is estimated to increase PPVT scores by about 5.55 raw score points. This represents an improvement of about 25% of the standard deviation for the control (No Preschool) group.

The estimated effect of state-funded preschool on children’s early math skills as measured by the Woodcock-Johnson-III Applied Problems subtest scores also is statistically significant for the New Mexico PreK initiative ($p < .05$). The increase in scores for New Mexico PreK children due to the program is about 2.26 raw score points. This represents an improvement of about 50% of the standard deviation for the control (No Preschool) group.

Table 4. Estimated Effects Based on Functional Form and Margin Around Kindergarten Cut-Off Date, Using Instrumental Variable Analyses

	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	6.31* (2.54)	7.94* (3.51)	11.39* (4.08)	15.82* (6.24)	11.57* (4.03)
Math	1.74* (0.57)	1.51* (0.73)	1.62* (0.79)	2.20+ (1.20)	1.92* (0.84)
Early Literacy	21.94* (3.42)	16.66* (4.29)	15.83* (4.40)	11.02 (6.92)	14.52* (4.92)

Note: Receptive vocabulary data represent PPVT raw score point increases. Math data represent WJ-III Applied Problems subtest raw score point increases. Early literacy data represent increases in percentage correct on the TOPEL Print Knowledge subtest.

Robust standard errors are shown in parentheses. Boldfaced terms show the selected estimates.

* Significant at $p < .05$.

Finally, the effect of state-funded preschool on children’s Print Knowledge scores is statistically significant for New Mexico PreK ($p < .01$). The effect of the New Mexico PreK initiative on children’s gains in Print Knowledge scores is 14% more items answered correctly. This increase represents approximately 59% of the control (No Preschool) group standard deviation on the Print Knowledge subtest.

The RDD estimated effects on raw scores can be roughly translated into changes in standard scores. Thus, a raw score change of 5 points translates into a standard score change of 2 to 4 points on the PPVT, and a raw score change of 13 points translates into a standard score change of about 10 points, depending on the age of the child. It is useful to compare the RDD estimates to the actual means for the preschool-entry and kindergarten-entry samples. This comparison shows an increase of only about 1 point on the PPVT. The implication is that without the preschool program, standard scores would actually have declined, that is, preschool-eligible children actually fall further behind other children if they do not attend the state prekindergarten program. Although this is plausible, such changes are not expected to be large. This reinforces the notion that the linear estimates in the “sharp” RDD are the most appropriate. In addition, the fact remains that children who attended PreK score had an average raw score of only 89 on the PPVT, which corresponds to the 23rd percentile. Thus, there is substantial scope to further increase the effectiveness of the New Mexico PreK initiative.

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

Although this study was initially designed with the goal of estimating the effects of the entire New Mexico PreK initiative (as a single entity), we also conducted additional RDD analyses to separately examine the impacts of participating in PreK sites offered through PED and those offered through CYFD. These analyses were based on the subgroups of 466 children who attended New Mexico PreK at PED sites and 427 children who attended New Mexico PreK at CYFD sites. The results of those analyses, and comparisons with analyses from the overall New Mexico PreK sample are shown in Table 5.

Table 5. Estimated Effects, Disaggregated for PED and CYFD Sites and Overall

Measure	PED	CYFD	Overall sample
Receptive Vocabulary			
Linear	7.63*	3.20	5.55*
Math			
Cubic	1.11	2.66 ⁺	2.26*
Linear	1.76**	1.83 ⁺	1.88**
Early Literacy			
Quadratic	12.41 ⁺	11.46 ⁺	13.66**
Linear	22.03***	20.17***	21.36***

Note: Receptive vocabulary data represent PPVT raw score point increases. Math data represent WJ-III Applied Problems subtest raw score point increases. Early literacy data represent increases in percentage correct on the TOPEL Print Knowledge subtest.

Boldfaced terms show the selected estimates.

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

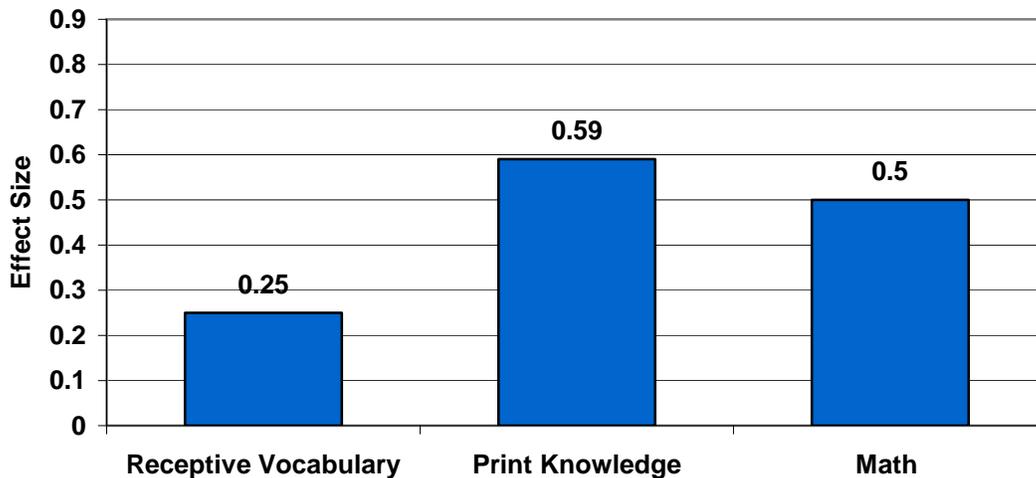
As previously reported, the New Mexico PreK initiative as a whole produced statistically significant impacts on measures of children's receptive vocabulary, mathematics, and early literacy at kindergarten entry. It should be noted that the separate analyses for CYFD and PED are based on smaller samples, and thus have less power to detect a statistical effect. As a result, the models based on data from the overall sample provide more precise estimates and a better fit to the data. Data from the overall sample may also take a different functional form than data restricted to the CYFD or PED subgroups. Due to these potential differences, in Table 5 we provide results from linear models and for the appropriate polynomial models selected for use with the overall sample for mathematics and early literacy scores. When results are examined separately for children attending PED or CYFD sites, the findings show patterns similar to those found for the entire New Mexico PreK sample, though statistical significance was not always attained for the site-specific analyses.

Summary and Discussion

Our findings show that the New Mexico PreK initiative produced positive effects on children's learning during its second year of operations, consistent with our RDD findings from the initiative's first year (Hustedt, Barnett, & Jung, 2007). The current analyses provide further evidence that New Mexico PreK has positive impacts on children's language, mathematics, and early literacy skills that are evident at the beginning of kindergarten. These types of initial effects are likely to lead to increased school success as well as continued advantages in reading and math skills.

Figure 1 provides an overview of our latest findings, showing estimated effect sizes for the impact of New Mexico PreK on children's receptive vocabulary, print knowledge, and early math scores. The effect sizes in this figure were calculated by comparing the regression coefficient to the standard deviation of the control group (analogous to Glass's delta), and are a common way of standardizing the estimated effects of the prekindergarten program. As reported for the previous school year, the estimated impacts of New Mexico PreK on children's learning are both substantive and meaningful when measured at kindergarten entry.

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



These effect size calculations also provide a basis for comparing results from the second year of the New Mexico PreK initiative to results from the first year (Hustedt, Barnett, & Jung, 2007). For our receptive vocabulary measure, the estimated effect size was 0.25 using Fall 2007 data, compared to 0.36 for Fall 2006. The effect size for our mathematics measure was 0.50 for Fall 2007, compared to 0.39 the previous year. While effect sizes from the two school years appear somewhat different at first glance, confidence intervals used to gauge the precision of our estimates show that estimates from the previous year fall within the confidence intervals for our current estimates. This

shows that the estimated impacts of New Mexico PreK on children's vocabulary and math scores are consistent across the two years of this study. The effect size for our measure of print knowledge measure was 0.59 for Fall 2007 compared to an effect size of 1.16 for print awareness the previous year; in that case the apparent decrease in effect size may be due to differences between the instruments used at the two assessment points.

Effect sizes for the New Mexico PreK initiative are comparable in magnitude to those reported for other well-regarded state preschool programs, such as the one in Oklahoma (Gormley et al., 2005). New Mexico's situation is unique, though, in that the evaluation of its state prekindergarten initiative began during the same school year as the initiative itself. The Oklahoma study and the other state prekindergarten evaluations conducted by NIEER (Frede et al., 2007; Hustedt, Barnett, Jung, & Thomas, 2007; Wong et al., 2008) have focused on well-established programs that had been operating for a number of years before the evaluation began. While the impacts found in this study are precursors to later success in school and may ultimately lead to the types of economic benefits found in long-term studies of preschool education (Schweinhart, Montie, Ziang, Barnett, Belfield, & Nores, 2005; Campbell, Ramey, Pungello, Sparling, & Miller-Johnson, 2002; Reynolds et al., 2002), New Mexico PreK should also be viewed as a work in progress. The children who took part in this research were participants during the first years of a new initiative that was not yet widely available to their peers across the state. In fact, only 8% of the state's 4-year-olds were enrolled in New Mexico PreK during the 2006-2007 school year (Barnett et al., 2007), when the Preschool group in this study completed PreK. As noted earlier, PPVT standard scores at kindergarten entry were only 89 on average, which puts them at the 23rd percentile. Thus, there remain substantial opportunities to increase the effectiveness of the New Mexico PreK initiative.

As the New Mexico PreK initiative matures, children's experiences may start to differ from those reported here. Although an income requirement is not used to determine eligibility, two-thirds of the children enrolled at a PreK site must live within the attendance area of a Title I school, and sites are also prioritized for funding based on local schools' progress toward adequate yearly progress under No Child Left Behind (Barnett et al., 2007; Pre-Kindergarten Act, NMSA 1978 § 32A-23-6, 2005). These current emphases lead to a focus on serving children from the most disadvantaged neighborhoods or communities. As further expansion occurs, priorities for funding may shift, with state-funded prekindergarten becoming more accessible to children who are less disadvantaged.

Also, teacher education requirements are being phased in as the New Mexico PreK initiative matures. All lead teachers at a PreK site must have a bachelor's degree and licensure in Early Childhood Education within the first five years of that site's establishment. Sites that began operating during the initiative's first year will be required to fully meet this requirement by the 2010-2011 school year. When teacher degree requirements are more fully implemented during the upcoming years, children are likely to benefit, as research supports the importance of having teachers with bachelor's degrees and specialized training in early childhood education (Barnett, 2003; Bowman, Donovan,

& Burns, 2001; Burchinal, Cryer, Clifford, & Howes, 2002; Whitebook, Howes, & Phillips, 1989). These and other improvements to program quality that occur in conjunction with the expansion of New Mexico PreK may have a positive effect on the future impacts of PreK on children's school readiness.

Although the results from CYFD and PED sites were similar in magnitude to the overall results from the entire sample, the individual CYFD and PED estimates of the impacts of PreK did not reach statistical significance in all cases. In general, though, these site-specific analyses lack the statistical power of the analysis utilizing the overall sample. Having a larger sample size allows us to have greater confidence in our estimates of the effects of the New Mexico PreK initiative, and for statistical models that better fit the data, including quadratic and cubic functions selected for the overall sample. The lack of statistical significance for some of the child outcome results for CYFD and PED is most likely due to the smaller sample sizes when analyses are limited to sites administered by a single agency, and not due to the lack of a relationship between CYFD or PED PreK participation and child outcomes. In fact, each of the non-significant relationships was positive and several approached significance. During the 2008-2009 school year, we will revise our methodology to provide a closer look at the impacts of New Mexico PreK at PED and CYFD sites. We will increase PED and CYFD samples to approximately 700 children each, which should allow for better estimates of the effects of attending PreK sites administered by each agency.

This is the second in a series of three planned regression-discontinuity reports on the effects of New Mexico PreK, focusing on child data collected in Fall 2007. The data presented here show that the New Mexico PreK initiative continues to have statistically significant effects on children's vocabulary, mathematics, and early literacy skills that are evident when they begin kindergarten. These findings are consistent with our findings from studies of other states' prekindergarten initiatives (Frede et al., 2007; Hustedt, Barnett, Jung, & Thomas, 2007; Wong et al., 2008). By repeating the RDD again during the 2008-2009 school year, we will gain further insight into changes in the impacts of New Mexico's state-funded prekindergarten initiative as it expands to serve even more children. As a result, our series of reports taken together will document progress by the New Mexico PreK initiative as it continues to mature.

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. (2003). Better teachers, better preschools: Student achievement linked to teacher qualifications. *Preschool Policy Matters*, 2. New Brunswick, NJ: National Institute for Early Education Research.
- Barnett, W. S., Hustedt, J. T., Friedman, A. H., Stevenson Boyd, J., & Ainsworth, P. (2007). *The state of preschool 2007: State preschool yearbook*. New Brunswick, NJ: National Institute for Early Education Research, Rutgers University.
- Bowman, B. T., Donovan, M. S., & Burns, M. S. (Eds.). (2001). *Eager to learn: Educating our preschoolers*. Washington, DC: National Academy Press.
- Burchinal, M. R., Cryer, D., Clifford, R. M., & Howes, C. (2002). Caregiver training and classroom quality in child care centers. *Applied Developmental Science*, 6, 2–11.
- Bureau of Business and Economic Research, University of New Mexico. (2007). *State and county population estimates by age, sex, race and Hispanic origin from the Census Bureau*. Retrieved May 7, 2008, from <http://www.unm.edu/~bber/demo/coestchar.htm>
- Campbell, F. A., Ramey, C. T., Pungello, E. P., Sparling, J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecedarian Project. *Applied Developmental Science*, 6, 42-57.
- 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.
- 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, Rutgers University.

- 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.
- 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, 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, Rutgers University. Available at <http://nieer.org/resources/research/ArkansasYear1.pdf>
- Judd, C. M., & Kenny, D. A. (1981). *Estimating the effects of social interventions*. New York: Cambridge University Press.
- Lonigan, C. J., Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (2007). *Test of Preschool Early Literacy (TOPEL)*. Austin, TX: Pro-Ed.
- Lonigan, C. J., Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (2002). *Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP)*. Tallahassee, FL: Department of Psychology, Florida State University.
- 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, Rutgers University.
- 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.ssc.wisc.edu/irp/pubs/dp124502.pdf>.
- Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., & Nores, M. (2005). *Lifetime effects: The High/Scope Perry Preschool study through age 40* (Monographs of the High/Scope Educational Research Foundation, 14). Ypsilanti, MI: High/Scope Educational Research Foundation.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin Company.
- 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.
- Whitebook, M., Howes, C., & Phillips, D. (1989). *Who cares? Child care teachers and the quality of care in America* (Final report of the National Child Care Staffing Study). Oakland, CA: Child Care Employee Project.
- 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, A. F. (1990). *Bateria Woodcock-Munoz Pruebas de Aprovechamiento – Revisados*. Itasca, IL: Riverside Publishing.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson Tests of Achievement*. Itasca, IL: Riverside Publishing.

Appendix: 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:

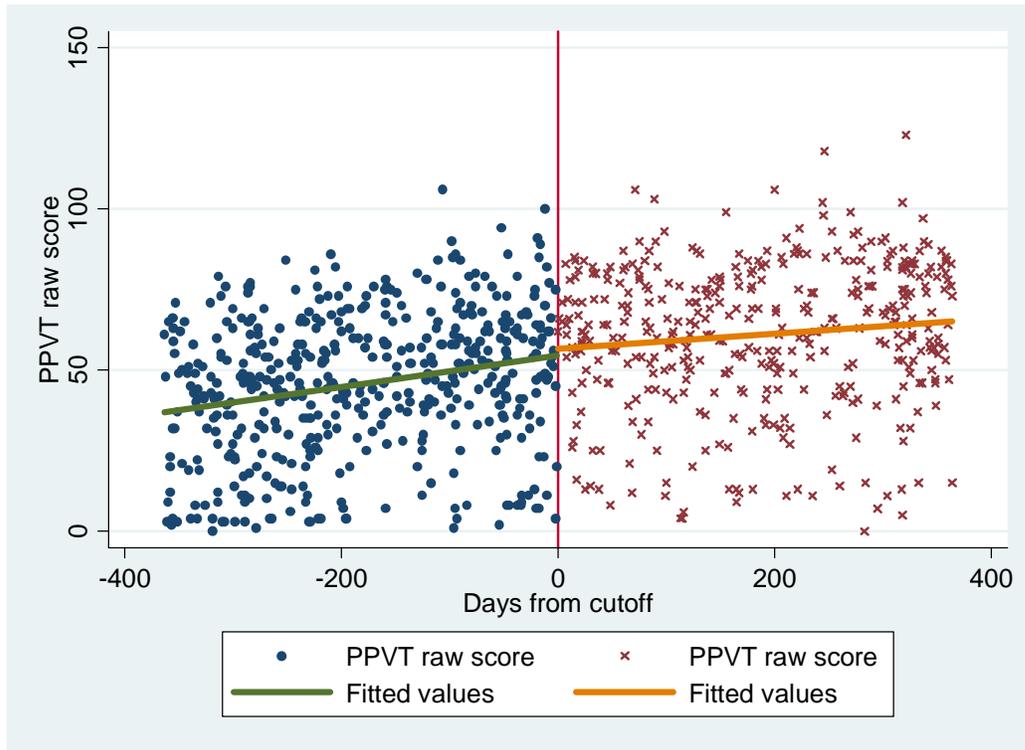
$$Y_{ij} = a + BX_{ij} + \beta_1(\text{Pre-K})_{ij} + g(\text{AV})_{ij} + \varepsilon_i$$

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 $g(\text{AV})_{ij}$ is a smooth function of the continuous assignment variable. We check the robustness of our estimates by considering alternative specifications for $g(\text{AV})_{ij}$, including polynomials and interaction terms. We determine the order of the polynomial approximation to the $g(\text{AV})_{ij}$ function by examining the statistical significance of the higher order and interaction terms.

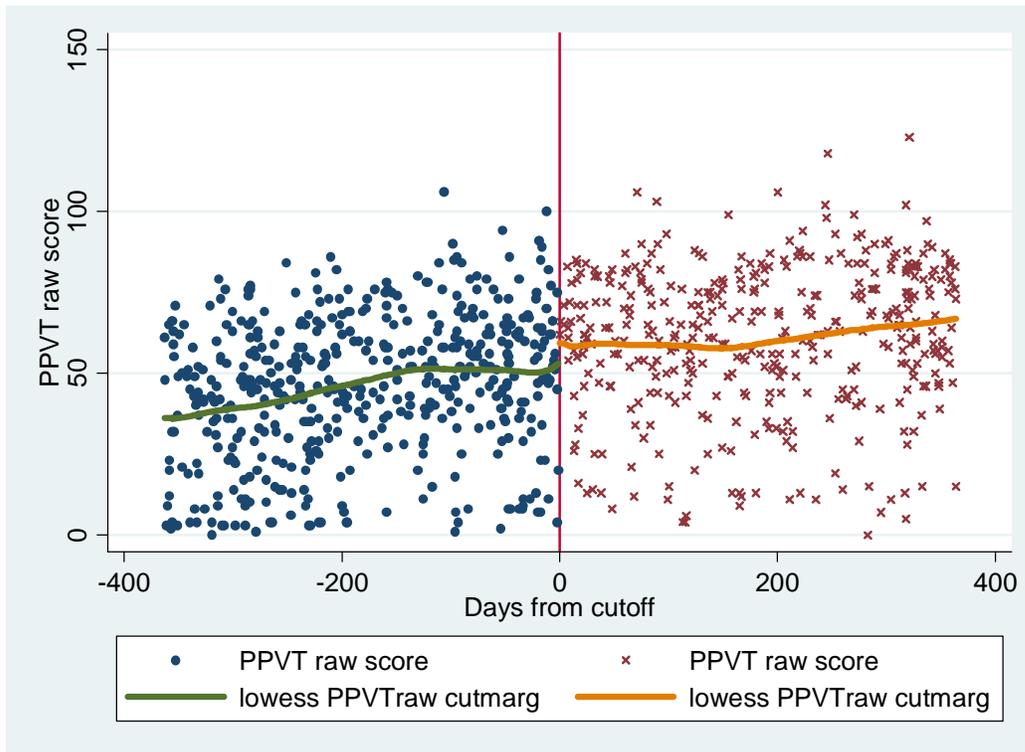
Following Trochim (1984), when the functional form of the regression model is ambiguous, we overfit the model by including more polynomial and interaction terms than needed, yielding unbiased but less efficient estimates. In all the parametric analyses we use Huber-White standard errors adjusted for 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

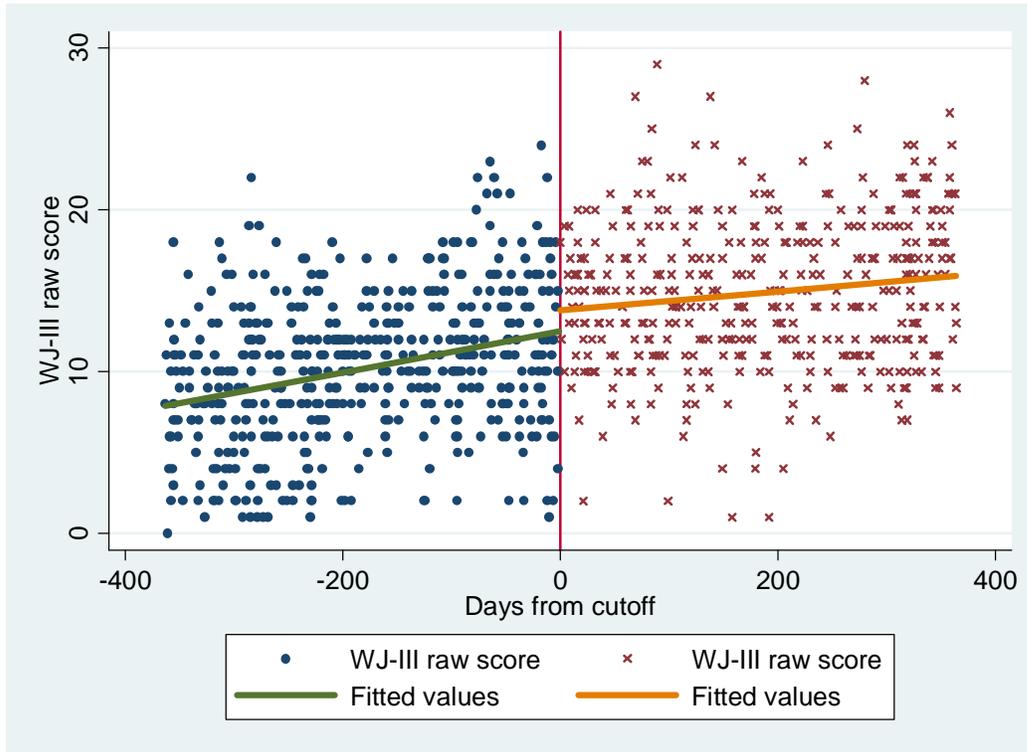


(1) Linear Plot

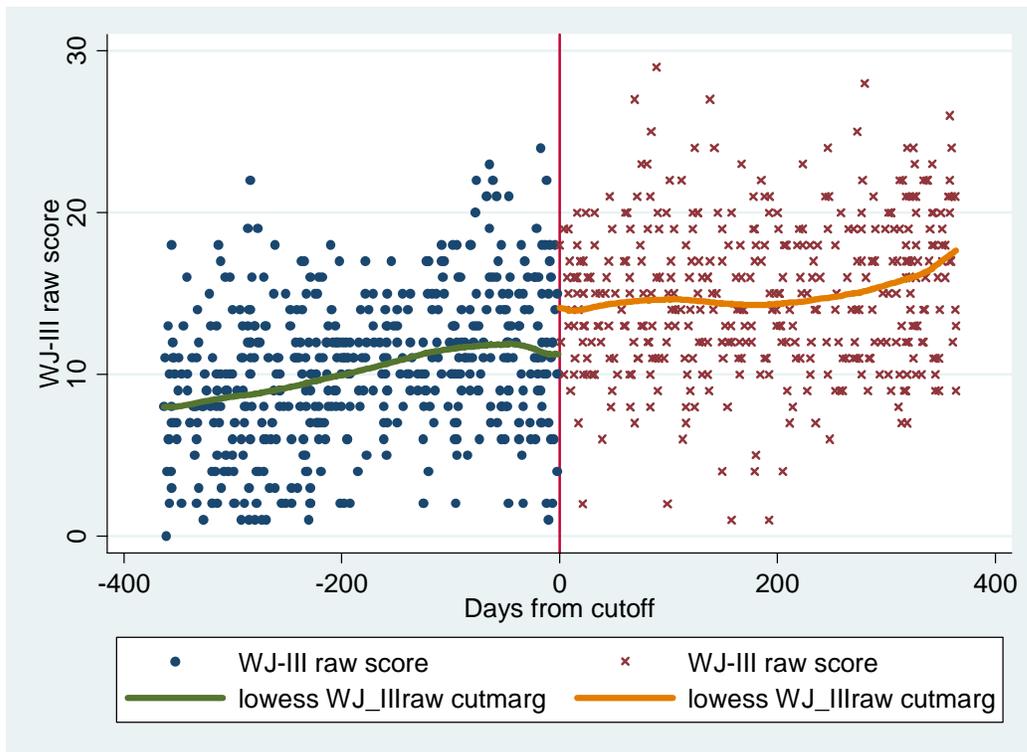


(2) Lowess Plot

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

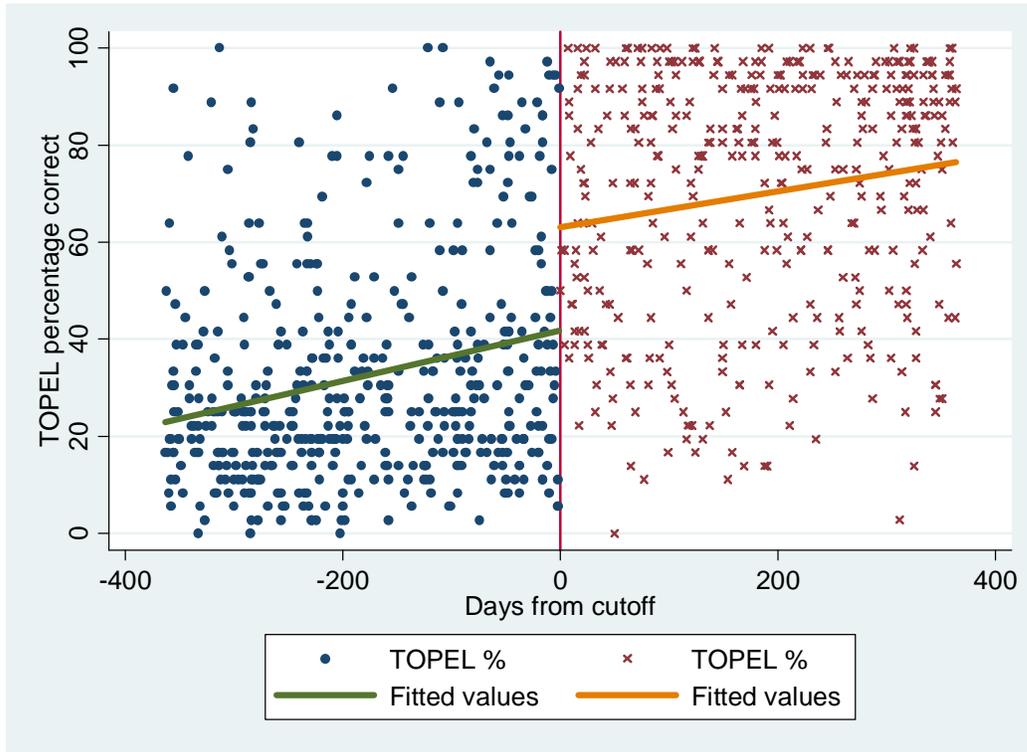


(1) Linear Plot

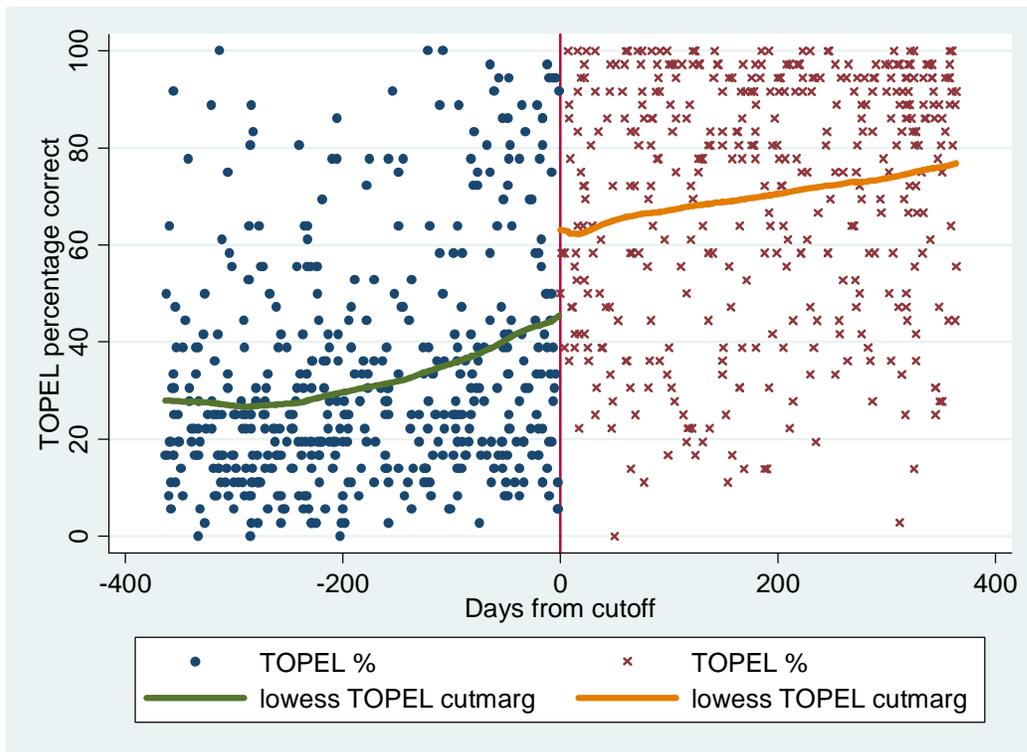


(2) Lowess Plot

Figure A3. Linear and Lowess Plots for TOPEL Print Knowledge Results



(1) Linear Plot



(2) Lowess Plot