PROMOTING SCHOOL READINESS IN OKLAHOMA:

AN EVALUATION OF TULSA'S PRE-K PROGRAM[**]

William T. Gormley, Jr.[*]   Ted Gayer[*]

Abstract

Since the mid-1990s, three states, including Oklahoma, have established a universal pre-kindergarten (pre-K) program. We analyze the effects of Oklahoma's universal pre-kindergarten (pre-K) program for four-year-olds on children in Tulsa Public Schools (TPS). The main difficulty with testing the causal impact of a voluntary pre-K program is that certain parents are more likely to select pre-K, and these parents might have other unobservable characteristics that influence the test outcomes of their children. Because TPS administered an identical test in September 2001 to children just beginning pre-K and children just beginning kindergarten, we can compare test outcomes of "old" pre-kindergarten students to test outcomes of "young" kindergarten students who attended pre-K the previous year. We find that the Tulsa pre-K program increases cognitive/knowledge scores by approximately 0.39 standard deviation, motor skills scores by approximately 0.24 standard deviation, and language scores by approximately 0.38 standard deviation. Impacts tend to be largest for Hispanics, followed by blacks, with little impact for whites. Children who qualify for a free lunch have larger impacts than other children.

PROMOTING SCHOOL READINESS IN OKLAHOMA:

AN EVALUATION OF TULSA'S PRE-K PROGRAM

I. Introduction

In recent years, several state governments have launched initiatives aimed at providing pre-kindergarten (hereafter, pre-k) programs for all four-year old children whose parents wish to enroll them. The program is thus universal, but voluntary. Georgia established such a program in 1995,[i] followed by New York in 1997[ii] and Oklahoma in 1998. The District of Columbia also has such a program. Other jurisdictions are moving in this direction.[iii]

A key premise behind these initiatives is that a universal pre-k program will enhance school-readiness and give all students the kind of "head start" that the Head Start program was designed to provide to eligible low-income students. But is this premise correct? Some studies show that well-funded, well-designed, and well-staffed early intervention programs can improve the academic achievement of disadvantaged children (Barnett, 1993; Campbell and Ramey, 1995; Reynolds et al., 2001). But can we extrapolate from intensive early intervention programs to less intensive programs? Would we see the same results from programs that serve disadvantaged children and programs that serve middle-class and upper-class children as well? Can we generalize from selective programs to universal programs, in view of the larger staffing challenges the latter face?

The primary difficulty with assessing a *voluntary* pre-k program is that certain parents are more likely to select the pre-k program for their children, and these parents (or their children) might also have other unobservable characteristics that influence the test outcomes of the children. In this paper, we assess Oklahoma's universal pre-k program by focusing on data from the Tulsa Public Schools (TPS), the largest school district in the state. TPS struck us as an

excellent research site for two reasons. First, TPS administered the same test to 4-year olds beginning TPS pre-k and 5-year olds beginning TPS kindergarten in September 2001. Second, TPS used a strict cut-off birthday requirement for enrollment in TPS pre-k. This strict cut-off birthday requirement creates a discontinuous relationship between age and whether the child attended TPS pre-k in the 2000 or 2001 academic year. Therefore we can compare children who just made the cut-off to children who just missed it. If other characteristics of the children are sufficiently smooth at this cut-off date, then this regression-discontinuity design can identify the effect of the TPS pre-k program on test scores.

II. Literature Review

The literature on the effects of Head Start, pre-k programs, and enriched early intervention programs is rich and intriguing. An enriched early intervention program that starts early in life, pays its staff well, and maintains low child-staff ratios, can have significant positive effects on student achievement and other desirable outcomes, at least for disadvantaged children. The Perry Preschool project from Ypsilanti, Michigan and the Abecedarian project from Chapel Hill, North Carolina -- both of which involved random assignment and several years of treatment -- yielded benefits to society and to the children themselves that substantially exceeded project costs (Barnett, 1993; Campbell and Ramey, 1995). Children in the Abecedarian treatment group, for example, performed much better on reading and math tests than children in the control group. They were also less likely to be retained in grade or to be placed in special education classes, more likely to finish high school and to be employed, and less likely to commit a crime.

Like these two exemplary projects, the Head Start program embeds pre-k readiness goals in the context of comprehensive services. In general, however, Head Start is less expensive than these model programs (Currie, 2001, p. 221), in part because it tends to be part-day and part-

year, usually for nine months only.[iv] According to a meta-analysis of the early Head Start

literature, the Head Start program has some positive effects on cognition, socio-emotional

development, and health status (McKey et al., 1985). More recent and more sophisticated

studies have taken a closer look at cognition, using standardized tests and grade retention

measures. Sibling-based comparisons, which control for family background effects, show that

Head Start improves school readiness for white and black children (Currie and Thomas, 1995).

Sibling-based comparisons also reveal positive Head Start effects for Hispanic children,

especially for native-born Hispanics and for Hispanic children of Mexican origin (Currie and

Thomas, 1999).

While Head Start and early intervention programs have been studied thoroughly, pre-k

programs have received less systematic attention. A Michigan study using a non-experimental

research design found favorable results: in kindergarten, teachers rated students who attended a

pre-k program higher in language, literacy, math, music, and social relations. Students who

attended a pre-k program also were more likely to pass the Michigan Educational Assessment

Program's reading and mathematics tests (Xiang and Schweinhart, 2002). Some national

studies, combining a variety of preschool programs (state-funded and otherwise), used sibling-

based comparisons and reached disappointing conclusions about the efficacy of preschool

(Currie and Thomas, 1995; Currie and Thomas, 1999). A recent national study, using the ECLS-

K data, found that kindergarten students who had attended a pre-k program scored higher on

reading and math tests than children receiving parental care (Magnuson et al., 2004). This same

study found that disadvantaged children benefited more from pre-k than other children and that

full-time enrollment was more beneficial than part-time enrollment. However, like so many

others, this study likely suffers from selection bias, since the characteristics that influence

parents' decisions to enroll a child in the pre-k program also may contribute to the outcome measures. In fact, while a meta-analysis of state-funded preschool programs in 13 states found statistically significant positive impacts on some aspect of child development in all of the states (Gilliam and Zigler, 2001, p. 453), none of the studies used random assignment and only one of the evaluations used a comparison group that constituted a credible control for selection bias.[v]

If we know little about the effects of pre-k programs, we know even less about the effects of universal pre-k programs. A Georgia study found that 82 percent of former pre-k students rated average or better on third-grade readiness (Henry et al., 2001), but it lacked an appropriate comparison group. A more recent Georgia study found that economically disadvantaged children attending Georgia's pre-k program began preschool scoring below national norms on a letter-and-word recognition test but began kindergarten scoring above national norms (Henry et al., 2003). The latest Georgia study compared Georgia pre-k, Head Start, and other preschool children but did not include children who attended no preschool. In one urban upstate city in New York state, universal pre-k classrooms averaged between 5.7 and 5.8 on the seven-point Early Childhood Environmental Rating Scale (ECERS), with five being good and seven being excellent (Cochran, 2002, p. 38). Although encouraging, these results fail to tell us whether New York children performed better because of their exposure to the universal pre-k program. An evaluation of New York's program was to have been conducted, but it never received funding from the New York state legislature.

III. Oklahoma's Universal Preschool Program

In 1990, Oklahoma established a pre-k program for disadvantaged children. Specifically, all school districts that wanted to offer a pre-k program received state aid for four-year olds who met Head Start income eligibility requirements. The program was well-received, and in 1998 the

State Legislature voted to make it universal, triggering substantial increases in enrollments (see

Figure 1).

Under the legislation and accompanying state regulations, each of the state's 543 public

school districts can choose to participate in the program. The state provides full funding, with no

match required. As of 2002-3, 494 (91 percent) of the state's school districts were participating.

An estimated 65 percent of all four-year old children in Oklahoma were participating in the

public pre-k program in 2002-3 -- a penetration rate lower than Georgia's but higher than New

York's. Thus, in Oklahoma, "universal" means that all school districts are eligible to participate

in the program, that almost all four-year olds have access to the program, and that most four-year

olds actually are enrolled in the program.

One key provision of Oklahoma's universal pre-k program is that all teachers must have a

college degree and a certificate in early childhood education. A corollary is that pre-k teachers

receive the same compensation as teachers in public elementary schools, which clearly

distinguishes these programs from day care centers, where wages are much lower. Group sizes

are set at 20 and child/staff ratios are set at 10/1. Although pre-k services are provided by public

schools, collaborative arrangements are possible. Approximately 18 percent of Oklahoma

children enrolled in a public pre-k program are enrolled in a collaborative program with Head

Start, a group day care center, a private school, or some other type of facility.

We decided to focus on the Tulsa Public Schools (TPS) pre-k program for three reasons.

First, the TPS is the state's largest school district, in terms of enrollment: 41,495 students, as

opposed to 37,231 students in Oklahoma City, the second largest district. Second, the ethnic

composition of the TPS student body is quite diverse. The student body is: 44 percent white; 36

percent black; ten percent Hispanic; nine percent Native American; and one percent Asian.

Third, beginning in September 2000, the TPS administered an annual Early Childhood Skills

Inventory to students entering TPS pre-k and to students entering TPS kindergarten. Together

with TPS's birthday cut-off policy, this testing allows us to estimate the effects of a universal

pre-k program using a regression-discontinuity approach that contrasts the performance of

children born just before the cut-off date (the treatment group) to the performance of children

born just after the cut-off date (the control group), at the same time controlling for continuous

age effects.

A. Tulsa's Database

In August 2001, TPS administered a 26-item test to most students about to enter their pre-

k and kindergarten programs.[vi] The tests were given to students, individually, by TPS teachers,

who informed parents in advance that their child would need to be tested at a particular time.[vii]

Approximately 76 percent of the district's 1,690 pre-k students (1,284) and approximately 66

percent of the district's 3,441 kindergarten students (2,276) were tested, yielding a total sample

of 3,560 children.[viii]

Did the test-takers make up a representative sample of TPS preschoolers? As Table 1

shows, the gender and free-lunch-eligibility of the tested children are quite similar to their

corresponding populations. There are some small, statistically significant differences in the

racial compositions: Hispanic students are under-represented in our TPS pre-k test sample (ten-

percent level); black students are under-represented in our TPS kindergarten test sample (five-

percent level); and white students are over-represented in our kindergarten test sample (five-

percent level).   There are no other statistically significant differences between the observed

characteristics of our sample and the observed characteristics of the universe of children, but

since few measured characteristics are available to us, we cannot say more definitively whether

our sample is fully representative of the population of Tulsa four-and five-year olds.

Thanks to the cooperation of the Tulsa Public Schools, we know which children who

enrolled in kindergarten in September 2001 participated in the TPS pre-k program during the

previous year. Unfortunately, for children who did not participate in the TPS pre-k program

during 2000-1, we do not know whether they participated in a private pre-k program or the Head

Start program, unless the Head Start program was involved in a collaborative relationship with

TPS.   Thus, for some of the children, we only know that they did not participate in the TPS pre-

k program, not whether they had no pre-k experience. This means that our empirical strategy can

only estimate the treatment-on-the-treated effect -- the effect on test scores of *attending* TPS

pre-k. Given the data limitations, we cannot estimate the intent-to-treat effect -- the effect on the

population's test scores of making the TPS pre-k program available.

IV. Empirical Strategy

A. Research Design

This paper examines whether attending a TPS pre-k program leads to short-term

improvements in test outcomes for children. The relationship between TPS pre-k and test scores

can be captured using the following model:

(1) $y_i = \gamma X_i + \theta T_i + u_i$

(2) $T_i = \Gamma X_i + v_i,$

where y_i is child i's test outcome; X_i is a vector of observed characteristics of child I; T_i is an

indicator variable of whether or not the child attended a TPS pre-k program; u_i captures the

unobservable determinants of child test outcomes; and v_i captures the unobservable determinants

of whether the child attended TPS pre-k. The observable characteristics are race (white, black,

Hispanic, Native American, and Asian), gender, and whether the child qualified for partial, full, or no free lunch. Note that the treatment we examine is "TPS pre-k," which is a pre-kindergarten program administered by TPS and funded by the state of Oklahoma. So the counterfactual to the treatment is either that the child gets no pre-k program or that the child gets another (perhaps private) day care or pre-k program or Head Start. The treatment of TPS pre-k also includes a possible income effect, since parents who would have paid for a pre-k program if the treatment were not available now have more disposable income to spend on their child.

In order for an OLS estimation of equation (1) to achieve an unbiased estimate of θ, it must be the case that $E[u_i v_i/X_i] = 0$. If the unobservable determinants of whether a child goes to TPS pre-k are correlated with unobservable determinants of test scores, then a naïve cross-sectional analysis would yield biased estimates of the impact of TPS pre-k on test scores.

The clearest way to estimate the causal impact of TPS pre-k on test scores would be to randomly select each child to be either in the treatment group (attend TPS pre-k) or the control group (do not attend TPS pre-k). Afterwards, the children would be given the same test to see whether the treatment affects test outcomes. Randomization assures that $E[u_i v_i/X_i] = 0$. In the absence of a randomized study, we exploit the structure of Tulsa's child testing program and strict age qualifications in order to identify the causal impact of TPS pre-k on test scores.

Within the Tulsa school district, children were qualified to attend TPS pre-k in academic year 2000-1 if, and only if, they were born before September 1, 1996 (and after September 1, 1995). Children whose birthdays were between September 1, 1996 and August 31, 1997 were required to wait until the following academic year to enroll in TPS pre-k; at that time, they automatically qualified.[ix] In September 2001, the Tulsa school district administered an aptitude

test to all TPS pre-k and kindergarten students: *the identical test was administered to both groups of students*. This gives us the following framework with which to work:

Starting TPS Pre-k in 2001	*Starting TPS K in 2001*
Not yet treated (Control2)	Had TPS Pre-k (Treatment)
	Did not have TPS Pre-k (Control1)

These test score data could be used in a cross-sectional analysis to compare the scores of kindergarten students who attended TPS pre-k the previous year (Treatment) to the scores of those who did not (Control1), controlling for other observable characteristics. As mentioned earlier, though, such an analysis would likely yield misleading results, because the former group could have unobservable characteristics that differ from the latter group.

Table 2 illustrates the problem associated with this cross-sectional regression. There are four different tests, measuring social/emotional skills, cognitive/knowledge skills, motor skills, and language skills. We also include the total test score, which is the sum of the scores from the four tests. While the treatment group of children who have had TPS pre-k shows statistically significant higher scores for the four different tests, the treatment and control samples differ along many of the observed characteristics. For example, the children who were in TPS pre-k are more likely to have been on the full free-lunch program and less likely to have not qualified for any free lunch program, and a much higher proportion of these children are black. The differences in observable characteristics suggest that Control1 is not a valid control group, and that there may well exist unobservable differences across the two groups. To the extent that these differences influence test outcomes, the regression strategy would yield biased estimates.[x]

Another estimation strategy would involve comparing the test scores of kindergarten students who attended TPS pre-k the previous year (Treatment) to the scores of the children just

beginning TPS pre-k (Control2). This has the advantage of choosing both a treatment and a

control group that selected into the treatment, with the latter not yet having been treated. But

even if one controls for the effect of age on test scores, a potential problem with this strategy is

that, while the selection criteria may be constant over the two years, the different populations

may have different characteristics. This could be because of changing socio-demographic

characteristics within Tulsa over time or changes in the selection process of parents over time.

The former suggests the possibility of omitted-variables bias (if unobservable determinants of

test scores also differ over time), and the latter suggests the possibility of selection bias (since

different populations are selecting into TPS pre-k over time).

The first set of columns in Table 3 confirms these concerns. The difference in test scores

between those who qualified and enrolled in TPS pre-k the previous year (denoted as "Before

9/1") and those who are currently qualified and enrolled in TPS pre-k (denoted as "After 9/1") is

quite substantial. The total test score is 5.7 points higher for those children who were previously

in TPS pre-k than for those who are currently in TPS pre-k. Keep in mind that much of this test

differential is likely attributable to age differences. But even after controlling for age, the

omitted-variable bias and selection bias are a concern in this analysis, because the two groups

differ substantially in their other observable characteristics. The treatment group has a higher

proportion of children in the full free-lunch program, a lower proportion of children with no or

only partial free-lunch program, and a higher proportion of non-whites (especially Hispanics).

Again, these differences suggest that the two groups probably also have different unobservable

characteristics, which could lead to biased estimates.[xi]

In the second and third set of columns, we can see that narrowing the margin closer to the

cut-off qualification birth date results in a decrease in test score differentials, which is likely

attributable to the reduced influence of age. We also see that some of the differences in observable characteristics disappear as the margin narrows. For example, there is no statistically significant difference in the proportion of children who are non-white once the margin is reduced to six or three months. However, some differences in observable characteristics still persist.

In order to address the concerns about using observational data, we base our empirical strategy on the strict birthday cut-off which generates a highly non-linear relationship between age and whether a child attended TPS pre-k in 2000 or 2001. Figure 2 thus illustrates the relationship between the age of the child (as measured in days from the September first cut-off date) and whether the child attended TPS kindergarten in 2001.[xii]

As demonstrated in Figure 2, the assignment of TPS kindergarten attendance in 2001 (versus TPS pre-k attendance in 2001) as a function of age lends itself to a regression-discontinuity design (see Cook and Campbell, 1979). Figure 2 shows both the control group (those children who missed the cut-off birth date and therefore were in TPS pre-k in 2001) and the treatment group (those children who made the cut-off birth date and therefore were in TPS pre-k the year before). The data points to the left of the cut-off date consist of the "young" children: those who qualified for TPS pre-k in 2001. These children make up our control group, because they selected into the program but (as of the test time of August 2001) they have yet to experience the treatment. The data points to the right of the cut-off date consist of the "old" children: those who qualified and enrolled in TPS pre-k in 2000 and are in TPS kindergarten in 2001. These children make up our treatment group because they have already received the TPS pre-k education in 2000: (as of the test time of August 2001) they are just beginning TPS kindergarten. Even though the older children are likely to be systematically different from the younger children (with these differences contributing to higher test scores), the children who

barely missed the September first cut-off birth date are likely to be comparable to the children who barely made the cut-off birth date.

The identifying assumption that needs to hold is that the unobservable characteristics of the children not vary discontinuously around the cut-off birth date. That is, the children who were in TPS kindergarten in 2001 and attended TPS pre-k in 2000 may be different from the children who attended TPS pre-k in 2001 (aside from differences in age), but the research design assumes that these differences are not discontinuous at the cut-off birth date. If there are differences in unobservable characteristics of the children near the cut-off birth date, then this identifying assumption may not hold. For example, parents with children born near the cut-off date might be concerned about their children being either the youngest or oldest in their grade; this may influence their children's test scores in unobservable ways. A discontinuity of unobservable characteristics at the cut-off date could bias our results. But we find that observable characteristics are not discontinuous at the cut-off birth date, which suggests (though not conclusively) that the identifying assumption holds. Therefore we believe that the assumption of constant treatment effects is reasonable. However, if the impact of TPS pre-k on test scores is different for children with different birthdays, then one cannot make causal inferences across the range of birth days.

B. Quadratic Specification

If the birthday cut-off requirement perfectly were enforced, then there would be a perfectly discontinuous relationship between birth date and enrollment. In that case, a properly specified OLS model, including a dummy variable for whether the child made the cut-off, would result in unbiased estimates of the effect of TPS pre-k (assuming, again, that all the other determinants do not vary discontinuously at the same cut-off date).

As mentioned earlier, there are 3,560 children in our sample. Because we are estimating the effect of the TPS pre-k program, our treatment group consists of children who were in TPS pre-k in 2000 and in TPS kindergarten in 2001. We therefore drop the 1,164 observations of kindergarten children who were never in TPS pre-k, leaving 1,112 children in our treatment group. Our control group consists of the 1,284 children who are just beginning TPS pre-k in 2001. We drop the 20 outlier observations that fall outside the 12 month plus-or-minus range of the cut-off date, leaving a total sample of 2,376 observations. Although the relationship between birth date and enrollment for the remaining 2,376 children is not perfectly discontinuous, it is close enough that we proceed as if it were.[xiii] That is, we drop the observations for the 18 children who qualified for TPS pre-k in 2000 but were in TPS pre-k in 2001, and we drop the observations for the four children (three born right on the cut-off date) who qualified for TPS pre-k in 2001 but were in TPS kindergarten in 2001. This leaves 2,354 observations in our sample.

Our goal is to estimate a flexible specification of test scores as a function of age leading up to and after the cut-off birth date. We accomplish this by separately regressing test scores on a second-order polynomial of the difference between birth date and cut-off date, for the sample of children born before and after the cut-off date. Figure 3 shows the scatter plots for the four different tests, as well as fitted cubic spline functions on both sides of the cut-off birth date.[xiv] While not conclusive, the figure suggests that the age/test score relationship is linear on both sides of the cut-off date.

We opted for a second-order, or quadratic, polynomial specification because it offers a more flexible fit for the age/test score relationship. We also tried higher-order polynomials, but the additional terms failed an F-test of joint significance. In addition, our results are fairly robust

to linear, quadratic, and higher-order polynomial specifications. Using the quadratic specification, we compare test scores at the limit approaching from the left to test scores at the limit approaching from the right, and any effect of TPS pre-k is then captured in the difference.

The final set of columns in Table 3 shows that this method balances the observable characteristics between the children who were in TPS pre-k in 2001 (control group) and the children who were in TPS pre-k the previous year (treatment group). In the sixth through 16th rows (which show the demographic characteristics), all of the differences between the treatment and control groups are statistically insignificant.[xv] This suggests that the quadratic regression-discontinuity design credibly replicates a randomized experiment, since the treatment and control groups are similar along observable characteristics.

The parents of children who just miss the TPS pre-k cut-off date might enroll them in a private pre-k while they wait a year to qualify for the TPS pre-k program. This would effectively give the children who just missed the cut-off date (control group) an extra year of pre-k relative to the children who just made the cut-off date (treatment group). As mentioned earlier, this study can only estimate the treatment-on-the-treated effect, so the counterfactual is what children did in the absence of TPS pre-k. We cannot estimate the intent-to-treat effect because we cannot distinguish between children who enroll in a private pre-k program, children who enroll in another program, or children who stay at home for the year awaiting TPS pre-k.

In addition to showing that the covariate observations balance, Table 3 shows that the differences in the cognitive/knowledge test scores and the language test scores remain large and statistically significant at the cut-off date. This indicates that TPS pre-k increases cognitive/knowledge test scores by 0.541 points (out of a possible seven points) and increases language test scores by 0.561 points (out of a possible ten points). These effects are equivalent

to 0.28 and 0.23 standard deviations, respectively. Again, these test score effects of TPS pre-k are relative to the control group of children who have not yet received TPS pre-k. Figure 4 shows the predicted values for the test scores at various ages, with a discontinuous increase in these test scores at the cut-off birth date. The figure also suggests that the relationship between test scores and age is primarily linear, so while the second-order specification offers added flexibility, it contributes little in terms of explanatory power. A linear specification leads to very similar results.

The quadratic regression-discontinuity estimates for test scores can be estimated in a single-equation model which includes the difference (in days) between birth date and cut-off date, the square of this term, a cut-off dummy variable, and interactions. In our specifications, we also include other observable covariates in order to estimate the effects of these covariates on test outcomes. These other covariates (denoted as the vector X_i) measure whether the child receives no free lunch, partial free lunch, or full free lunch (no free lunch is the omitted category); whether the child is a boy or girl; and the race of the child (white, black, Hispanic, Native American, or Asian, with white as the omitted category). The estimation equation is as follows:

$$(3) \qquad y_i = \gamma_i W_i + \gamma_2 W_i^2 + \gamma_3 (W_i \times T_i) + \gamma_4 (W_i^2 \times T_i) + \theta T_i + \gamma_5 X_i + u_i,$$

where W_i measures the number of days between child i's birthday and the cut-off date, T_i is an indicator variable of whether the child received the TPS pre-k treatment (is born before the cut-off date); and the other variables are as defined earlier. The coefficient estimate of θ gives the mean test score difference at the cut-off date between those who did and those who did not attend TPS pre-k.

V. Results

A. Quadratic Regression-Discontinuity

The last set of columns in Table 3 show the difference in the predicted probabilities of the test scores using a quadratic parametric fit on both sides of the cut-off date. In Table 4, we estimate equation (3), which includes the observable covariates in the regression equation. Since the observable covariates are balanced at the cut-off (as shown in Table 3), adding these variables to the regression equation should not significantly change the estimation results. However, they do allow for tests of the impacts of these characteristics on test scores. Thus, we estimate equation (3) in order to test the impact of TPS pre-k on the different test scores as well as the impact of the observable characteristics on the test scores.

Table 4 suggests that TPS pre-k does have an impact on three of the four tests. The largest effects are on language scores (0.817 increase; 0.33 standard deviations higher) and on cognitive/knowledge scores (0.756 increase; 0.39 standard deviations higher), both significant at the one-percent level. There is a smaller increase of 0.413 (0.24 standard deviations higher) on the motor skills score.[xvi]

Appendix B replicates the results for cases in which the window around the cut-off date has been narrowed. Because identification in the regression discontinuity research design comes from the observations in the neighborhood around the cut-off, narrowing the margin should reduce any bias (although reducing the number of observations increases the standard errors). The results in Appendix B demonstrate that the point estimates are rather robust as the margin around the cut-off date narrows.[xvii]

The absence of statistically significant effects on socio-emotional development is worth noting here. It is possible that the measures of socio-emotional development are too few (only three) to capture the actual effects of TPS pre-k. Indeed, the mean score on this test was 2.74

(2.67 for the children who missed the cut-off birthday and 2.81 for the children who made the cut-off). And, as the first graph of Figures 3 and 4 shows, there is virtually no variation in this test score across ages. Therefore, the actual impact of TPS pre-k on socio-emotional development may be masked by "ceiling effects" of the testing instrument.

In Table 5 we report results separately for Hispanics, blacks, and whites, depending on whether they were enrolled in a half-day or full-day TPS pre-k program. Each cell of Table 5 contains the estimated TPS pre-k treatment effect for race by half-day/full-day program. The table does not report the coefficient estimates for the other covariates. For each estimation the control group's race and half-day/full-day designation corresponds to the treatment group's race and half-day/full-day designation. In the case of Hispanics, we observe higher cognitive scores, higher language scores, and higher motor skills scores for those enrolled in a full-day TPS pre-k program than for those just beginning a full-day TPS pre-k program. In the case of blacks, we find higher language scores and higher cognitive scores for those enrolled in a full-day TPS pre-k program than for those just beginning a full-day TPS pre-k program. We also observe lower socio-emotional scores for black children enrolled in a half-day TPS pre-k program than for black children just beginning a half-day TPS pre-k program. In the case of whites, those enrolled in a half-day TPS pre-k program have higher language scores than those just beginning a half-day TPS pre-k program.

It is difficult to determine the impact of full-day versus half-day programs, because different racial groups tend to sort into these programs. For example, for white children in a half-day program, there is a statistically significant effect on language test scores, but no such effect is found for white children in the full-day program. This might be because most of the white children select into a half-day program. Thus the high standard errors for the impact of the

full-day program on white children's test scores may be attributable to the small sample size. For Hispanic children – where the sample sizes are relatively even across full and half-day programs – there is a larger estimated impact of full-day relative to half-day programs. However, even here, it is possible that the unobservable characteristics of Hispanic children who selected into the full-day program differ from those of Hispanic children who selected into the half-day program. If so, one cannot know for sure that a full-day program would be more effective than a half-day program if these Hispanic children – or other Hispanic children – were randomly assigned to full-day and half-day programs. What we do know is that the full-day program in Tulsa is more effective than the half-day program for Hispanic children who enrolled in these programs. The same logic applies to black children, where we also find the full-day program to be effective for black children who selected into it.

While not reported in the tables, we also find large and statistically significant improvements in cognitive/knowledge, motor skills, and language scores for children who qualified for the full free-lunch program. We find no statistically significant effect for the children who did not qualify for free lunch. Also, we find no impact of TPS pre-k on any test score for white children in the full free-lunch program, or for black children who were not receiving full or partial free lunch. This suggests that the advantages of TPS pre-k tend to accrue to minorities of low economic status.

VI. Conclusion

Because Oklahoma's pre-k program is universal, it is impossible to conduct a randomized experiment to assess its impact on test scores. Relying instead on observational data, we use the strict cut-off age qualification for TPS pre-k in order to replicate randomization.

Oklahoma's universal pre-k program offers tangible benefits to young children in Tulsa, especially low-income and minority children. For three of the four dimensions we examine – cognition, language, and motor skills – the effects of TPS pre-k on child development are clear. For the entire sample, we find an increase in cognitive/knowledge scores of approximately 0.39 standard deviation, an increase in language scores of approximately 0.33 standard deviation, and an increase in motor skills scores of approximately 0.24 standard deviation. We do not find a statistically significant impact of the TPS pre-k program on socio-emotional development, but that might be because of the restrictive survey instrument used and the resulting lack of variation in the social/emotional test scores.[xviii]

The positive effects of TPS pre-k are greatest for Hispanic children, followed by black children. For whites, there is only a modest positive impact on language test scores for those in the half-day program. The actual impact of TPS pre-k on white children may be greater than we have reported here because of "ceiling effects" from the testing instrument.[xix] We also find a positive impact of TPS pre-k on test scores for children who qualify for the full free-lunch program. Thus both minority children and children of lower socio-economic status gain the most from TPS pre-k. These findings support Currie (2001, p. 27), who argued that the "payoff to early intervention is greatest for the most disadvantaged children."[xx]

However, we do extend Currie's findings in two respects. First, the positive effects of early intervention are not limited to programs that begin with infants or toddlers. A program that targets four-year olds, like the Oklahoma pre-k program, can have positive effects on the cognitive, language, and motor skills of young children, especially disadvantaged children. Second, the positive effects of early intervention on disadvantaged children are not limited to

model programs or Head Start programs but also extend to other preschool programs, even massive programs available to children across the socio-economic spectrum.

Whether the benefits of Oklahoma's pre-k program exceed the costs is beyond the scope of this paper. However, we can say that the costs of the program appear to be relatively low in comparison to those of the Head Start program or to such celebrated early intervention programs as the Perry Preschool Program and the Abecedarian Project. According to an estimate from the TPS Budget Office, the cost of the Tulsa pre-k program for fiscal year 2003 was: $3,488 per child for the full-day program and $1,714 per child for the half-day program.[xxi] Because these figures exclude federal spending (Title I, special education, the school lunch program), they understate the actual cost somewhat. For the full-day program, the actual cost probably falls somewhere in between $3,488 per child and $6,088 per child, which is the average cost for all children enrolled in the Tulsa public school system in fiscal year 2003.[xxii]

The focus of this paper has been on school readiness. Questions remain as to whether these effects persist over time. Some research suggests that this may vary by ethnic group. For example, Head Start's positive test score impacts on white and Hispanic children seem to persist over time (Currie and Thomas, 1995; Currie and Thomas, 1999), while impacts on black children seem to fade after an initial burst (Currie and Thomas, 1995). Other studies indicate that student test scores have important implications for future earnings, especially for disadvantaged children. For example, Currie and Thomas (2001) found that test scores at the age of seven are good predictors of test scores at the age of 16, and that the latter are good predictors of wages and employment at the age of 33. It is important to look beyond test scores when assessing the effects of Head Start, pre-k, and similar programs. For example, blacks who attended Head Start are less likely to have been booked or charged with a crime (Garces, Thomas, and Currie, 2002).

In future research, we hope to assess whether short-term improvements in school readiness experienced by Oklahoma pre-k participants translate into longer-term improvements, both inside and outside the classroom.

REFERENCES

Barnett, W. Steven. 1993. "Benefit-Cost Analysis of Preschool Education," *American Journal of Orthopsychiatry* 63, pp. 500-508.

Berls, Abbey and Irene McEwen. 1999. "Battelle Development Inventory," *Physical Therapy* 79, pp. 776-783.

Campbell, Frances and Craig Ramey. 1995. "Cognitive and School Outcomes for High-Risk African-American Students at Middle Adolescence: Positive Effects of Early Intervention," *American Educational Research Journal* 32 (Winter), pp. 743-772.

Cochran, Moncrieff. 2002. "The Implementation of Universal Prekindergarten in an Urban, Upstate City." Unpublished manuscript. Ithaca, N.Y.: Cornell University, Early Childhood Program.

Cook, Thomas D., and Donald T. Campbell. 1979. *Quasi-Experimentation: Design and Analysis Issues for Field Settings* (Houghton Mifflin Co.: Boston, MA).

Currie, Janet. 2001. "Early Childhood Education Programs," *Journal of Economic Perspectives* 15 (Spring), pp. 213-238.

Currie, Janet and Duncan Thomas. 2001. "Early Test Scores, School Quality and SES: Longrun Effects on Wage and Employment Outcomes," *Research in Labor Economics* 20: 103-132.

Currie, Janet and Duncan Thomas. 1999. "Does Head Start Help Hispanic Children?" *Journal of Public Economics* 99 (November), pp. 235-262.

Currie, Janet and Duncan Thomas. 1995. "Does Head Start Make a Difference?" *American Economic Review* 85 (June), pp. 341-364.

Daleiden, Eric and Steve DeBois. 2001. "School Readiness: Early Childhood Skills Inventory (ECSI) Validation Study Report." Presented to the Community Service Council of Greater Tulsa. Tulsa: Department of Psychology, University of Tulsa, July 31.

Diamond, Laura. 2002. "Pre-K Plan Raises Issues," *Florida Times-Union*, November 8, p. 1.

Garces, Eliana, Duncan Thomas, and Janet Currie. 2002. "Longer Term Effects of Head Start," *American Economic* Review 92 (September), pp. 999-1012.

Gilliam, Walter and Ed Zigler. 2001. "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, pp. 441-473.

Head Start Bureau. 2003. Statistical Fact Sheet for the Fiscal Year 2002. Washington, D.C.: Administration on Children and Families, U.S. Department of Health and Human Services, June 2.

Henry, Gary, Craig Gordon, Andrew Mashburn, and Bentley Ponder. 2001. "Pre-K Longitudinal Study: Findings from the 1999-2000 School Year." Atlanta, GA: Georgia State University, Applied Research Center.

Henry, Gary, Laura Henderson, Bentley Ponder, Craig Gordon, Andrew Mashburn, and Dana Rickman. 2003. "Report of the Findings from the Early Childhood Study: 2001-02." Atlanta: Andrew Young School of Policy Studies, Georgia State University.

Hoxby, Caroline M. 2000. "The Effects of Class Size on Student Achievement: New Evidence from Population Variation." *Quarterly Journal of Economics* 115(4): 1239-1286.

Love, John. 2001. "Instrumentation for State Readiness Assessment: Issues in Measuring Children's Early Development and Learning." Paper presented at symposium on Assessing

the State of State Assessments, Atlanta, Georgia, December 12-14. Available through

www.mathematica-mpr.com/PDFs/assapaper.pdf.

Ludwig, Jens, G.J. Duncan, and P. Hirschfield. 2001. "Urban Poverty and Juvenile Crime:

Evidence from a Randomized Housing-Mobility Experiment." *Quarterly Journal of

Economics* 116(2): 655-680.

Ludwig, Jens, H.F. Ladd, and G.J. Duncan. 2001. "Urban Poverty and Educational Outcomes."

Brookings-Wharton Papers on Urban Affairs (The Brookings Institute: Washington, D.C.).

Lytal, Gary. Tulsa Public Schools. 2003. Telephone interview, March 17.

Magnuson, Katherine, Marcia Meyers, Christopher Ruhm, and Jane Waldfogel. Forthcoming

2004. "Inequality in Preschool Education and School Readiness." *American Educational

Research Journal*.

Masse, Leonard and W. Steven Barnett. 2003. "A Benefit Cost Analysis of the Abecedarian

Early Childhood Intervention." Unpublished manuscript. New Brunswick, N.J.: National

Institute for Early Education Research.

McKey, Ruth, Larry Condell, Harriett Ganson, Barbara Barrett, Catherine McConkey, and

Margaret Planz. 1985. "The Impact of Head Start on Children, Families and Communities:

Final Report of the Head Start Evaluation, Synthesis and Utilization Project." Washington,

D.C.: CSR, Inc.

Reynolds, Arthur, Judy Temple, Dylan Robertson, and Emily Mann. 2001. "Long-term Effects

of an Early Childhood Intervention on Educational Achievement and Juvenile Arrest,"

Journal of the American Medical Association 285 (May 9), pp. 2339-2346.

Sacerdote, Bruce. 2001. "Peer Effects with Random Assignment: Results for Dartmouth

Roommates." *Quarterly Journal of Economics* 116(2): 681-704.

Wade, Linda. Tulsa Public Schools. 2003. Personal interview, March 13.

Wetzstein, Cheryl. 2002. "The Preschool Push," *Washington Times*, November 26, p.2.

Xiang, Zongping and Lawrence Schweinhart. 2002. "Effects Five Years Later: The Michigan

    School Readiness Program Evaluation Through Age 10." Paper prepared for the Michigan

    State Board of Education. Ypsilanti, Mich.: High/Scope Educational Research Foundation,

    January 23.

APPENDIX A: SURVEY INSTRUMENT

Teacher Name_____

School_____

Test ID# (optional)_____

EARLY CHILDHOOD SKILLS INVENTORY
SCORING SHEET

See "Guidelines for Administration" for item descriptions and complete scoring criteria.

TODAY'S DATE_____ BIRTH DATE_____ AGE_____ MALE/FEMALE_____

CHECK IF COMPLETED:
VISION__ HEARING__ ECSI PARENT FORM__ IMMUNIZATION RECORD__

SCORE '0' OR '1' FOR EACH ITEM
Total scores for each section are the sum of item scores in that section.

SOCIAL/EMOTIONAL

1.   Separates from caregiver……………… YES___ NO___ NOT OBSERVED___
       (Do not include in total Social/Emotional score)

2.   Responds to greeting appropriately…………………………………….   _____

3.   Displays appropriate behavior for situation (Assess at end of test)…….   _____

4.   Displays adequate attention span (Assess at end of test)………………   _____

SOCIAL/EMOTIONAL SCORE_____/3

COGNITIVE/GENERAL KNOWLEDGE

1.   Names basic colors (circle correct responses)…..# correct___   Score (if>=6)_____
       Red   Blue   Green   Yellow Orange Purple Black   Brown

2.   Names basic shapes (circle correct responses)….# correct___   Score (if>=3)_____
       Circle   Square   Triangle Rectangle

3.   Identifies numerals (circle correct responses)…..# correct___   Score (if>=5)_____
       8   5   7   1   3   10   9   2   4   6

4.   Names objects and tells meaning of following words:……….   Score (if>=7)_____
                                          Name   Meaning
                              Apple   _____ _____
                              Book   _____ _____
                              Pencil _____ _____
                              Car   _____ _____

5.   Counts out 5 objects…………………………………………   Score (if=5)_____

6.   Sorts pictures by category……………………………………..   Score (if all)_____

7.   Recalls digit sets (circle each group child recalls correctly)   Score (if>=6)_____
       Practice Item   Set 1:   2-9   4-6   3-1
           5-3   Set 2:   3-8-6   6-1-2   4-1-7
                               Set 3:   8-2-4-1 5-3-7-2 6-9-3-1

COGNITIVE SCORE_____/7

ECSI SCORING SHEET

<u>MOTOR SKILLS</u>

Gross

1.   Stands on one foot (for 5 seconds)………………………………………   _____

2.   Hops on one foot (5 times without touching)…………………………   _____

Fine

1.   Copies………………………………..………O + □ ∇ Score (if>=3)   _____

2.   Colors circle with control………………………………………..   _____

3.   Prints first name……………………………………………………   _____

4.   Uses scissors to cut on a straight line…………………………………   _____

RIGHT HANDED ____   *LEFT HANDED* ____   **MOTOR SKILLS SCORE**_____/6

<u>LANGUAGE</u>

1.   When asked, "Do you see grown-ups reading at home?" child answers: **No   Yes**
        (Do not include in total Language score)

2.   States first and last name when asked………………………………   _____

3.   States age and birthday…………………………………………..   _____

4.   Repeats an 8 word sentence………………………………………...   _____
     (*Practice item:* <u>The bird sings.</u>)   *The big dog barked at the gray cat.*

5.   Speaks in complete sentences when describing a picture…………….   _____

6.   Able to follow a series of 3 directions after being told once…………   _____

7.   Letter recognition…… **R S T L E** ………………………… Score (if>=4)   _____

8.   Demonstrates knowledge of books.   *Opens book in proper orientation*   _____
                                          Correctly identifies first word on page   _____

9.   Identifies rhyming words………………………….……… Score (if <u>all</u>)   _____
          Practice.   *1. fit—sit*
          Cat—bat   *2. glad—dog*
          Sock—ball   *3. honey—money*

10. Produces rhyming words………………………………… ...Score (if>=2)   _____
          Practice: nose   *1. car   2. sing   3. tire*

                                          LANGUAGE SCORE_____/10

                                          TOTAL SCORE   _____/26

<u>ADDITIONAL INFORMATION</u> CHECK ANY THAT APPLY:
(Not Included in Total Score)

____ PHYSICAL HANDICAP   ____ SPEECH/LANGUAGE PROBLEMS (Need Referral? YES NO)
____ ATTENTION PROBLEM   ____ VISUALLY IMPAIRED   ____ HEARING IMPAIRED

OTHER COMMENTS OR CONCERNS: _____

APPENDIX B: The Effect of Being Born Before Cut-Off Date on Test Scores:
Quadratic Parametric Fit

(Kindergarten Children are Conditional on Having Been in Tulsa Pre-kindergarten)

	Total	Soc/Emt	Cog/Know	Motor	Language
|Margin|<=1 Year	1.994 [a]	-0.043	0.756 [a]	0.413 [b]	0.817 [a]
	(0.557)	(0.082)	(0.197)	(0.175)	(0.259)
	N=2246	N=2286	N=2298	N=2283	N=2280
|Margin|<=9 Months	1.929 [a]	-0.011	0.750 [a]	0.373 [c]	0.724 [b]
	(0.647)	(0.097)	(0.229)	(0.203)	(0.301)
	N=1633	N=1659	N=1674	N=1661	N=1660
|Margin|<=6 Months	2.192 [a]	0.000	0.822 [a]	0.455 [c]	0.846 [b]
	(0.791)	(0.124)	(0.278)	(0.249)	(0.368)
	N=1084	N=1099	N=1111	N=1102	N=1101
|Margin|<=3 Months	2.348 [b]	-0.039	1.050 [a]	0.616	0.627
	(1.193)	(0.194)	(0.409)	(0.377)	(0.548)
	N=559	N=568	N=578	N=572	N=572

Notes: Standard errors in parentheses. Each of the rows represents a different set of regressions, each pertaining to a different margin of the data, as indicated in the row headings. Statistical significance at the 1%, 5%, and 10% levels (2-tailed) are are denoted by "a," "b," and "c," respectively.

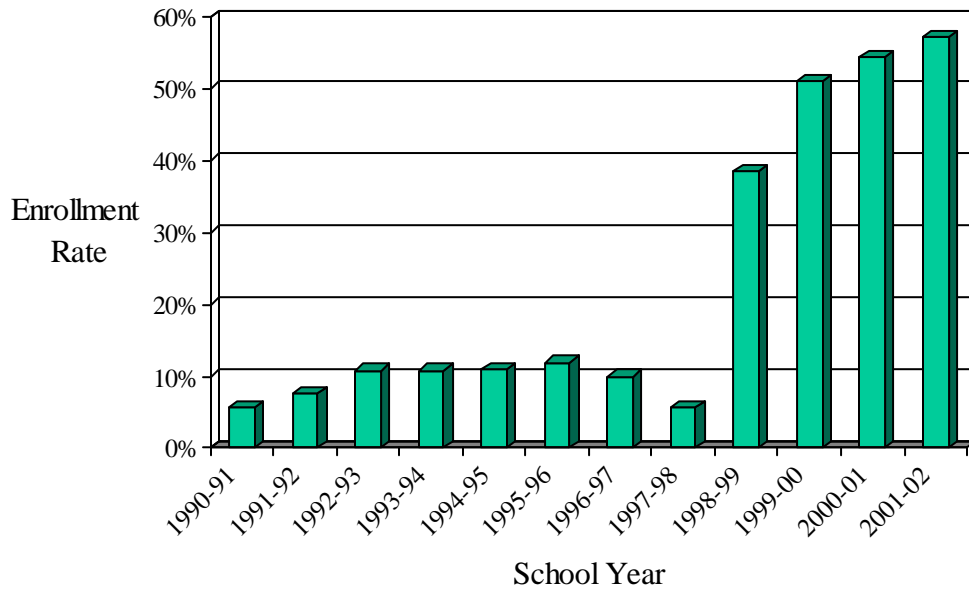**Figure I: Oklahoma Pre-K Enrollment Rate, By Year**

Figure II:

The Relationship between Kindergarten Enrollment in 2001 and Margin of Birthday from Cut-off Date
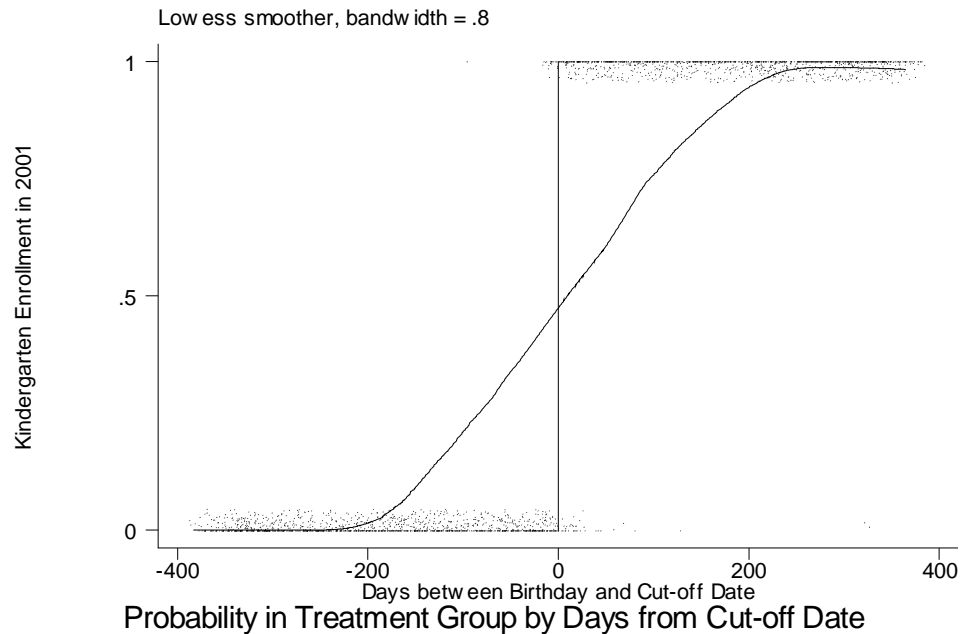
Probability in Treatment Group by Days from Cut-off Date

Notes: The data points to the left of the cut-off date consist of the "young" children, who are those who qualified for TPS pre-k in 2001. These children make up our control group because they selected into the program but (as of the test time of August 2001) have yet to experience the treatment. The data points to the right of the cut-off date consist of the "old" children, who are those who qualified and enrolled in TPS pre-k in 2000. These children make up our treatment group because they have already received the TPS pre-k education in 2000 and (as of the test time of August 2001) are just beginning TPS kindergarten. There are 18 children who qualified for TPS pre-k in 2000 but were in TPS pre-k in 2001, and there are 4 children (3 born right on the cut-off date) who qualified for TPS pre-k in 2001 but were in TPS kindergarten in 2001. These 22 observations make for a "fuzzy" discontinuity.

Figure III:

Scatter Plot of Test Scores with Fitted Cubic Splines
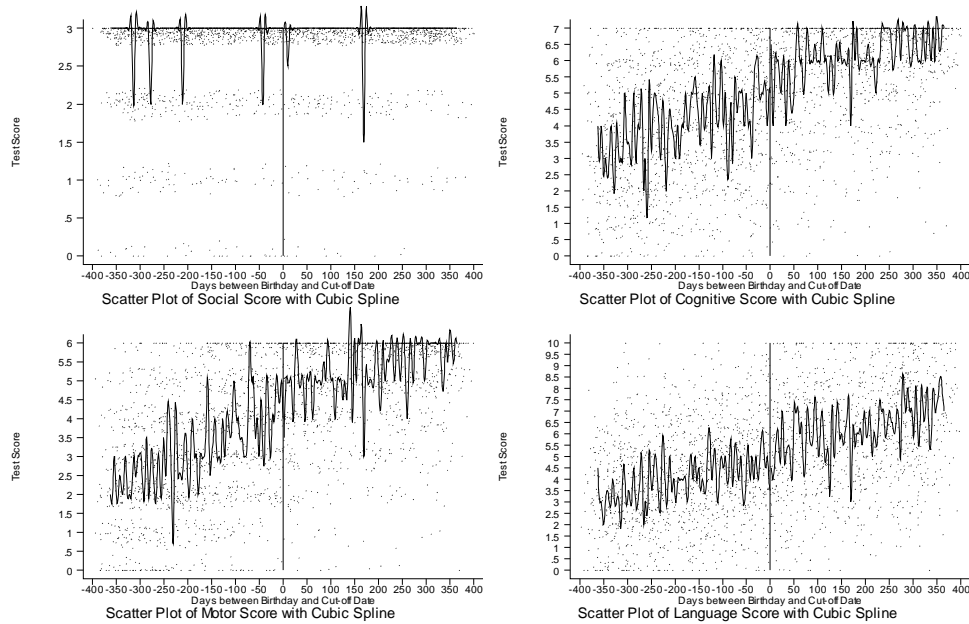
Scatter Plot of Social Score with Cubic Spline

Scatter Plot of Cognitive Score with Cubic Spline

Scatter Plot of Motor Score with Cubic Spline

Scatter Plot of Language Score with Cubic Spline

Notes: The data points to the left of the cut-off date consist of the "young" children, who are those who qualified for TPS pre-k in 2001. These children make up our control group because they selected into the program but (as of the test time of August 2001) have yet to experience the treatment. The data points to the right of the cut-off date consist of the "old" children, who are those who qualified and enrolled in TPS pre-k in 2000. These children make up our treatment group because they have already received the TPS pre-k education in 2000 and (as of the test time of August 2001) are just beginning TPS kindergarten.

Figure IV:

Predicted Test Scores by Margin of Birthday from Cut-off Date (Quadratic Parametric Fit)
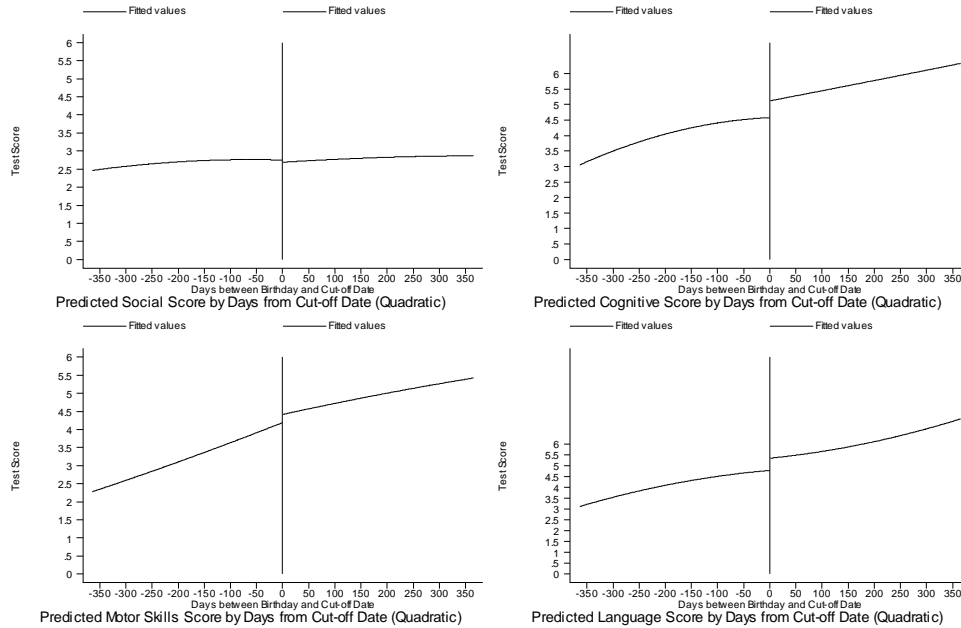

Notes: The data points to the left of the cut-off date consist of the "young" children, who are those who qualified for TPS pre-k in 2001. These children make up our control group because they selected into the program but (as of the test time of August 2001) have yet to experience the treatment. The data points to the right of the cut-off date consist of the "old" children, who are those who qualified and enrolled in TPS pre-k in 2000. These children make up our treatment group because they have already received the TPS pre-k education in 2000 and (as of the test time of August 2001) are just beginning TPS kindergarten.

Table I:
Comparison of Tested Children to Universe of Children

Variable	Pre-K Tested Children		Pre-K Universe		Diff.	Kindergarten Tested Children		Kindergarten Universe		Diff.
Female	0.519		0.499		-0.020	0.486		0.479		-0.007
	(0.500)	N=1284	(0.500)	N=1690	P=0.2800	(0.500)	N=2276	(0.500)	N=3441	P=0.6044
White	0.406		0.407		0.001	0.448		0.421		-0.027
	(0.491)	N=1284	(0.491)	N=1690	P=0.9561	(0.497)	N=2275	(0.494)	N=3441	P=0.0437
Black	0.389		0.370		-0.019	0.290		0.317		0.027
	(0.488)	N=1284	(0.483)	N=1690	P=0.2902	(0.454)	N=2275	(0.465)	N=3441	P=0.0301
Hispanic	0.103		0.124		0.021	0.151		0.153		0.002
	(0.304)	N=1284	(0.329)	N=1690	P=0.0750	(0.358)	N=2275	(0.360)	N=3441	P=0.8368
Native American	0.084		0.083		-0.001	0.096		0.095		-0.001
	(0.278)	N=1284	(0.276)	N=1690	P=0.9223	(0.295)	N=2275	(0.294)	N=3441	P=0.9000
Asian	0.019		0.016		-0.003	0.015		0.013		-0.002
	(0.135)	N=1284	(0.125)	N=1690	P=0.5312	(0.120)	N=2275	(0.115)	N=3441	P=0.5271
Free Lunch	0.466		0.460		-0.006	0.513		0.529		0.016
	(0.499)	N=1265	(0.499)	N=1822	P=0.7425	(0.500)	N=2239	(0.499)	N=3442	P=0.2380
Reduced Price Lunch	0.118		0.109		-0.009	0.084		0.087		0.003
	(0.322)	N=1265	(0.311)	N=1822	P=0.4358	(0.277)	N=2239	(0.281)	N=3442	P=0.6925
Full Price Lunch	0.416		0.431		0.015	0.404		0.384		-0.020
	(0.493)	N=1265	(0.495)	N=1822	P=0.4070	(0.491)	N=2239	(0.486)	N=3442	P=0.1312

Notes: Standard deviations in parentheses. The tested children are those children who were tested and whose test results were intelligble. Most children were tested in August 2001. The universe of children represents those children who were enrolled in TPS as of October 2001. Neither the tested children nor the universe of children for the pre-k cohort includes children enrolled in Head Start programs, some of which collaborate with TPS.

Table II:
Comparison of Mean Test Scores and Covariates for K Children
Who Were in TPS Pre-K vs. K Children Who Were Not in TPS Pre-K

	|Margin|<=1 year			
Variable	**No Pre-K**		**Pre-K**		**Difference**
Total Test Score	18.587		19.685		1.098
	(0.168)	N=974	(0.136)	N=1044	P=0.0000
Soc./Emt. Score	2.730		2.806		0.075
	(0.020)	N=990	(0.017)	N=1076	P=0.0037
Cog./Know. Score	5.411		5.754		0.343
	(0.055)	N=992	(0.046)	N=1075	P=0.0000
Motor Skills Score	4.761		4.979		0.219
	(0.046)	N=986	(0.039)	N=1063	P=0.0003
Language Score	5.765		6.143		0.378
	(0.083)	N=986	(0.072)	N=1059	P=0.0006
Female	0.500		0.493		-0.006
	(0.016)	N=1007	(0.015)	N=1097	P=0.7715
Full Price Lunch	0.491		0.335		-0.156
	(0.016)	N=985	(0.014)	N=1082	P=0.0000
Reduced Price Lunch	0.078		0.091		0.012
	(0.009)	N=985	(0.009)	N=1082	P=0.3120
Free Lunch	0.430		0.574		0.143
	(0.016)	N=985	(0.015)	N=1082	P=0.0000
Non-White	0.448		0.655		0.207
	(0.016)	N=1007	(0.014)	N=1093	P=0.0000
White	0.552		0.345		-0.207
	(0.016)	N=1007	(0.014)	N=1093	P=0.0000
Black	0.179		0.402		0.223
	(0.012)	N=1007	(0.015)	N=1093	P=0.0000
Hispanic	0.151		0.151		0.000
	(0.011)	N=1007	(0.011)	N=1093	P=0.9991
Native American	0.105		0.085		-0.020
	(0.010)	N=1007	(0.008)	N=1093	P=0.1149
Asian	0.012		0.017		0.005
	(0.003)	N=1007	(0.004)	N=1093	P=0.2996

T Notes: Standard errors in parentheses.

Table 3: Comparison of Mean Test Scores and Covariates Before and After Cut-Off Birth Date
Kindergarten Classes are conditional on having been in TPS Pre-K.

Variable	|Margin|<=1 year			|Margin|<=6 months			|Margin|<=3 months			Quadratic Parametric Fit		
	After 9/1 (Control)	Before 9/1 (Treatment)	Diff.	After 9/1 (Control)	Before 9/1 (Treatment)	Diff.	After 9/1 (Control)	Before 9/1 (Treatment)	Diff.	After 9/1 (Control)	Before 9/1 (Treatment)	Diff.
Total Test Score	13.950	19.685	5.735	15.372	18.626	3.254	15.833	18.254	2.421	16.269	17.625	1.355
	(0.150) N=1229	(0.136) N=1044	P=0.0000	(0.207) N=608	(0.208) N=486	P=0.0000	(0.304) N=293	(0.293) N=268	P=0.000	(0.412) N=1233	(0.389) N=1062	P=0.0169
Soc./Emt. Score	2.674	2.806	0.132	2.750	2.757	0.007	2.766	2.731	-0.035	2.749	2.688	-0.062
	(0.020) N=1237	(0.017) N=1076	P=0.0000	(0.026) N=612	(0.027) N=497	P=0.8643	(0.037) N=295	(0.039) N=275	P=.5087	(0.057) N=1241	(0.049) N=1094	P=0.4082
Cog./Know. Score	4.006	5.754	1.748	4.443	5.460	1.018	4.465	5.294	0.829	4.577	5.118	0.541
	(0.055) N=1252	(0.046) N=1075	P=0.0000	(0.076) N=619	(0.072) N=502	P=0.0000	(0.111) N=301	(0.100) N=279	P=0.0000	(0.153) N=1256	(0.132) N=1093	P=0.0075
Motor Skills Score	3.199	4.979	1.780	3.686	4.741	1.056	3.937	4.653	0.717	4.192	4.417	0.225
	(0.048) N=1249	(0.039) N=1063	P=0.0000	(0.067) N=617	(0.059) N=495	P=0.0000	(0.095) N=300	(0.080) N=274	P=0.0000	(0.129) N=1253	(0.114) N=1081	P=0.1925
Language Score	4.090	6.143	2.052	4.532	5.688	1.157	4.677	5.540	0.863	4.781	5.342	0.561
	(0.061) N=1250	(0.072) N=1059	P=0.0000	(0.087) N=617	(0.107) N=494	P=0.0000	(0.129) N=300	(0.147) N=274	P=0.0000	(0.171) N=1254	(0.206) N=1077	P=0.0357
Female	0.522	0.494	-0.028	0.522	0.513	-0.009	0.521	0.491	-0.030	0.543	0.475	-0.068
	(0.014) N=1261	(0.015) N=1093	P=0.1793	(0.020) N=621	(0.022) N=507	P=0.7658	(0.029) N=303	(0.030) N=283	P=0.4645	(0.040) N=1265	(0.044) N=1111	P=0.2546
Full Price Lunch	0.415	0.335	-0.080	0.043	0.329	-0.105	0.415	0.322	-0.094	0.419	0.371	-0.048
	(0.014) N=1243	(0.014) N=1082	P=0.0001	(0.020) N=614	(0.021) N=504	P=0.0003	(0.028) N=301	(0.028) N=283	P=0.0190	(0.040) N=1246	(0.042) N=1099	P=0.4021
Reduced Price Lunch	0.117	0.091	-0.027	0.130	0.085	-0.045	0.136	0.106	-0.030	0.117	0.108	-0.009
	(0.009) N=1243	(0.009) N=1082	P=0.0349	(0.014) N=614	(0.012) N=504	P=0.0168	(0.020) N=301	(0.018) N=283	P=0.2651	(0.026) N=1246	(0.025) N=1099	P=0.8090
Free Lunch	0.467	0.574	0.107	0.435	0.585	0.150	0.449	0.572	0.124	0.463	0.521	0.057
	(0.014) N=1243	(0.015) N=1082	P=0.0000	(0.020) N=614	(0.022) N=504	P=0.0000	(0.029) N=301	(0.029) N=283	P=0.0027	(0.040) N=1246	(0.044) N=1099	P=0.3361
Non-White	0.600	0.655	0.055	0.596	0.651	0.055	0.617	0.654	0.037	0.621	0.620	0.000
	(0.014) N=1261	(0.014) N=1093	P=0.0062	(0.020) N=621	(0.021) N=507	P=0.0580	(0.028) N=303	(0.028) N=283	P=0.3593	(0.039) N=1265	(0.042) N=1111	P=0.9952
White	0.400	0.345	-0.055	0.404	0.349	-0.055	0.383	0.346	-0.037	0.379	0.380	0.000
	(0.014) N=1261	(0.014) N=1093	P=0.0062	(0.020) N=621	(0.021) N=507	P=0.0580	(0.028) N=303	(0.028) N=283	P=0.3593	(0.039) N=1265	(0.042) N=1111	P=0.9952
Black	0.393	0.402	0.008	0.380	0.383	0.003	0.396	0.392	-0.004	0.380	0.349	-0.031
	(0.014) N=1261	(0.015) N=1093	P=0.6813	(0.019) N=621	(0.022) N=507	P=0.9285	(0.028) N=303	(0.029) N=283	P=0.92550	(0.039) N=1265	(0.043) N=1111	P=0.5938
Hispanic	0.103	0.151	0.048	0.114	0.168	0.053	0.125	0.177	0.051	0.142	0.175	0.034
	(0.009) N=1261	(0.011) N=1093	P=0.0005	(0.013) N=621	(0.017) N=507	P=0.0098	(0.019) N=303	(0.023) N=283	P=0.0828	(0.025) N=1265	(0.032) N=1111	P=0.3997
Native American	0.085	0.085	0.000	0.084	0.089	0.005	0.086	0.074	-0.012	0.096	0.093	-0.003
	(0.008) N=1261	(0.008) N=1093	P=0.9838	(0.011) N=621	(0.013) N=507	P=0.7650	(0.016) N=303	(0.016) N=283	P=0.6060	(0.023) N=1265	(0.024) N=1111	P=0.9227
Asian	0.019	0.017	-0.002	0.018	0.012	-0.006	0.010	0.011	0.001	0.003	0.003	0.000
	(0.004) N=1261	(0.004) N=1093	P=0.7658	(0.005) N=621	(0.005) N=507	P=0.4206	(0.006) N=303	(0.006) N=283	P=0.9331	(0.011) N=1265	(0.011) N=1111	P=0.9807
Read at Home	0.858	0.813	-0.046	0.837	0.820	-0.017	0.834	0.810	-0.024	0.839	0.830	-0.009
	(0.011) N=1093	(0.013) N=955	P=0.0053	(0.016) N=546	(0.018) N=444	P=0.4755	(0.023) N=271	(0.025) N=242	P=0.4779	(0.985) N=1096	(1.144) N=970	P=0.8569

Notes: Standard errors in parentheses.

Table IV:
The Effect of TPS Pre-K on Test Scores: Quadratic Parametric Fit

Variable	Total	|Margin|<=1 Year Soc/Emt	Cog/Know	Motor	Language
Born Before Cut-Off (Treated)	1.994 [a]	-0.043	0.756 [a]	0.413 [b]	0.817 [a]
	(0.557)	(0.082)	(0.197)	(0.175)	(0.259)
Qualify (days)	0.009 [c]	-3.9E-04	0.001	0.006 [a]	0.002
	(0.005)	(0.001)	(0.002)	(0.002)	(0.002)
Qualify^2	-1.5E-05	-3.2E-06	-7.9E-06	2.0E-06	-6.5E-06
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Qualify * Cut-Off	-0.003	0.001	0.002	-0.004 [c]	-5.3E-04
	(0.007)	(0.001)	(0.002)	(0.002)	(0.003)
Qualify^2 * Cut-Off	2.7E-05	2.2E-06	7.9E-06	-1.3E-07	1.5E-05 [c]
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female	1.593 [a]	0.153 [a]	0.255 [a]	0.597 [a]	0.560 [a]
	(0.187)	(0.026)	(0.068)	(0.059)	(0.086)
Black	-0.319	-0.026	-0.309 [a]	-0.063	0.064
	(0.227)	(0.032)	(0.081)	(0.074)	(0.103)
Hispanic	-3.651 [a]	-0.095 [b]	-1.427 [a]	-0.203 [b]	-1.943 [a]
	(0.322)	(0.047)	(0.129)	(0.097)	(0.141)
Native American	-0.056	-0.009	-0.099	-0.079	0.064
	(0.350)	(0.048)	(0.131)	(0.108)	(0.165)
Asian	-1.496 [c]	0.126 [b]	-0.569 [c]	0.353	-1.408 [a]
	(0.783)	(0.058)	(0.308)	(0.231)	(0.412)
Reduced Price Lunch	-0.606 [c]	0.005	-0.186	-0.067	-0.351 [b]
	(0.319)	(0.044)	(0.118)	(0.106)	(0.146)
Free Lunch	-1.685 [a]	-0.053 [c]	-0.523 [a]	-0.313 [a]	-0.753 [a]
	(0.219)	(0.030)	(0.079)	(0.069)	(0.100)
Number of Obs.	2246	2286	2298	2283	2280

Notes: Robust standard errors in parentheses. Statistical significance at the 1%, 5%, and 10% levels (2-tailed) are denoted by "a," "b," and "c," respectively.
Kindergarten Children are Conditional on Having Been in TPS Pre-kindergarten.

Table V
The Effect of TPS Pre-K on Test Scores by Race and Half-/Full-Day: Quadratic Parametric Fit

Total	Soc/Emt	|Margin|<=1 Year Cog/Know	Motor	Language
		Hispanic, Half-Day Children		
2.013 [b]	-0.153	0.825	0.294	1.119 [b]
(2.286)	(0.479)	(0.824)	(0.635)	(0.902)
N=140	N=142	N=143	N=143	N=142
		Hispanic, Full-Day Children		
6.435 [a]	0.176	2.757 [a]	1.688 [b]	1.801 [b]
(2.385)	(0.424)	(0.892)	(0.760)	(0.908)
N=131	N=132	N=138	N=137	N=139
		Black, Half-Day Children		
2.046	-0.585	0.540	0.962	1.194
(2.056)	(0.269)	(0.770)	(0.709)	(1.010)
N=145	N=148	N=148	N=144	N=146
		Black, Full-Day Children		
2.599 [a]	-0.032	1.271 [a]	0.318	0.962 [b]
(1.003)	(0.160)	(0.336)	(0.317)	(0.468)
N=745	N=760	N=763	N=762	N=749
		White, Half-Day Children		
1.657 [c]	-0.094	0.370	0.187	1.020 [b]
(0.972)	(0.154)	(0.314)	(0.323)	(0.480)
N=610	N=622	N=616	N=612	N=616
		White, Full-Day Children		
0.294	0.038	-0.208	0.561	-0.030
(1.748)	(0.189)	(0.617)	(0.610)	(0.861)
N=217	N=221	N=229	N=226	N=227

Notes: Robust standard errors in parentheses. Each row represents a different set of regressions, each pertaining to a different race by half-day/full-day designation. Each regression estimation includes the relevant covariates from Table 4. Statistical significance at the 1%, 5%, and 10% levels (2-tailed) are denoted by "a," "b," and "c," respectively. Kindergarten children are conditional on having been in TPS Pre-k.

Notes:

i Georgia's universal program replaced a means-tested program established in 1993.

ii New York's program, though established in 1997, was not intended to become universal until 2002. Because of state budget shortages, only 200 out of 700 school districts were participating in 2002. Governor George Pataki proposed discontinuing the pre-k program in 2003, but the New York state legislature voted to continue the program, overriding a gubernatorial veto.

iii In August 2002, Los Angeles County created a universal preschool program using revenue from tobacco taxes (Wetzstein, 2002). In November 2002, Florida voters approved an initiative requiring state officials to establish a universal pre-k program by 2005 (Diamond ,2002).

iv In 2001, the cost per child for Head Start was $6,934 (Head Start Bureau 2003). Currie (2001) estimates that Head Start's costs were 71 percent of the Perry Preschool Project's costs, which would make the Perry Preschool Project's costs $9,763 in 2001 dollars. Masse and Barnett (2003) have estimated that the Abecedarian project cost $13,900 per child in 2002 dollars.

v A New York evaluation compared program participants with wait-listed children; this is an appropriate technique for controlling for differences in parental motivation. Unfortunately, that evaluation, published in 1977, was also the oldest of those reviewed (Gilliam and Zigler, 2001, p. 452).

vi The testing instrument (see Appendix A) used in August 2001 correlates rather well with nationally-normed tests, including the Brigance Screens (0.85) and the Battelle Developmental Inventory (0.80) (Daleiden and DeBois, 2001, p.8). Performance on the Battelle Developmental Inventory among preschoolers correlates well with subsequent performance on the Woodcock-Johnson Achievement Test in elementary school (Berls and McEwen 1999). In addition, we were able to establish substantial congruence between the Tulsa testing instrument and Woodcock-Johnson (Working Paper # 1, http://www.crocus.georgetown.edu).

vii Due to scheduling challenges, 19 percent of the TPS pre-k children and 23 percent of the TPS kindergarten children were tested in a month other than August 2001: approximately half of them were tested in July 2001.

viii The pre-k statistics do not include children enrolled in Head Start programs that "collaborate" with TPS.

ix In a very small number of cases, TPS pre-k deviates from this general rule. For example, a special education child who needs speech therapy may be admitted to pre-k ahead of schedule (Wade, 2003). Also, three parents lied about their children's ages in order to get them enrolled in a TPS pre-k class ahead of schedule (Lytal, 2003). Once this was discovered, the children were asked to leave.

x A naïve regression comparing kindergarten children who were in TPS pre-k to kindergarten children who were not in TPS pre-k results in estimated test impacts lower than in the regression-discontinuity estimations reported later (see Working Paper # 1, http://www.crocus.georgetown.edu).

xi A naïve regression comparing kindergarten children who were in TPS pre-k to current TPS pre-k children results in estimated test impacts lower than in the regression-discontinuity estimations reported later (see Working Paper # 1, http://www.crocus.georgetown.edu).

xii The curve fitted through the points in Figure 2 is estimated using a lowess smoother with a band width of 0.8. All the points in Figure 2 have a y-axis value of either zero or one. Since these points would be on top of one another, we added spherical random noise to the data in the figure to make it easier to see each point.

xiii In a separate analysis (Working Paper # 1, http://www.crocus.georgetown.edu), we include the aberrational observations in the full sample and conduct an instrumental-variables estimation. The findings are very similar to the analysis in this paper, which is expected given the very small number of aberrational observations.

xiv The categorical nature of the test data leads to plotting several points on top of each other. In order to make the graphs more readable, we added spherical random noise to each point before graphing.

xv The last row of Table 3 compares the mean values for the treatment and control groups for whether the child sees an adult reading at home. We find no statistically significant difference in this measure between the treatment group and the control group at the discontinuity, which is what one would expect in a randomized framework. Because this variable is missing values for many children, we did not include it in our later regressions.

xvi If, instead, we estimate natural logs (rather than levels) of the test scores, we find that TPS pre-k leads to a 17.2 percent increase in cognitive/knowledge scores and a 16.5 percent increase in language scores (both significant at the 1-percent level), as well as an 8.4 percent increase in motor skills (significant at the 10-percent level).

xvii Although we rely on the quadratic results throughout this paper, in another draft we conducted analyses for different functional forms (Working Paper # 1, http://www.crocus.georgetown.edu). We found that the results are robust across linear, quadratic, cubic, and quartic specifications. We also conducted estimations of these alternative

functional forms in which the sample was restricted to birthdays within six months of the cut-off birth date. This analysis found very similar point estimates to the results with the one-year margin; however, the reduction in degrees of freedom did increase the standard errors.

xviii This is not uncommon. As Love (2001: 6) has noted when discussing the five school readiness goals of the National Education Goals Panel (physical well-being and motor development; social and emotional development; approaches toward learning; language development; and cognition), "The areas of social-emotional development and approaches toward learning are typically the least-well measured."

xix Among white children, 35% receive the maximum cognitive score. Quite possibly, some of these children would receive an even higher score if a more versatile testing instrument were used. In contrast, 24% of black children and 15% of Hispanic children receive the maximum cognitive score.

xx While the positive impacts of the program are higher for minority and low-income children, one cannot rule out the possibility that having children with higher socio-economic status in TPS pre-k contributes to these higher impacts through spillover effects. For empirical evidence of such spillover effects, see Hoxby (2000); Ludwig, Duncan, and Hirschfield (2001); Ludwig, Ladd, and Duncan (2001); and Sacerdote (2001). Of course, it is possible that children with higher socio-economic status may experience negative spillover effects, in which case one would want to know whether the positive spillover effects for disadvantaged children outweigh the negative spillover effects for advantaged children.

xxi Figures supplied by Barbara Whisenhunt, Director, Budget Office, Tulsa Public Schools, October 22, 2003.

xxii TPS officials agree that the cost of educating a pre-K child is not as high as the cost of educating older children, because the latter require more expensive facilities and a wider variety of specialized personnel. The average per-pupil expenditure for TPS (all grades) was $6,088 for the 2002-3 school year.