The effects of Texas's Targeted Pre-Kindergarten Program on Academic Performance[1]

Rodney J. Andrews
The University of Texas at Dallas

Paul Jargowsky
Rutgers University–Camden

Kristin Kuhne
Communities Foundation of Texas

November 19, 2012

[1]We gratefully acknowledge that this research was made possible through data provided by the University of Texas at Dallas Education Research Center. The conclusions of this research do not necessarily reflect the opinions or official positions of the Texas Education Agency, the Texas Higher Education Coordinating Board, or the State of Texas.

Abstract

There has been a resurgence in research that investigates the efficacy of early investments as a means of reducing gaps in academic performance. However, the strongest evidence for these effects comes from experimental evaluations of small, highly enriched programs. We add to this literature by assessing the extent to which a large-scale public program, Texas's targeted pre-Kindergarten (pre-K), affects scores on math and reading acheivement tests, the likelihood of being retained in grade, and the probability that a student receives special education services. We find that having participated in Texas's targeted pre-K program is associated with increased scores on the math and reading sections of the Texas Assessment of Academic Skills (TAAS), reductions in the likelihood of being retained in grade, and reductions in the probability of receiving special education services. We also find that participating pre-K increases mathematics scores for students who take the Spanish version of the TAAS tests. These results show that even modest, public pre-K program implemented at scale can have important effects on students educational achievement.

Introduction

A number of recent papers—for example, Heckman and Masterov (2007) and Knudsen, Heckman, Cameron, and Shonkoff (2006)—strongly suggest that early investments in children are an effective means of reducing gaps in academic performance between disadvantaged children and their more advantaged counterparts. The estimates of the impacts obtained from the study of model programs, such as the Perry Preschool Program or the Carolina Abcedarian Project, have fueled the interest in the efficacy of early childhood investment. Heckman, Moon, Pinto, Savelyev, and Yavitz (2010) find that the social returns to the Perry Preschool Project are on the order of 7 to 10 percent, which is greater than the average return to equity and Anderson (2008) reports that the Abcedarian Project results in a .45 standard deviation increase for girls on a summary index of outcomes that include IQ, grade repetition, special ed., high school, college attendance, employment, earnings, receipt transfers, arrests, convictions, drug use, teen pregnancy and marriage.

The characteristics of these model programs—namely, random assignment and the magnitude of resources directed towards the treatment group—make them particularly amenable to study, but also limit the policy relevance of the findings. First, while random assignment bolsters internal validity, the small samples involved hinder the generalizability of the studies. The Perry Preschool Program and the Carolina Abcedarian projects started with small samples—123 children and 111 infants, respectively—of disadvantaged children in a single location.

Second, the treatment that the model programs offered are more intensive than the interventions offered by other early intervention programs. The Carolina Abcedarian project targeted infants with the treated children attending a preschool center for 8 hours per day, 5 days per week, 50 weeks per year until reaching schooling age, while the treated children from the Perry Preschool Program attended the program 5 mornings per week from October through May and received one 90-minute home visit per week. Given budget constraints, it is highly unlikely that any new public programs will approach these levels of investment. Relative to the model programs, the most prevalent existing early intervention programs—for example, Head Start and state funded pre-K programs—attempt to treat a broader audience and offer treatments that are not nearly as intense.

Recent research on the effects of the more moderate early intervention programs have used both a variety of data sources and identification strategies to investigate the effects of these programs on a number of outcomes. A number of papers use nationally representative data sets—such as the National Longitudinal Survey of Youth or the Panel Study of Income Dynamics. Currie and Thomas (1995) use the National Longitudinal Mother-Child supplement and exploit within family differences in Head Start participation to determine the effects of the program on a variety of outcomes. They find that Head Start increases test scores among blacks and whites, decreases the likelihood that a white child will be retained, and increases access to health services. Garces, Thomas, and Currie (2002) use the Panel Study of Income Dynamics and exploit within family variation in Head Start Attendance to determine the effects of Head Start participation on a number of later-life outcomes and find that, relative to the sibling who did not participate in Head Start, whites are more likely to complete high school, attend college, and have higher earnings in their early twenties, while for blacks the sibling who participated in Head Start is less likely to be charged with a crime. Deming (2009) uses the the National Longitudinal Mother-Child Supplement and, like Currie and Thomas (1995) and Garces et al. (2002), exploits within family difference in Head Start participation to estimate the effects of Head Start on a summary index of adult outcomes. He finds that Head Start participation results in a .23 standard deviation increase for the sibling who participated in Head Start. Puma et al. (2010) use a randomized control study to examine the effects of Head Start and find that Head Start participation increased the scores obtained in the first grade on the Peabody Picture Vocabulary Test for 4-year old participants and increased the scores on a test of oral comprehension for the 3-year old head start participants.

Gormley and Gaye (2005) use eligibility based on the date of birth in a regression discontinuity research design to estimate the effects of Tulsa's universal pre-K program. They find that Tulsa's pre-K program increased cognitive scores .39 standard deviations, motor skills by .24 standard deviations, and language scores by .38 standard deviations; moreover, the impacts are largest for Hispanics and blacks with little impact for whites. The children who are eligible for free lunch benefit more from pre-K than their more affluent peers. Fitzpatrick (2008) uses data from the National Assessment of Educational Progress in a difference-in-differences framework to evaluate Georgia's universal pre-K program. Using other states as a counterfactual, she finds that the availability of universal pre-K increases the math and reading scores at the fourth-grade level and increases the probability of students being on-grade for their age. Gormley and Gaye (2005) and Fitzpatrick (2008) are the most comparable to the research here as they consider locally sponsored early intervention programs that are similar to Texas's targeted pre-K program.

Texas began offering pre-K during the 1985–1986 academic year. The purpose of state-sponsored pre-K in Texas is to bolster the academic performance of at risk children. The risk factors include the following: free or reduced-price lunch eligibility, limited English proficiency, homelessness or unstable housing, foster care participation, or parents who are on active military duty or who have been injured or killed on duty. In 2011, Texas's pre-K program provided services for 6 percent of 3-year old children and 52 percent of 4-year old children, a total that exceeds 224,000 children, while Head start accounted for 8 percent of 3-year old children and 10 percent of 4-year old children (Barnett et al., 2011).

The Texas program is large and well established, but the program is not considered high-quality. The National Institute for Early Education Research (NIEER) ranks state pre-K programs on numerous criteria. The Texas program ranks low in terms of class size, staff-to-pupil ratios, and spending per capita (Barnett et al., 2011). As such, an evaluation of this program's impact on student outcomes can provide guidance on whether modest programs, perhaps the best that can hoped for in the current budgetary environment, are worth implementing.

We exploit the growth of the program over time, using differences in the availability of pre-K within districts over time to help identify the effects of pre-K on third grade math examinations, third grade reading examinations, retention in grade, and assignment to special education. If the change in the districts' offering of pre-K is unrelated to other factors that influence the outcomes under consideration, then our estimates have a causal interpretation.

We add to the literature that considers the effects of locally sponsored early intervention programs in several ways. First, as our analysis considers a large number of heterogenous school districts across the state of Texas, our results are more generalizable than the single-district results obtained in Gormley and Gaye (2005). Second, our use of a school district before it provides pre-K as the counterfactual is a more natural comparison relative to using other states as counterfactuals for Georgia as is done in Fitzpatrick (2008). Third, while other studies—for example, Gormley and Gaye (2005) and Currie and Thomas (1999)—analyze the effects of early interventions on the subset of Hispanic children who are fluent enough in English to be tested in English, we obtain results for both Hispanic children who are facile enough with English to take the English version of the examination and Hispanic children who take the Spanish version of the examination. Given the demographic changes that this country is experiencing, our ability to examine Hispanics of varying English ability increases the policy relevance of our research.

To preview results, we find that having participated in pre-K is associated with increased scores on the math and reading sections of the third grade version of the Texas Assessment of Academic Skills, reductions in the likelihood of being retained, and reductions in the probability of receiving special education services. We also find that participating in pre-K increases the math scores for students who take the Spanish version of TAAS.

The remainder of this paper is organized as follows. The second section describes the data. We present our empirical methodology in the third section. The fourth section discusses the results. The fifth section concludes.

Data

The study uses archival administrative data known as the Texas Schools Microdata Panel (TSMP) that is administered by the Texas Schools Project (TSP) located at the University of Texas at Dallas. This longitudinal panel consolidates individual level student data from several state agencies. The panel encompasses 13 years of individual data for more than 10 million students enrolled in Texas public schools between 1990 and 2002. Enrollment, attendance, test scores and other public school data is available for grades pre-K–12, along with key student demographics including age, ethnicity, language and economic status (TSP 2006).

Data is linked via encrypted personal identification numbers. This makes it possible to follow students, as long as they remain enrolled in a public school in Texas, throughout their academic career. Grade level and campus can be identified for each student by year; however, student-teacher links are not included in the data. Several TSMP files were combined to capture the student and district characteristics employed in the study. The primary source of data was the enrollment files from 1992–2002 and the TAAS files (Texas Assessment of Academic Skills) from 1997–2002. This data was appended with data characterizing the locale of Texas school districts from the Common Core of Data (CCD), a program of NCES under the auspice of the

United States Department of Education.

Available files allowed for the construction of five cohorts, capturing five years of treatment in a mature program. Children are not required to attend pre-K, so the first time we can observe both those who attended the state-funded pre-K and those who did not is when they attend kindergarten. Thus, cohorts are defined by the year a student first attended Kindergarten. We look two years back in the enrollment files to determine if the child was ever observed in pre-K. We then look forward to find the students' thrid grade test scores and information about retention in grade and special education placement. Data was not available to measure 3rd grade TAAS scores for both English and Spanish until 1997; therefore, the first cohort we can observe enrolled in kindergarten in 1994. The TAAS test was not given after 2002, so 1998 is the last available kindergarten cohort.

Data was not available to control for the educational experiences of the students who left and then re-enrolled prior to third grade. Therefore, students who were not continuously enrolled were excluded from the sample to limit treatment to Texas public schools. The sample is further limited to eligible students, since they are the target population for the program. Our determination of a student's eligibility for pre-K is based on eligibility for free and reduced price lunch and limited English proficiency in the kindergarten year. While it would be better to determine eligibility in the pre-K year, we do not observe these characteristics in the pre-K year for non-attenders, since they are not in the data. The degree of measurement error thus introduced is likely small, especially for limited English proficiency. Five year old children are not eligible to enroll in state-funded pre-K and if enrolled are considered ineligible for state funding since the program was specifically established to serve children under age five (Jones 2004). Based on this guideline, pre-K students who were five years or older were also excluded from the sample.

Thus, all students in our sample were eligible for the program, did not attend pre-K after age 5, did attend kindergarten, remained continuously enrolled in Texas public schools until the third grade, and took a standardized test that year. The sample includes $682,749$ students, 49 percent of all students in Texas attending kindergarten in 1994-1998. The large, heterogeneous sample reflects the ethnic, socioeconomic, and geographic diversity of the state, unlike the homogenous groups of participants found in studies of model programs.

Fifty-seven percent of these eligible students attended state funded pre-K. Seventy-five percent of these students were economically disadvantaged, and 30 percent had Limited English proficiency, and 5 percent were eligible for both reasons. The total sample pool is evenly divided across each cohort; and nearly 60 percent of the sample participated in pre-K as four year old children, but only 1 percent participated as three year old children.

Methodology

To evaluate the effects of the pre-K, we first compare students who attend the program with students who did not, controlling for as many covariates as possible. We examine five cohorts of kindergarten students who are either LEP, economically disadvantaged, or both – the target population for the program – from 1994–1998. This period was marked by a substantial growth in the Pre-K program.

Our base model for estimating the effect of Pre-K on student achievement is as follows:

$$Y_{icj} = \alpha + \beta_E PK + \beta_L PK * L + \beta_B PK * B + \beta_2' X_{ijc} + \gamma_c + \gamma_j + \varepsilon_{icj} \quad (1)$$

Y_{icj} is any outcome variable, such as a score on the reading section for the Texas Assessment of Academic skill for student i in cohort c from school district j.[1] α is a constant term, X_{icj} is a vector of individual, school, and district controls—for example, gender, socioeconomic status, whether the district is urban or rural, an indicator for whether full-day kindergarten is offered. X_{icj} also includes indicators for the reason for program eligbility: limited English proficiency only (L), or both limited English proficiency and economic disadvantage (B); eligibiity due to economic disadvantage (E) only is the reference category. γ_c is a cohort fixed effect that accounts for differences in across cohorts. γ_j is a district fixed effect that controls for fixed differences across districts. ε_{icj} is an idiosyncratic error term.

PK assumes a value of one if child i in cohort c in district j attended pre-K and zero otherwise. β_E is the difference between the mean score of eligible students who attended pre-K and those who did not, controlling for the covariates and fixed effects specified, for students who were eligible for

[1]The use of a test score is an example to fix ideas. The discussion that follows holds for other academic outcomes—for example, retention or assignment to special education status—that this research will explore. In the case of binary outcomes, we use estimate logistic regressions and linear probability models.

the program due to economic disadvantage only. By interacting the pre-K indicator with the reason for eligibility indicators (L and B), we allow the pre-K effect to vary by reason for program elibility. β_L and β_B indicate how the program effect varies from the reference group by reason for program eligibility. $\beta_E, \beta_E + \beta_L, and \beta_E + \beta_B$ are estimates of the effect of the program on those who participated in the program who were eligible due to economic disadvantage, limited English proficiency, or both, repectively.

This estimate may be subject to selection bias, however, if there is a systematic difference between those who enrolled and those who did not for which we have not controlled. Students are not required to attend pre-K. Families with eligible children choose to enroll their children in pre-K if that option is available to them. If families who enroll their children in the targeted pre-K program are systematically different in ways that the researcher can not observe and these differences are related to academic performance, then we can not assert that the pre-K program is the reason that the performance of participants and non-participants are different.

To the extent that the enrollment decision is based on whether the program is available in the family's school district, then enrollment is exogenous to the circumstances of individual children. When the program is available, then selection bias may occur. It is not possible to know a priori which direction this selection bias will operate. On the one hand, it is possible that the parents most interested in their child's education may seek out the public program. On the other hand, families with other potentially better options—a stay at home mother, a grandmother, private pre-K through a church, etc.—may opt out. Given that, by design, we have already controlled for economic disadvantage, LEP status, key individual covariates, cohort effects, and district fixed effects, there may be no systematic selection effects. Technically, as long as the attendance variable PK is uncorrelated with the disturbance term, the estimate of the program's effects are unbiased.

Nevertheless, as test of the robustness of our findings, we estimate a second set of models based solely on whether the student lived (in his or her kindergarten year) in a district that offered the pre-K program.[2] What is required is a source of variation in targeted pre-K enrollment that is orthogonal to ε_{icj}. We strongly curtail the potential for selection bias by estimating the Intent To Treat parameter (ITT). The ITT approach ignores take-up of

[2] Ideally, we would measure this in pre-K year, but we have no data prior to kindergarten on the location of students who did not attend pre-K.

the program and only estimates what happens to children who have been exposed to targeted pre-K in the sense that the program was available to them (Bloom, 1984). Thus, the ITT is not biased by selection at the family level. Consider the following model:

$$Y_{icj} = \alpha + \beta_{\tilde{E}} PO + \beta_{\tilde{L}} PO * L + \beta_{\tilde{B}} PO * B + \beta_2' X_{ijc} + \gamma_c + \gamma_j + \varepsilon_{icj} \quad (2)$$

Y_{icj}, α, X_{icj}, γ_c, γ_j, and ε_{icj} retain the definitions given above. PO is an indicator variable that assumes a value of one if a student is in a district that offers pre-K. $\beta_{\tilde{E}}$ represents what we can expect to happen to test scores for economically disadvantaged students if a district offers targeted pre-K regardless of who takes up the program. It is a weighted average of the effect of the program on those who enrolled and the effect of the program on those who did not.[3] Similarly, $\beta_{\tilde{L}}$ and $\beta_{\tilde{B}}$ are the differences in the effect of offering the program to those eligible for limited English proficency or both economic disadvantage and limited English proficiecy, respectively.

If the assumption that families who reside in a particular district can not willfully induce districts into offering pre-K holds, then this indicates that, conditional on X_{ijc} γ_c, and γ_j, PO is orthogonal to ε_{icj}, which implies that variation in program offering is exogenous to unmeasured student characteristics related to the outcome variable. This assumption is reasonable as it is unlikely that a given family with eligible children is able to intentionally alter the population of eligible children such that the district is compelled to provide targeted pre-K. Estimating the ITT models is a way to assess whether self-selection into the program at the family level has biased the program effects estimated based on those who selected to participate in the program.

As discussed below, the program was growing during the time period we study. Thus, given this variation in offering and our assumptions, we can obtain unbiased estimates of $\beta_{\tilde{E}}$, $\beta_{\tilde{L}}$ and $\beta_{\tilde{B}}$. These are conservative estimates of the effect of the program as they represent exposure to treatment and ignores consideration of who complies with the assignment to treatment, or as is the case here, we avoid having to consider why certain families elect to enroll their children in pre-K. Policy makers, however, are likely to be more interested in knowing the effect of targeted pre-K on children who actually enroll in the program.

[3] Those who did not enroll may still benefit from the program due to spillover effects in grades K-3, given that peer effects are well established in the education literature.

In effect, in the ITT model, the between-district and within-district variation in the availability of targeted pre-K is an instrument for enrolling in targeted pre-K. That is, these estimates only use the variation in the likelihood of enrolling in pre-K that is correlated with a district providing pre-K. If the variation in pre-K provision is uncorrelated with ε_{icj}, then we obtain unbiased estimates of the effect of the program at the expense of the lack of precision introduced by ignoring the information on actual program participation. While unbiased, the ITT estimator is obviously less precise then the estimator based on actual program participation, and provides a weighted average of the effect for those who attend with those who did not.

Our estimated program effects, whether based on program participation or the offer of the program, should be understood in the context of the other options avialable to families. With our data, we can determine if a student was exposed to targeted pre-K and if a child participated in targeted pre-K. A value of zero for PK, does not mean that the child received no early intervention. There are three possibilities that lead to $PK = 0$: 1) the child stays in the home and does not participate in any sort of early intervention; 2) the child participates in a private pre-K, which includes, for example, church-based care or informal care by neighbors; and 3) the child participates in another public option—such as, Head Start. Absent Texas's targeted pre-K, these are the counterfactual states for an eligible child, as these states represent what the child would have done had there been no targeted pre-K.

The introduction of targeted pre-K in Texas results in the crowding out of students from these alternative states. Conceptually, there is an implicit, unobserved treatment effect for going from no intervention to targeted pre-K, a treatment effect for going from private pre-K to targeted pre-K, and a treatment effect for going from Head Start to the targeted pre-K. As we don't observe these three states, the program effects estimated here are weighted averages of the three aforementioned effects where the weights are the percentages of the children that would be in each of the three unobservable states absent the newly available public option. This means that we can potentially find any result depending on whether targeted pre-K is of higher or lower quality than the other options. Still, the program effects estimate here are policy relevant parameters as they gives you the effectiveness of introducing another option given the existing alternatives available to parents. Careful consideration of "crowd out" offers a more nuanced understanding of the sources of variation that produce the parameters that we estimate.

Results and Discussion

Table 1 presents evidence of the variation that we exploit to identify the effects of targeted pre-k on academic outcomes. During the time period that we consider, the number of districts in Texas that offered targeted pre-K grew from 688 districts to 784 districts and the number of campuses—i.e. school buildings that housed a pre-K program—grew from 1,944 to 2,287. When a district offers pre-K in any school, students from the whole district are eligible to attend. To the extent that the enrollment decision is based on whether the program is available in the family's school district, then enrollment is exogenous to the circumstances of individual children. When the program is available, selection bias may occur. However, a non-trivial proportion of the variance in program participation is due simply to whether the program was offered in a given district in a given year.

Table 2 presents the regression results for the English language version of the 3rd grade TAAS Reading and Math tests. The key variable is PK which indicates that the student attended the public pre-K program. The reference case consists of students who did not attend the program, which includes those who stayed at home with relatives, informal care arrangements, Head Start, and private child care programs.

For the 3rd grade TAAS reading test, the OLS model reveals a statistically significant effect of 0.0552 for public pre-K attendance for those with economic disadvantage only. In other words, economically disadvantaged students who participated in public pre-K scored about 0.06 standard deviations higher on their third grade reading test than students who did not attend the program. For students whose reason for eligibility was limited English proficiency only, the effect is 0.0874 (obtained by adding the base level effect and the coefficient of the appropriate interaction term); the difference in the effect sizes for the two groups, 0.0295, is statistically significant. The largest effect size was experienced by students eligible for the program due to both economic disadvantage and limited English proficiency, 0.1107; again, the difference in the size of the effect compared to students with economic disadvantage only was significant.

A rule of thumb in education research is that one tenth of a standard deviation is considered a large effect. Thus, these effect sizes are substantively meaningful, particularly for an intervention that occurred four years prior to the outcome measure. The fact that the program's effect was largest for the students with two forms of disadvantage is also an encouraging result.

While these effects are smaller than those reported for model programs and resource-intensive programs, they indicate that even a modest program can help to boost student achievement.

Other covariates, included in the models but not shown in Table 2, serve as controls. These include indicator variables for race and gender, whether the student changed districts at any time, whether their kindergarten was full day, whether the student's district was rural or suburban, and a set of dummies identifying the students cohort year. The results are generally in line with expectations. The full model results are not shown, but are available from the corresponding author upon request.

Districts vary enormously in terms of their resources, institutional arrangements, demographics, and neighborhood characteristics. Many of these district level variables could affect the achievement of third graders and could also be related to whether or not the district offers a pre-K program and whether a given family chooses to use a program given that one is offered. Thus, we augment our basic model with a model that includes district fixed effects. This model implicitly controls for factors common to all the students within a district. The inclusion of district fixed effects slightly attenuates the estimated impact of public pre-K, but does not materially affect the results. The estimated effects for reading, shown in the second column of Table 2, are 0.0417 for economically disadvantaged students; 0.0657 for limited English proficiency students, and 0.0871 for students with both eligibility conditions. While the program impact is significantly greater than zero for all students, the difference in the size of the effect between the economically disadvantaged and limited English proficiency students is not significant in the fixed effect model, although the point estimate is similar in size to that estimated in the OLS model.

The story is quite similar for 3rd grade math test scores. The effect for students with economic disadvantage only is 0.0523 in the OLS model, and larger for LEP and students who are both economically disadvantaged and LEP. The main effect is slightly smaller when district fixed effects are added, 0.0394, but remains statistically significant. The differences in the effect sizes by eligibility class are significant in the OLS model and for students with both forms of disadvantage in the district fixed effects model.

The results discussed above are for tests conducted in English, even for those students classified as LEP. The TAAS tests are also administered in Spanish for students who English is so limited they would not be able to take the test in English at all. The sample size is smaller, at about 54,000, compared to the 493,000 who took the tests in English. Nevertheless, public pre-K was found to be effective for this group as well, with an effect of 0.0503 for reading and 0.0882 for mathematics in the OLS models. When district fixed effects are added, the reading effect drops to 0.0413 and is not significant at conventional levels (p=0.093); the reading score drops to 0.0620 but remains significant. In this group, no statistically signifcant difference was found between those who were LEP only compared to both LEP and economically disadvantaged.

The effects of the program were not limited to higher scores on standardized tests. We also estimate models for the probability of retention and special education designation. For grade retention, we analyze the probability that a student is retained in grades 1, 2, or 3 as a function of public pre-K controlling for covariates. Repeating of kindergarten is not considered retention, because kindergarten is voluntary and the decision to hold a student back in kindergarten is usually made by the parent, not the school.

Logit regression results reported in Table 4 indicates that attendance in public pre-K, relative to the alternatives, significantly reduces the probability of retention. The logit coefficient of -0.279 indicates that odds of retention are 24 percent lower for those who attended public pre-K. The odds of retention for students who qualify for the program due to limited English proficiency are 40 percent lower for those who did not attended public pre-K than for those who do not. The difference in retention among those who qualified due to both LEP and economic disadvantage were between those two. All the program effects were significantly different from zero, and the difference in the program effects by eligibility classes were also significant. These are large, substantively meaningful effects with important educational consequences for the students and for the costs of education in the state.

As in the case of the results for standardized tests, it is desirable to control for factors that are constant within district that could be correlated with pre-K availability or attendance. However, in the case of a logit regression estimated with maximum likelihood, more than one thousand fixed effects makes the analysis intractable. One issue is the sheer size of the maximization problem in estimating the logit, with close to one half million students. A second issue concerns that fact that within many districts, especially smaller ones and ones that did not offer public pre-K, there is no or very little within-district variation in the dependent variable, which is coded either one or zero indicating grade retention. To get around this, we estimate a linear probability model with OLS that is comparable to the logit model. The coefficient on public pre-k is -0.032, indicating the probability of retention is 3.2 percentage points lower for an economically disadvantaged student who attended public pre-k relative to a similar student who did not. (The logit equation produces a similar marginal effect when the probability of retention is 13 percent.) We then added district fixed effects to the OLS linear probability model. The resulting coefficient are nearly identical to the logit equation, confirming that these results are robust with respect to control of district-level factors.

Special education designation is a controversial dependent variable. On the one hand, pre-K might serve to provide earlier and better evaluation of students, leading to a higher level of appropriate placements. On the other hand, in some cases students who are borderline may be designated as special education if they perform very poorly or behave disruptively; if pre-K improves performance, emotional maturity, or social skills, it could reduce special education assignment. The results in Table 4 show that students who attended the Texas pre-K program were less likely to be assigned to special education in third grade; the odds of assignment were 13 percent lower for those who attended public pre-K other things equal. This result is confirmed in the comparable OLS model and the OLS model with district fixed effects.

So far the results indicate substantial positive benefits for students who participated in Texas public pre-K program. The greatest threat to the validity of these results resides in the selection of students into the program. Given that the selection into the program includes students choosing no child care and those choosing private child care, and given that care in the home by a relative can be a good option depending on the home situation, there is no way to tell a priori how this selection would bias the results, if at all.

The ITT results based on Equation 2 address this concern by removing all effects of selection, at the expense of losing information about actual participation in the program. Table 5 presents the results for the English language versions for third grade TAAS math and reading tests. For the reading test, the offer of pre-K is positve in both the OLS and fixed effects model, although it is only significant in the later. The effect size of 0.0509 for economically disadvantage students in the fixed effects model is similar to the estimate of the effects on those who actually participated. No statistically significant differences in the effect of pre-K are observed depending on reason for eligibility.

In mathematics, the effect for economically disadvantaged students is postive and signifcant in the OLS model but not in the fixed effect model.

The effect for students eligible due to limited English proficiency or both economic disadvantage and limited English proficiency are even larger than those eligible due to economic disadvantage only, and the differences are significant in the OLS model but no the fixed effects model. For those taking tests in Spanish, there is a large and statistically signifincant effect for mathematics, but not reading, and not when fixed effects for districts are included as shown in Table 6. In summary, despite the huge loss of information due to discarding knowledge about which students actually took the test, the ITT results are broadly consistent with the estimates based on Equation 1 in terms of the direction of the effect on student achievement, although the level of significance of the coefficients is lower as would be expected when information about program participation is discarded.

Table 7 presents the ITT results for grade retention and assignment to special education. Again, these results are broadly consistent with the estimates from the estimates based on actual program participation. The results from the ITT models do not support the notion that self-selection into program participation, for those student living in districts that offered the program, produced any significant bias in one direction or the other in the models based on actual program participation.

Conclusion

Evaluation of experiments is considered by many to be the gold standard in education research. However, experimental studies have limitations as well. For example, the experimental evaluation of the Tennessee Star program showed important effects of classroom size on student achievement. To implement this proposal at a large scale, however, requires hiring many new, inexperienced teachers. The new teachers are those who were at the margin and would not have been hired before the change. On average, they may be less skillful than the teachers already in the system. Moreover, the large expenditure on new teacher salaries may displace expenditures on other resources and alternative policy intiatives. Due to these macro effects, the experimental results on class size reductions may not be acheived in a large large scale implementation. To understand the effect of an educational intervention as actually implemented, it is important to conduct evaluations based using administrative data on programs in the field.

This paper has shown that targeted pre-kindergarten programs, even a mediocre program implemented state-wide, can have a positive impact on a number of academic outcomes even if they lack the resources or intensiveness of the model programs that have featured so prominently in the literature on pre-K. We found consistent effects on math and reading test scores of economically disadvantage and LEP students ranging from 0.05 to 0.1 standard deviations, depnding on reason for eligibility. Similar effects were found for students whose English was so poor they were tested in Spanish, a group of particular concern to policymakers. We also found reductions in the probability of retention in grade and assignment to special education. The results are robust to the inclusion of district fixed effects, and the ITT estimates suggest that the results are not driven by selection bias.

Given the importance of early intervention and the difficult fiscal environment that many states are experiencing since the 2008 recession, it is encouraging that Texas's Targeted Pre-Kindergarten program demonstrates such promise. Even modest programs can achieve important gains for economically disadvantaged and limited English proficiency students. States should strive for excellent, resource-intensive programs, but programs that fall short of this goal are still worthwhile for many students.

References

Anderson, M. (2008, December). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association 103*(484), 1481–1495.

Barnett, W., M. Carolan, J. Fitzgerald, and J. Squires (2011). The state of preschool 2011. *National Institute for Early Education Research, New Brunswick, NJ. Retrieved from http://nieer. org/yearbook*.

Bloom, H. S. (1984). Accounting for no-shows in experimental evaluation designs. *Evaluation Review 8*, 225–246.

Currie, J. and D. Thomas (1995, June). Does head start make a difference? *American Economic Review 84*(3), 341–364.

Currie, J. and D. Thomas (1999). Does head start help hispanic children? *Journal of Public Economics 74*(2), 235 – 262.

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from head start. *American Economic Journal: Applied Economics 1*(3), 111–34.

Fitzpatrick, M. D. (2008). Starting school at four: The effect of universal pre-kindergarten on children's academic achievement. *The B.E. Journal of Economic Analysis & Polic 8*(1 (Advances)).

Garces, E., D. Thomas, and J. Currie (2002). Longer-term effects of head start. *American Economic Review 92*(4), 999–1012.

Gormley, William T., J. and T. Gaye (2005). Promoting school readiness in oklahoma: An evaluation of tulsa's pre-k program. *Journal of Human Resources 40*(3), 533–558.

Heckman, J. J. and D. V. Masterov (2007). The productivity argument for investing in young children. *Review of Agricultural Economics 29*(3), 446–493.

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Yavitz (2010). The rate of return to the highscope perry preschool program. *Journal of Public Economics 94*(1-2), 114 – 128.

Knudsen, E. I., J. J. Heckman, J. L. Cameron, and J. P. Shonkoff (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Science 103*(27), 10155–10162.

Puma, M., S. Bell, R. Cook, C. Heid, G. Shapiro, P. Broene, F. Jenkins, P. Fletcher, L. Quinn, J. Friedman, J. Ciarico, M. Rohacek, G. Adams, and E. Spier (2010, January). Head Start Impact Study. Final Report. Technical report, U.S. Department of Health and Human Services, Administration for Children and Families, Washington, D.C.

Table 1: Changes in Pre-Kindergarten Offering Over Time

	DISTRICT DATA				CAMPUS DATA			
Year	TTL # DIST	TTL # OFFRNG PK	% OFFRNG PK	% CHG PY	TTL # CAMPUS	TTL # OFFRNG PK	% OFFRNG PK	% CHG PY
1990	1,057	549	52 %	–	5,978	1,537	26%	–
1991	1,053	567	54 %	3.28%	6,062	1,583	26%	2.99 %
1992	1,050	613	58 %	8.11 %	6,417	1,728	27%	9.16%
1993	1,048	677	65 %	10.44 %	6,283	1,875	30 %	8.51%
1994	1,046	688	66 %	1.62 %	6,369	1,944	31 %	3.68%
1995	1,045	723	69 %	5.09 %	6,500	2,051	32 %	5.50%
1996	1,044	741	71 %	2.49 %	6,819	2,133	31 %	4.00%
1997	1,059	761	72 %	2.70 %	7,035	2,210	31 %	3.61%
1998	1,061	784	74 %	3.02 %	7,222	2,287	32 %	3.48%
1999	1,103	816	74 %	4.08 %	7,394	2,341	32 %	2.36%
2000	1,183	851	72 %	4.29 %	7,549	2,414	32 %	3.12%
2001	1,199	884	74 %	3.88 %	7,598	2,505	33 %	3.77%
2002	1,234	925	75 %	4.64 %	7,672	2,610	34 %	4.19%

Table 2: TAAS Reading and Math: English Version

	Reading		Mathematics	
	OLS	FE	OLS	FE
PK	0.0552***	0.0417***	0.0523***	0.0394***
	(0.00320)	(0.00612)	(0.00317)	(0.00549)
$PK \times L$	0.0295*	0.0240	0.0418**	0.0259
	(0.0135)	(0.0184)	(0.0134)	(0.0202)
$PK \times B$	0.0555***	0.0454***	0.0536***	0.0383***
	(0.00714)	(0.00995)	(0.00706)	(0.00861)
L	0.0253*	0.00947	0.0931***	0.0722**
	(0.0115)	(0.0254)	(0.0114)	(0.0240)
B	-0.146***	-0.150***	-0.0195**	-0.0364*
	(0.00601)	(0.0176)	(0.00594)	(0.0168)
R^2	0.039	0.029	0.044	0.032
N	493028	493028	503761	503761

Notes: Robust Standard errors are in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: TAAS Reading and Math: Spanish Version

	Reading		Mathematics	
	OLS	District FE	OLS	District FE
PK	0.0503*	0.0413	0.0882***	0.0620*
	(0.0246)	(0.0248)	(0.0249)	(0.0291)
$PK \times B$	-0.0187	-0.0198	-0.0256	-0.0112
	(0.0262)	(0.0287)	(0.0265)	(0.0320)
B	-0.0482*	-0.00644	-0.0243	0.00449
	(0.0206)	(0.0281)	(0.0209)	(0.0328)
R^2	0.038	0.038	0.025	0.027
N	54134	54134	53554	53554

Notes: Robust standard errors are in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Retention and Special Education Designation

	Retention			Special Education		
	Logit	OLS	District FE	Logit	OLS	District FE
PK	-.279***	-.032***	-.036***	-.144***	-.02***	-.022***
	(.009)	(.001)	(.002)	(.008)	(.001)	(.002)
$PK \times L$	-.228***	-.014***	-.013***	.052	.013***	.013**
	(.035)	(.004)	(.004)	(.038)	(.004)	(.004)
$PK \times B$	-.067***	-.005**	-.005	.014	.010***	.008***
	(.017)	(.002)	(.003)	(.018)	(.002)	(.003)

Notes: Where appropriate, robust standard errors are in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: ITT—-TAAS Reading and Math: English Version

	Reading		Mathematics	
	OLS	FE	OLS	FE
PO	0.0164	0.0509**	0.0192**	-0.0066
	(0.00912)	(0.0248)	(0.009)	(0.0244)
$PO \times L$	-0.0512	0.0295	0.0418**	0.0071
	(0.0615)	(0.0707)	(0.0605)	(0.0549)
$PO \times B$	0.0340	-0.0006	0.0988**	0.0462
	(0.00317)	(0.0390)	(0.0313)	(0.0410)
L	0.0891	0.0790	0.0962	0.0852
	(0.0612)	(0.0698)	(0.0603)	(0.0544)
B	-0.138***	-0.117***	-0.0775**	-0.0539
	(0.0316)	(0.038)	(0.0313)	(0.0397)
R^2	0.037	0.036	0.043	0.041
N	493028	493028	503761	503761

Notes: Robust Standard errors are in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: ITT—TAAS Reading and Math: Spanish Version

	Reading		Mathematics	
	OLS	District FE	OLS	District FE
PO	-0.0194	-0.1240	0.4276***	-0.0229
	(0.1195)	(0.1392)	(0.1232)	(0.0521)
$PO \times B$	0.0301	0.0973	-0.1632	-0.0108
	(0..1371)	(0.0814)	(0.1408)	(0.0438)
B	-0.0902	-0.1155	0.1251	0.0069
	(0.1365)	(0.0789)	(0.1402)	(0.0399)
R^2	0.038	0.033	0.001	0.022
N	54134	54134	53554	53554

Notes: Robust standard errors are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: ITT—Retention and Special Education Designation

	Retention			Special Education		
	Logit	OLS	District FE	Logit	OLS	District FE
PO	-0.088***	-0.010***	-0.001	-0.137***	-0.027***	-0.005
	(0.025)	(0.003)	(0.010)	(0.021)	(0.003)	(0.009)
$PO \times L$	-.0395**	-0.041*	-0.035	0.062	0.025	0.036*
	(0.141)	(0.017)	(0.027)	(0.159)	(0.017)	(0.018)
$PO \times B$	-0.317***	-0.040**	-0.041*	0.208**	-0.005	0.003
	(0.070)	(0.009)	(0.016)	(0.073)	(0.009)	(0.013)

Notes: Where appropriate, robust standard errors are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$