Public Investments in Early Childhood Education and Academic Performance: Evidence from Head Start in Texas

Esra Kose*

May 15, 2019

Abstract

Do early childhood investments for low-income children narrow the academic achievement gap in elementary school? I study this question in the context of Head Start, by utilizing previously unexplored variation in the federal funding expansions across counties in the 1990s. Using student-level data from Texas, I find that exposure to more generous Head Start funding during childhood led to substantial gains in test scores, particularly for low-income Hispanic students. Hispanics benefited from early childhood investments through increased access to the Head Start program and improvements in program inputs during early childhood. These advances, in turn, led to enhancement in their language proficiency and reduction in their likelihood of special education needs during elementary school.

JEL classification codes: H53, I32, I38, J13

Keywords: Early Childhood Education, Head Start, Program Evaluation

---

*I am thankful to Doug Miller, Marianne Page, Marianne Bitler, and Scott Carrell for their help and support. I also thank Nicholas J. Sanders, Elira Kuka, Na'ama Shenhav, Hilary Hoynes, Chris Walters, Aline Buetikofer, Gaetano Basso, Cassie Hart, Vasco Yasenov, Michel Grosz, Ariel Pihl, and brownbag participants at UC Davis for providing input. I am grateful to Matthew Neidell and David Frisvold for generously providing data. Administrative student-level data are from the Texas Education Agency. These data are currently administered by the University of Texas Education Research Centers in Austin and Dallas and are available for use to approved researchers. An earlier version of this paper circulated under the title "Local Area Spending Exposure to Head Start and Academic Performance: Evidence from Texas." All errors are my own. Esra Kose, Department of Economics, Bucknell University. E-mail: esra.kose@bucknell.edu.

1 Introduction

Racial, ethnic, and socioeconomic disparities in academic achievement have been ongoing issues of the U.S. education system. These disparities are of concern to policymakers, since black and Hispanic students perform much worse than the average white student on standardized tests (Reardon et al., 2008). Similarly, students from low income backgrounds score below the average high income student on such tests. These gaps persist over time, raise serious concerns about the life prospects of these children and the state of social mobility in the U.S. (Duncan and Murnane, 2011).

Do early childhood investments play a role in narrowing the academic achievement gap? Early childhood is a period of physical, cognitive, socioemotional, and language development; a child's environment can dramatically influence the degree and pace of these advances. There is strong theoretical support in the fields of economics, neuroscience, and child development that points to early childhood as a particularly important time to invest. Despite the strong theoretical support, regardless of the scale of the program many early childhood interventions that target skill development demonstrate initially promising results that disappear by the time students reach third grade (Bailey et al., 2017). Given these findings, whether early childhood investments close the achievement gap in elementary school is still an open question.

I study this question in the context of the largest federal early childhood program in the U.S., Head Start (HS). HS began in 1965 to provide education, health, and social services to low-income children ages three to five, as well as their families. The goal of the program is to reduce education and health disparities between low-income children and their more advantaged counterparts. National studies have indicated that HS has not fulfilled its mission of closing the achievement gap across socio-economic groups when children are in elementary school (Barnett, 2011). The best available evidence is based on a random assignment of 4,442 children to a national sample of HS centers in the early 2000s. It reports notable impacts on test scores at the end of the HS year that fade out by third grade (Puma et al., 2012). Studying the period of the 1990s when federal HS funding quadrupled, I show that HS reduced the achievement gap significantly in third grade between Hispanic and white students in Texas.

The 1990s were a golden age for the HS program, as they were a period of immense expansion, with bi-partisan efforts in Congress to pass two important acts aimed at improving the quality and capacity of the program. Specifically, this paper investigates to what extent federal HS funding expansions during this time period affect student performance in Texas. How does the effect of public investments relate to the way funds are spent? The investigation of the first question provides new insight into the old debate. The second question presents new evidence of whether public investments were spent as they were planned, by linking administrative data on HS program characteristics and budgets to funding expansions. The evidence concerning this question has been limited, primarily due to data availability and data quality (Currie and Neidell, 2007). Furthermore, the existing research on early childhood education quality is still seeking to understand which measures of program quality matter most for child outcomes (Blau and Currie, 2006; Walters, 2015).

My empirical strategy utilizes a panel-fixed effects model, which leverages variation across counties in the timing and intensity of HS funding expansions in the 1990s. A similar methodology has been used in a few recent papers; these papers take advantange of the roll-out of HS grants at the county level in the first few years of the program's introduction, in order to identify the long-run socio-economic impacts of HS (Thompson, 2018; Bailey et al., 2018; Barr and Gibbs, 2018; Johnson and Jackson, 2019). The main identification assumption is that funding expansions are exogenous to other underlying geographic-level trends in test scores. I employ various analyses to check the robustness of the estimates, and show that the results are not driven by the expansion of the alternative early childhood programs during this time period, selection bias, or other endogenous factors.

My analyses use a large, demographically and socioeconomically diverse student population in Texas. Student-level administrative data on student characteristics and achievement for third graders are provided by the Texas Education Agency (TEA) from 1994 to 1999. To construct a unique county-by-year dataset on HS spending per age-eligible child from 1988 to 1994, I match grantee-level HS spending data from the Consolidated Federal Funds Reports (CFFR) with administrative data that describe the serving counties for each grantee. Combining student-level data with information on HS funding generosity, I estimate the effects of exposure to HS funding at age four on third grade standardized test scores in math and reading (at ages nine and ten) for children born between 1984 and 1990.

The main finding is that exposure to HS funding expansions significantly improved academic performance in third grade. For students who were free-or-reduced-lunch certified, being exposed to an average-sized HS program led to a 0.04 standard deviation increase in average test scores in math and reading combined.[1] Using several sources on program characteristics and budgets, I show that additional HS funding led to significant increases in HS participation rates. Moreover, federal funding expansions are associated with increases in child-teacher ratios, child-staff ratios, full-time enrollment, and education spending in the HS programs. In totality, these analyses suggest that both program capacity and quality improvements are important pathways for the ultimate effect on test scores.

Estimates by race and ethnicity show that improvements among Hispanics are the main driver of these results. Overall, my findings suggest that, among low-income students, the funding increase during the 1990s led to a 15% reduction of the test score gap in math and reading combined between Hispanics and whites. There are at least three testable channels through which HS could be beneficial for Hispanics. First, HS could improve their language skills by exposing them to English at an early age, which could in turn affect their academic performance. My analysis shows that HS funding exposure significantly improved language proficiency for Hispanics. Second, HS could help special needs children during childhood; HS programs have been required to serve 10% of their enrollment for children with disabilities, according to the federal law. I show that additional funding led to a reduction in the

[1] This intent-to-treat estimate of 0.04 standard deviation corresponds to treatment effect of the treated of a 0.4 standard deviation increase in test scores as an upper bound.

likelihood of having special education status for Hispanics. Finally, Hispanics have higher participation in HS in Texas compared to whites and blacks combined. I find that additional funding induced more Hispanic children to participate in the program.

This paper contributes to economics literature in a number of ways. Firstly, my analyses provide new evidence on the effects of HS on academic achievement by taking advantage of a different policy lever (spending, not participation), a different time period, and a different population. Previous studies have exploited within-family comparisons of siblings who have and have not participated in the program (Currie and Thomas, 1995, 1999; Garces et al., 2002; Deming, 2009; Bauer and Schanzenbach, 2016). This approach potentially suffers from spillover effects across siblings (Black et al., 2017) and measurement error in the retrospective report of participation in HS, which may bias estimates toward zero.[2] Other papers that exploit discontinuities due to program funding and eligibility rules significantly improve upon earlier studies, but still suffer from limited sample sizes (Ludwig and Miller, 2007; Carneiro and Ginja, 2014). Relative to existing studies, which are underpowered to detect reasonable effect sizes in test scores, I find significant improvements using the whole student population of Texas.

Secondly, my findings align well with a growing number of papers which find that public investments and education policies in early childhood, or in the elementary school, are more effective for Hispanics and children with low levels of baseline scores (e.g., Currie and Thomas, 1999; Bitler et al., 2014; Gibbs, 2017; Figlio and Ozek, 2019). My study explores the use of public funds in early childhood for low-income children during the 1990s, which has so far not been fully studied.

Finally, I contribute to the growing literature which examines the effect of public spending on academic success. The most recent causal evidence shows that public spending during the formal school year improves student outcomes (Jackson, 2018). What is less known is whether public spending during early childhood makes a difference in academic performance.

[2]Moreover, a new paper by Miller et al. (2018) points out the challenge with the interpretation of family fixed effect estimates, as they rely on the assumption of the switchers in the sample.

My paper provides new evidence on this question in the context of HS.

The rest of the paper is organized as follows. I present background on the HS expansions and review the prior research in Section 2. Section 3 describes data sources, followed by an overview of the methodology in Section 4. I report my results in Section 5 and discuss potential mechanisms in Section 6. I then present robustness checks in Section 7, a cost-benefit analysis in Section 8, and conclude in Section 9.

2 Background and Prior Literature

2.1 The Head Start Program

Head Start (HS) is a federally-funded early childhood education program that provides education, health, nutrition, and other services to economically disadvantaged children and their families. The program is designed to reduce the disparities in school readiness, health, and other social aspects between low-income children and their more advantaged peers. The main eligibility criteria are that children must be between ages three and five, and come from families with income at or below the poverty level. Children automatically qualify if they are homeless, in foster care, or their family receives Supplemental Security Income (SSI) or Temporary Assistance for Needy Families (TANF) (U.S. Department of Health & Human Services (HHS)).[3] In addition, at least 10% of the children served in each center must have some type of disability, regardless of the income eligibility.

HS began as a part of the "War on Poverty" initiative in 1965 as a summer-only program. Soon after its implementation, HS became a nine-month program that operated part-day. In the 1990s, the program expanded substantially and shifted to more of a full-day program. Appendix Figure A.1 shows that, between 1990 and 2000, enrollment increased by about 60%, and the federal funding per enrolled child doubled (HHS, 2017). The 1990s expansion was a

[3] However, eligibility does not imply attendance. In 2015-16 fiscal year, 31% of eligible children ages three to five had access to HS (National Head Start Association, 2017).

result of a significant effort by the Bush and Clinton Administrations to improve quality and capacity constraints. In particular, additional funding was appropriated toward increases in teacher salaries and training, expansion of services for families of children attending the program, and the local HS agencies to purchase facilities. As a result of these policy efforts, there was a substantial ramp-up in federal funding per child, which provides a natural experiment with which to study the program.

HS is a federal-local matching grant program. The federal government determines overall HS funding annually, as a component of the federal budget, and allocates it to states on the basis of the relative number of public assistance recipients, unemployed persons, and children from families below the poverty line (Community Services Act of 1974). To receive funding, local agencies must write grant proposals directly to the Head Start Bureau in the Administration for Children and Families (ACF) of the Department of Health and Human Services (HHS). Grants are issued through a competitive process, with priority given to agencies that are able to demonstrate the most cost-effective operation. Existing programs have priority when re-applying. Grantees must provide at least 20% of the funding, which may include in-kind contributions through community partnerships such as facilities in which to hold classes.[4]

Most HS grantees operate through community action agencies, local school systems, private/public non-profits, government agencies, and Indian Tribes. In Texas, for example, during 1988-1994, the distribution of the programs was approximately 38% community action agencies, 33% local school systems, 23% private/public non-profits, 5% government agencies, and less than 1% Indian Tribes (Program Information Reports). Therefore, grantees are heterogeneous in several dimensions, such as costs of personnel and space (depending on the geographic location, for example), and type of sponsoring agency (school system or private nonprofit) (Currie and Neidell, 2007).[5] As a result, there is a great deal of geographic

[4]Additional funds for cost of living adjustments, quality improvements, and other initiatives could be made available by the federal government depending on the needs of the communities (HHS, 2000).

[5]However, each center must comply with publicly known standards, which are described in the Head Start Act.

variation in funding levels, which provides part of the identification in this paper.

As stated by Currie and Neidell (2007), a local grantee can obtain additional funding in three main ways: (1) the federal government allocates more funding in a given year; (2) program directors write better grant proposals and attract more funds; or (3) grantees attract larger local funds from the state or other local community agencies based on need or better connections. Because part of the variation in funding stems from the qualifications of the grantees (and is not due to exogenous policy changes), careful consideration is needed to isolate exogenous variation to identify the effectiveness of HS. These issues will be addressed in detail in Section 4.

2.2 Prior Literature

As the largest federal early childhood program, Head Start (HS) has been evaluated extensively over its existence.[6] Nevertheless, the evidence is limited to cohorts who participated the program in the first 15 years, and to cohorts who partipated the program in the 2000s. My paper contributes to the literature by providing new evidence from the 1990s. During this time, substantial policy efforts to provide services full-time, while expanding enrollment and improving program quality, transformed the HS program. This was also a time period when state-funded preschool alternatives were not fully available to low-income families.

The best available evidence on the short-term impacts of HS on children comes from the Head Start Impact Study (HSIS), which took place in 2002 among oversubscribed HS centers. Puma et al. (2012) show that short-term positive gains for most measures of cognition disappear by age nine (third grade). It is important to note that around 40% of the control group participated in some center-based care, and failing to account for this would bias the estimates towards zero. Therefore, simple treatment and control group comparisons as reported in Puma et al. (2012) are not too informative about the effects of the program

[6] A number of studies have reviewed the literature on the HS's effectiveness (Barnett, 1995; Currie, 2001; Barnett and Hustedt, 2005; Ludwig and Phillips, 2008; Shager et al., 2013; Duncan and Magnuson, 2013; Gibbs et al., 2013).

relative to no early childhood alternative. On the contrary, the small short-term test score gains and quick fade-out in cognitive outcomes have been highly influential as evidence against the success of the HS program.

More recent papers, which address the alternative counterfactuals in the HSIS in a more systematic way, find positive significant effects of HS (relative to home-care) for the short-term test score impacts up to first grade (Feller et al., 2016; Zhai et al., 2014; Kline and Walters, 2016). Additionally, Bitler et al. (2014) examine the distributional effects of HS using the HSIS. They find larger short-term impacts at low quantiles of the test score distribution, and persistent effects through the first grade for Spanish-speakers at the bottom of the test score distribution. However, these recent papers do not provide evidence on test-score effects in third grade. Overall, in terms of time period, my paper is likely to be comparable to the evidence in the HSIS.

Since the longest follow-up for the HSIS is the effects in third grade, the best available evidence on the longer-term effects comes from quasi-experimental studies. Previous literature has exploited within-family comparisons of siblings who have and have not participated in the program (e.g., Currie and Thomas, 1995, 1999; Garces et al., 2002; Deming, 2009; Bauer and Schanzenbach, 2016).[7] This approach suffers from spillover effects across siblings (Black et al., 2017) and measurement error in the retrospective report in participation to HS, which may bias estimates toward zero. Other papers that exploit discontinuities due to program funding and eligibility rules significantly improve upon earlier studies, but still suffer from limited sample sizes (Ludwig and Miller, 2007; Carneiro and Ginja, 2014).

A few recent papers exploit variation in HS funding expansions to analyze the long-term effects of HS on socio-economic well-being and health for participants in the early introduction of the program (Thompson, 2018; Bailey et al., 2018; Barr and Gibbs, 2018; Johnson

[7]In terms of the outcome measures, my paper is comparable to Deming (2009), which studies the impact of HS on cognitive test scores for elementary school children (in addition to the long-term impacts). Deming (2009) examines the participants before the 1990s, and finds that short-term significant test score gains which persist through elementary school, then fade out. He analyzes the effects on the cognitive tests separately for different age groups: age 5-6 (kindergarten), age 7-10 (elementary school), and age 11-14 (adolescent).

and Jackson, 2019). For example, Bailey et al. (2018) and Barr and Gibbs (2018) use the roll-out of HS grants in the first few years of the program's introduction to examine the long-run effects on the first generation and the second generation, respectively. Thompson (2018) and Johnson and Jackson (2019) examine the long-term socio-economic well-being of early participants by exploiting HS funding expansions across counties over time, and find positive effects of the program. My paper uses similar methodology, stemming from the 1990s program expansions to analyze the impact of HS on academic performance. Additionally, I utilize student-level administrative data in my study, which is advantageous to overcome statistical inference issues.

3 Data Construction and Summary Statistics

3.1 Data Construction

I combine several data sets on HS spending, student test scores and demographics, economic conditions, and school quality to analyze the effect of HS funding expansions on academic performance.

Administrative student-level data include the universe of exam takers in Texas in third grade from 1994 to 1999. These data are from the Texas Education Agency (TEA) and include test scores monitored through the Texas Academic Assessment System (TAAS).[8] Texas became the first state to use achievement tests to measure school performance in 1993, with the goal of ensuring that student achievement at each school meets specific, minimum standards (Richardson, 2010; Carnoy and Loeb, 2002). With the standardized testing requirement, and through robust data collection, the state of Texas provides high-quality student-level data on academic performance. Relevant for this paper, these data contain information on the year of birth, gender, ethnicity, free-or-reduced lunch status, language

[8]The TAAS was administered from 1994 to 2002; however, the TEA started offering a Spanish version of the standardized tests for students with limited language proficiency in 2000. To track students' performance consistently, I only keep test years from 1994 to 1999.

proficiency, and special education status for each student at each school district.[9] From the school district information, I determine the county of residence, which serves as a proxy for the child's county of residence at age four. Along with the year of birth, this determines each student's exposure to HS funding.

I conduct the following sample restrictions. Third grade students make up 1,000,524 observations between 1994 and 1999. First, I drop observations with missing demographic information, missing test scores, exempt testing status, or nonstandard test administration (739,427 observations remain). Next, using the description in the administrative data, I restrict my analysis to students who are certified for free-or-reduced lunch or who are identified as economically disadvantaged based on their families' welfare eligibility, because they are more likely to be eligible for HS (332,910 observations remain).[10] From here, I will refer to this sample as free-or-reduced-lunch certified.

I develop a measure of HS funding per age-eligible child in a local community using a number of sources. HS spending data are from the Consolidated Federal Funds Reports (CFFR), which include information on local appropriations for federally funded programs for each grantee from every year starting in 1983. Similar to other programs that are financed by federal grants, a HS grantee could oversee one county or a group of counties. A detailed description of data construction is in Appendix Section A.1. The years of HS spending are restricted to 1988 and 1994, when the analysis sample would be age-eligible to attend HS. The data structure section in Appendix A.1 shows the mapping of test score years with HS funding data for third grade cohorts. To construct the population denominator of children ages three and four years old,[11] county-level population counts for each age group are extracted from the Surveillance, Epidemiology, and End Results Program (SEER). Together

[9]See Appendix A.1 for more detailed information.

[10]Students from families reporting income between 130 and 185 percent of the federal poverty line are certified for reduced price meals, while children from families with incomes below 130 percent of poverty are eligible for a fully subsidized or free meal (U.S. Department of Agriculture, 2014). Also, they are automatically eligible if their families collect Food Stamps or TANF benefits. This is similar to eligibility for Head Start.

[11]Three- and four-year-old children make up to approximately 95% of the HS enrollment.

with these two data sources, I construct the "HS funding per child" variable at the local community-year level.

I augment these data with additional data sources to bring information on (1) HS enrollment and program budget items; (2) school quality and other pre-k alternatives in Texas; and (3) economic conditions. First, I compiled data from the Program Information Reports (PIR). Starting in 1988, the Office of HS Program has collected comprehensive data from all grantees and delegates on the services, staff, children and families served by the program. These data provide information on the number of funded enrollees, the number of staff, demographic composition of children and staff, and qualifications of directors. Second, I source school-level pupil teacher ratio and information on school-level state-funded pre-K (outside of HS) enrollment from the Common Core of Data (CCD).

Next, I compiled data on county-level economic conditions. I use the Regional Economic Information Systems (REIS) to construct county-year data on per capita income, per capita transfer payments for cash income support (AFDC and SSI), medical benefits (Medicare, Medicaid and Children's Health Insurance Program), food assistance (Food Stamps), retirement, and disability programs. Using the 1980 City and County Data Book (before 1990s HS spending expansion), I construct other county demographics. These include the 1980 population living in an urban area, black, Hispanic, single parent, less than age 5, ages 65 or older, percentage of the 0-18 year-olds living in poverty, as well as income, education, welfare spending per capita (in 2014$). To control for exposure to business cycles at birth, I use the county-year unemployment rate from the Bureau of Labor Statistics (BLS). Finally, to control for the composition of the demographics of the population, and population counts at the county-level by racial and age groups, I use data from SEER.

3.2 Summary Statistics

My analysis centers on children exposed to HS funding in Texas between 1988 and 1994. Figure 1 presents the average HS spending per child in the 15 most populous counties in

Texas from 1980 to 2000.[12] The vertical lines in Figure 1 highlight the period of this study. This figure shows that there is substantial variation in HS spending per child across counties and within a county over time. To get more insight into the geographic variation in HS funding, Figure 2 presents maps of: (i) the levels of funding in 1988 and (ii) growth in the HS spending per child from 1988 to 1994, respectively. These maps show that there is a great deal of variation across local communities in both levels and growth in funding. My basic identification strategy uses this geographic and time variation to identify the effect of HS on academic achievement.

Appendix Figure A.2 shows the distribution of student test scores in math and in reading by free-or-reduced lunch status. The solid line shows the passing score of 70 that is determined by the TEA. This figure shows that free-or-reduced-lunch certified students were academically behind compared to the rest of the sample, with a lower passing rate.[13]

Table 1 presents summary statistics for third grade students, first for the full sample and economically advantaged students, then for students who are certified for free-or-reduced lunch. The first three columns show a significant discrepancy in average test scores by economic status. Relative to free-or-reduced-lunch certified students, economically advantaged ones have substantially higher test scores. Importantly, Hispanics make up more than 50% of students who are certified for free-or-reduced lunch. They have the lowest test scores on average, relative to other economically disadvantaged students, and tend to live in counties with higher funding for HS per child. Figure 3 shows the evolution of standardized test scores across birth cohorts by their status of free-or-reduced lunch. The most striking observation is that even among the non-certified students, there is a significant achievement gap among racial and ethnic groups. For those who were certified for free-or-reduced lunch, academic performance of Hispanics and blacks seems to be improving across cohorts. Overall, these observations suggest that there is value in analyzing the potential effects of early childhood

[12] There are 254 counties in Texas. I chose the 15 biggest counties for the purpose of clear visualization of the variation. These counties in total make up around 60% of the student population in Texas.

[13] Probability of passing was 70% for non-certified students relative to 30% for certified ones.

investments separately by race and ethinicity.

To demonstrate the relationship between average test scores and HS funding, I restrict the sample to children certified for free-or-reduced lunch and collapse the data on test scores to county-level averages. Appendix Figure A.3 shows raw correlations during the period of this study, and implies that there is a positive relationship between HS spending per child and average test scores for free-or-reduced-lunch certified children. The purpose of this paper is to provide evidence on the effect of HS funding expansions on test scores by implementing a careful methodology, which is described in the next section.

4 Empirical Strategy

To study the effect of HS on academic performance, I exploit variation in community-level HS funding per child in the 1990s. Following Ludwig and Miller (2007) and Sanders (2012), I assign HS funding exposure based on each student's county of residence and year of birth.[14] My empirical strategy utilizes a panel-fixed-effects approach, which relies on variation within communities and over time in HS spending per child, conditional on observables. Formally, I estimate the following equation for the sample of free-or-reduced-lunch certified students using a newly assembled dataset on community-level HS spending per child:

$$Y_{iscbt} = \alpha + \beta HSfunding_{c(b+4)} + X_{iscbt}\gamma + Z_{ct}\lambda + W_{c(b+4)}\psi + \theta_s + \xi_b + \eta_t + \pi_c * b + \epsilon_{iscbt}, \quad (1)$$

where Y_{isgbt} denotes the outcome variable (standardized test scores) for student i in school s in local community c in birth year b and in test year t. $HSfunding_{c(b+4)}$ represents HS funding per child in local community c when student i was four years old ($= b+4$). X_{iscbt} is a vector of individual-level demographic controls including gender, ethnicity, and an in-

[14]If low-income families make migration decisions depending on the availability of the services provided by HS, the assignment using county of residence would create bias in my estimates. I test this directly in Appendix Table A.6 and find no evidence that the HS spending affects the composition of students within school over time.

dicator for bilingual, English as a second language, and gifted/talented. Z_{ct} has three sets of community-level controls, including per capita income transfers, average characteristics of students collapsed from the main data, and additional controls, which include fraction of 0-18 living in poverty, percent urban, share of 0-5 year-olds by ethnicity, log population, and unemployment rate. $W_{c(b+4)}$ includes community-level controls at the time of HS such as income per capita, share of pre-K enrollment, and income transfers per capita. Finally, θ_s, ξ_b, η_t are school, birth year, and test year fixed effects respectively, and $\pi_c * b$ is a community-specific linear trend. Standard errors are clustered at the county level.[15] The coefficient of interest is β, which is interpreted as the conditional change in the outcome variable from a unit increase in exposure to federal HS funding per child at age four.

For this research design to be valid, funding expansions must be exogenous to other underlying geographic-level trends in test scores. Threats to identification are any differential trends among communities that are correlated with spending changes, which may also influence student outcomes. I use several methods to probe the validity of the key identification assumption.

First, I take county-level characteristics measured in 1980, before the expansions occurred, and use them to predict the levels and changes in HS spending per child from 1988 to 1994 (similar to Hoynes and Schanzenbach (2012)). For this analysis, I collapse the data to the local community-level. The independent variables include percent of the 1980 population living in an urban area, black, Hispanic, single parent, less than age 5, age 65 or older, population, percent of the 0-18 year-olds living in poverty and income, education, and welfare spending per capita (in 2014\$). The results are presented in Appendix Tables A.1 and A.2.[16] Simple correlations imply that communities that have a larger Hispanic population and a higher share of single mothers, poor, very young, or elderly have more HS funding. In contrast, communities with a larger black population tend to have less funding for HS, which could be because blacks in Texas live in more urban areas with high population

[15]Clustering at the local community level does not affect the estimated standard errors.
[16]See Online Appendix Figure OA.1 for visual presentation of correlations.

density. Additionally, communities with more social spending have higher funding for HS, and communities with higher per capita income have less HS funding. The determinants of the change in HS funding from 1988 to 1994 are also similar (see Appendix Table A.2). After controlling for characteristics that are described above, the last columns of Appendix Tables A.1 and A.2 show that income per capita and percent children living under poverty are significant determinants of HS expansions, and together all variables explain around 25 percent of the overall variation (R-squared ≈ 0.25). Nevertheless, to control for possible differences in trends across communities, I include the observable determinants of the funding variation and community-specific time trends in all of my models. The results are robust to exclusion of these trends.

Second, director quality could be a possible confounding factor if directors who are able to obtain more funds may also run programs that are better in other respects. For example, a bias will occur if better quality directors write better grants to obtain additional funds, and operate higher quality programs. In this case, funding levels could be correlated with child outcomes (Frisvold, 2006; Currie and Neidell, 2007). To rule out this possibility, I show that federal funding increases are not significantly associated with directors' qualifications in Online Appendix Table OA.1.[17] A bias could also occur if some communities devote local resources toward children, and if these resources during childhood also help children succeed in school years. Inclusion of school fixed effects controls for this possibility, as well as other neighborhood characteristics that could be correlated with student success.

Next, HS may have been introduced or expanded with other local policies that affect children's outcomes, such as other War on Poverty programs. For instance, the timing of the introduction of HS corresponds to the foundation of the other government programs including Medicaid, Medicare, Food Stamps, and the Supplemental Nutrition Program for Women, Infants and Children (WIC). To address the concerns regarding contemporaneous

[17] One caveat to this analysis is that data on directors' characteristics are available from the PIRs starting in 1992. Therefore, the sample is restricted to individuals who were exposed to the HS program between 1992 and 1994.

policy changes that targeted four-year-olds, I directly control for county-level spending for other social programs.[18] Moreover, inclusion of local community-specific linear time trends also accounts for the fact that some communities may be improving over time.

Finally, important to my identification strategy is that HS spending expansions should not systematically change the composition of a particular cohort within a school (similar to Carrell and Hoekstra (2010)). For example, if low-income families made decisions to move based on the generosity of HS, and HS generous communities had better schools, such nonrandom selection would misattribute higher student performance to exposure to HS. In Appendix Table A.6, I formally test this and other types of self-selection by examining whether student characteristics such as gender, race/ethnicity, and county-level income per capita are correlated with HS spending per child after conditioning on school fixed effects. I find that variation in HS funding does not predict changes in the composition within school, suggesting that the estimates are not biased by self-selection of families to particular cohorts within a school.

4.1 State-Provided Early Childhood Education Alternatives in Texas

A potential threat to my identification strategy is the existence of state-provided preschool alternatives to HS, as all preschools provide similar services to improve children's school readiness. Texas has offered a half-day public pre-Kindergarten (pre-K) program since the 1985-1986 academic year. State-provided pre-K aims to improve the academic performance of at risk children by providing early childhood education for four-year-olds who are identified as at risk (TASB, 2010).[19] Although it is mandatory for any district that has at least 15 eligible children to offer a half-day education-based program for four-year-old children,

[18] See Section 3 for detailed discussion about which controls are being added.

[19] At risk population includes: children unable to speak and comprehend the English language; children certified for free-or-reduced lunch program; those that are homeless as defined by federal law; a child whose parents are either on active military duty, in an activated reserve unit, or who were killed or wounded while serving on active duty; and children in the Texas foster care system (Texas Education Code 29.1531).

attendance is voluntary.[20,21] Thus, starting in 1985, an eligible child in Texas could attend a public pre-K as an alternative to HS. While funding for pre-K is allocated directly to school districts by the state of Texas, districts are encouraged to partner licensed child care centers and HS programs to provide preschool services (NIEER, 2012).[22,23] This raises the possibility that HS and pre-K might have operated jointly to some degree during the 1990s.

Appendix Figure A.4 plots the number and the share of children enrolled in HS and pre-K in Texas between 1988 and 1994. Relative to HS which served a 6% share of age-eligible enrollment in 1988, pre-K was larger, with a 14% share. This figure also shows that the timing of the expansions in both programs coincide. This is a concern, as the existence of a large-scale pre-K is a potential threat to identification, as a confounding factor. Hence, it is important to control for the availability of other preschool alternatives, as failing to control for it could misattribute the effects of other preschools to HS.

To diagnose and address this concern, I directly analyze the relationship between HS spending and pre-K expansions between 1988 and 1994. It is true that pre-K expansions coincide with the timing of HS expansions at the aggregate state level. However, Appendix Table A.3 shows that, after controlling for community and year fixed effects, there is not a significant relationship between a share of age-eligible (Column (1)), and age- and income-eligible (Column (2)) children enrolled in pre-K and HS funding per child. While this analysis confirms that variation in HS per child does not predict pre-K enrollment, there is still a possibility that, in terms of facility usage and program operations, pre-K and HS cooperated in the 1990s.[24] To take this possibility into account, I control for the share of age-eligible

[20] Texas Education Code 29.1532.

[21] Andrews et al. (2012) evaluate this program and show that it has been effective on improving math and reading test scores, reducing the likelihood of being retained in grade, and decreasing the probability of receiving special education services.

[22] `http://nieer.org/sites/nieer/files/Texas_0.pdf`

[23] State funds eligible children through the Foundation School Program based on average daily attendance (TEA, 2014).

[24] Starting in 2003, the state law requires the new pre-K establishments to coordinate and cooperate with HS (TEC §29.1533; TEC §29.158). For more details, see `http://www.statutes.legis.state.tx.us/Docs/ED/htm/ED.29.htm#29.1533` and `http://www.statutes.legis.state.tx.us/Docs/ED/htm/ED.29.htm#29.158`.

children enrolled in pre-K in a given county during the time of exposure to HS and show that the results are not sensitive to addition of this control.

5 Results

Having documented the plausible exogeneity of the Head Start (HS) funding variation, I now present the results of Equation 1. As noted above, the main sample is restricted to free-or-reduced-lunch certified students and HS exposure is assigned at the time when a child was four years old.[25] To simplify the interpretation of the coefficient of interest, HS spending per child is scaled by the mean spending ($517), thus, the coefficients should be interpreted as the effect of exposure to an average-sized program (similar to Thompson (2018)). All monetary values are converted to 2014 dollars. Test scores are standardized using the entire student population.

5.1 Federal Head Start Funding and Test Scores

Panel (A) of Table 2 shows the effect of HS spending per child on combined test scores in math and reading. Column (1) reports the results for the main sample. The estimated coefficients in Panel (A) indicate that being exposed to an average-sized HS program at age four leads to a statistically significant 0.04 standard deviation increase in third grade test scores.[26]

Previous studies show that the returns to public education investments are higher for groups in the lower end of the skill distribution (e.g., Bitler et al., 2014). Given that male students tend to be lower achieving relative to females and are more likely to participate in special education programs, HS may yield larger returns for males compared to females. In

[25] Additional results are presented with HS exposure at ages three through eight in Section 7.

[26] I present the results on combined test scores in math and reading for simplicity. To provide additional insight, in the Appendix Table OA.2, I examine the effect of HS in math and reading test scores separately and show that the estimated coefficients are larger and more statistically precise for math scores (a 0.05 standard deviation increase, significant at 1%) relative to reading scores (a 0.03 standard deviation increase, significant at 10%).

the next two columns of Table 2, I analyze whether exposure to more generous HS funding improves third grade test scores differentially by gender. In Columns (2) and (3), I find that additional HS funding exposure is associated with improvements in test scores for both males and females, with slightly larger point estimates for males. However, the estimated coefficients are not statistically different from one another.

In the rest of the table, I present the estimates for non-Hispanic whites, blacks, and Hispanics, respectively. The race and ethnicity breakdown reveals that improvements for Hispanics are the main driver of the results. In particular, I find that exposure to an average size HS program leads to a 0.06 standard deviation increase in test scores in math and reading. These improvements are statistically significant at the 1% level. Similar to Currie and Thomas (1999), who find that participation in HS closes at least 25% of the gap in test scores between Hispanics and whites, my results suggest that around a 500 dollar increase in federal HS funding exposure closes more than 15% of the gap relative to the raw mean difference in test scores (= 0.06/0.378).[27]

While the results for Hispanics are positive, large, and statistically significant, the analogous estimates for whites suggest that these cohorts did not experience improvements in test scores with exposure to additional HS spending. In addition, the estimated effects for blacks are positive and economically significant, but statistically imprecise. This discrepancy is worth noting, given that previous literature shows positive findings of HS, especially for black children (e.g., Deming, 2009). However, it is important to keep in mind that the demographic composition of HS enrollees in Texas does not reflect the demographic composition at the national level. The next section reconciles potential explanations for the large effects among Hispanic students.

[27] Using Table 2, the average standardized test score for Hispanics is -0.449 and -0.071 for whites. The raw difference is 0.378.

5.1.1 Why does Federal Funding Improve Test Scores for Hispanics?

As noted in Section 3.2, Hispanics lag behind both blacks and whites academically. In contrast to black children, who are historically under-privileged compared to whites, Hispanics often live in immigrant, Spanish-speaking families and communities (Currie and Thomas, 1999). During my study period, Hispanics make up more than 50% of free-or-reduced-lunch certified third grade students in Texas; 37% of them have limited language proficiency. Table 1 also reports that Hispanics were exposed to more generous HS funding compared to whites and blacks. For example, average HS funding exposure for Hispanics was $686, with a standard deviation of 1174; by contrast funding exposure for blacks was approximately $300, with a standard deviation of 211. Additionally, communities with a high fraction of Hispanics experienced more funding expansions relative to the ones with high fraction of blacks, as shown in Appendix Table A.2.

There are at least four channels by which HS could be beneficial for Hispanics. First, HS could increase their exposure to English and develop their language skills at an early age. This could, in turn, affect their educational performance in both reading and math. Panel (C) of Table 2 reports that HS spending exposure significantly increases the likelihood of becoming language proficient. One might expect that improved language proficiency would improve reading skills more than math skills. Given that the results on math scores are larger than reading, this channel alone does not explain the overall pattern of the results (see Online Appendix Table OA.2).

Second, HS could help special needs children during childhood, reducing their risk of worsening later in their school lives. This is possible, as HS programs have been required by federal law to reserve 10% of their enrollment for children with disabilities. Panel (D) of Table 2 shows that additional funding significantly reduces the likelihood of having special education status in third grade for Hispanic students.

Third, considering that Hispanics have higher participation in HS in Texas compared to whites and blacks, additional funding could induce more Hispanic children into the program.

Table 3 presents enrollment effects of HS funding expansions among whites, blacks and Hispanics separately. Indeed, the enrollment effects of additional HS funding are much larger for Hispanics (4.8 percentage point increase in enrollment per poor Hispanic children) relative to whites (2 percentate points increase) and blacks (1.3 percentage points increase).

Finally, although not statistically testable, HS could promote cultural assimilation for Hispanics which, in turn helps children adapt to school more easily (Currie and Thomas, 1999; Bitler et al., 2014).

5.2 Interpretation and Magnitude of Estimates

The estimated coefficients reported in Table 2 represent intent-to-treat (ITT) effects, which can be interpreted as the average effect of HS funding exposure on economically disadvantaged children. My main sample consists of students who were free-or-reduced lunch certified. These students would have been eligible for HS as children if their families' socio-economic status had stayed appoximately the same from childhood to age nine. However, eligibility does not imply attendance. Due to capacity constraints, HS only serves about up to 50% of the eligible children in a given year. There is also an issue with incomplete take-up, meaning that not all eligible children actually enroll in the progam.

To make accurate comparisons with the literature, I attempt to convert the ITT estimate of HS funding on test scores to treatment effect on the treated (TOT). This conversion requires having a "first-stage" that provides an estimate of exposure to HS funding on the likelihood of participating in HS for the sample. Unfortunately, HS participation is not observed in the main dataset. As a proxy, I estimate the effect of HS funding increase on community-level enrollment to HS using the data compiled from the Program Information Reports (PIR). Table 3 reports the results using HS enrollment as a share of age- and income-eligible children from 1988 to 1994 as a dependent variable (unit of observation is community-year). This table shows that, after controlling for community and year fixed effects, an increase in funding to an average-sized HS program is associated with a 9.2

percentage point increase in enrollment in HS for all poor children.

Before taking the next step, it is important to remember that this exercise requires a strong assumption that all the funding increases translated into enrollment expansion. However, in the 1990s, additional funding was appropriated toward improvements in teacher qualifications, expansion of services for families, provision of more resources to the local agencies, and transition from part-time to full-time enrollment, as described in detail in Section 2.1. Because HS funding expansions also affect program quality, attempted conversion here could be interpreted as an upper bound of the effect of HS participation.

Scaling up the main test score impact of a 0.037 standard deviation increase in math and reading, the implied effect of HS enrollment (0.092) at the local community level corresponds to a 0.4 standard deviation increase in test scores. This exercise provides a comparison point with the literature; however, caution is needed when interpreting these estimates as the true effect of participating in HS.

I next compare the test score impact of HS from my paper with the impact from Deming (2009) and the impact for the four-year-old cohort from the HSIS (Puma et al., 2012). Overall, my study reports a substantially larger impact on third grade test scores compared to the estimated effects reported in Deming (2009), who finds a 0.133 standard deviation increase in test score in elementary school for children who participated in HS compared to their siblings who did not. Deming's findings potentially ignore spillover effects across siblings, which may bias the results downward.

Furthermore, the follow-up of the HSIS reports null effects for third grade test scores by comparing mean differences between treatment and control groups (Puma et al., 2012). As noted above, HSIS took place in the 2000s after the program expansions of 1990s. Additionally, reported effects in Puma et al. are not adjusted for the fact that around 40% of the control group get exposure to another type of preschool.

My analysis and the other papers mentioned above differ in terms of the study period, sample of interest, and identification strategy, which could potentially explain the larger findings here. Whereas Deming (2009) studies the HS program in the 1980s (pre-1990) and Puma et al. (2012) examines the random assignment in the 2000s at the national level, my paper focuses on the 1990s, during which HS evolved significantly in terms of quality and capacity. Moreover, the other two papers analyze nationally representative samples, but my sample consists of low-income children in Texas who are predominantly Hispanic. Lastly, my identification strategy differs from the other two papers by taking advantage of the funding expansions across counties and over time.

6 Discussion of Mechanisms

One channel by which HS funding increases could improve test scores is by serving more children. As expressed above, Table 3 reports the results which show that additional funding led to significant increases in HS enrollment for economically disadvantaged children in Texas.

A second potential channel is through improvements in existing program characteristics that could then lead to better academic outcomes. Prior research has shown that reductions in student-teacher ratios benefit students, particularly children from disadvantaged backgrounds (e.g., Krueger and Whitmore, 2001). To my knowledge, there is only one paper that examines the effect of program inputs on cognitive and non-cognitive skills in the HS literature. Examining the impact of different inputs in HS centers using the HSIS, which took place in the 2000s, Walters (2015) finds that teacher education, teacher certification, and class size are not associated with improvements in test scores. He shows that the key input which improves children's cognitive skills is the provision of full-time services at the center level.

To explore this, I examine the relationship between federal funding increases and program inputs such as child-teacher ratios, child-staff ratios, share of full-time enrollment, and director's salary,[28] as well as spending for education, health, nutrition, and social services. For this analysis, I employ data on child-teacher ratios and full-time enrollment from the

[28]Online Appendix Table OA.2 plots the average of these inputs from 1988 to 1994.

PIRs (available for 1988-1994), director's salary from the PIRs (available for 1992-1994), and program budgets for various types of spending available in the PCCOST data (available for 1993-1994 for some programs).

Table 4 reports the findings on program inputs. The first two columns show that federal funding increases are associated with significant reductions in child-teacher and child-staff ratios. Column (3) shows that there is a significant relationship between funding expansions and the share of full-time enrollment at the local community-level. This finding is consistent with Walters (2015).

If additional funding were used to increase directors' salary, it may be considered wasteful, since directors already made considerably high salaries relative to the rest of the HS staff. Column (4) shows that funding increases are not significantly associated with directors' salary increases.

In the last four columns of Table 4, I show that when the federal government doubles HS funding, spending on education increases by around \$259,000, spending on health services increases by \$43,000, spending on nutrition increases by \$4,000, and spending on social services increases by \$59,000. Spending for educational services makes up to 75% of all spending, and it accounts for about 7% of the marginal increase. Overall, these results suggest that on the margin, federal funding partially goes to spending for services that might improve education and health development for children. However, due to data limitations, estimates are lacking statistical precision.

While not testable, the HS program has aimed to increase teacher qualifications by setting aside funds to improve the quality of teachers since the 1990s. Historical facts indicate that HS teachers took the opportunity to earn Associate or Bachelor degrees on child development and various related fields in the 1990s (HHS, 2000). Although it is not possible to provide quantitative evidence on this due to data limitation, quality improvements may have been partly driven by the improvements of teacher qualifications.[29]

[29]PIR data have information on the number of teachers with AA or BA degrees starting in 1999.

7 Robustness

In this section, I conduct a variety of robustness exercises to address potential concerns with the estimation strategy. Table 5 presents some sensitivity checks to the main specification. As a point of comparison, I include the baseline estimates in Column (1) that show the effect of HS exposure on third grade test scores. In Column (2), I omit local community-specific linear trends, which increases the size of the effect. Column (3) presents results with no pre-K controls, which barely decreases the magnitude of the effect. In Column (4), I excluded the controls for average county-level income per capita at the time of birth, at the time of HS, and at the time of test year. With this restriction, the estimated coefficient remains significant but the magnitude decreases slightly. In the next column, I omit controls for the county-level measures of per-capita transfer payments for cash income support, Food Stamps, medical care, retirement, and disability programs, and find that the magnitude of the estimate goes down without these controls. This is expected, because HS funding expansions are positively correlated with generosity of transfer payments. Excluding these important controls biases the main effect downward.

In Column (6), I add language proficiency and special education as additional controls. This analysis is essentially to check the sensitivity of the main results to the addition of "bad controls." Angrist and Pischke (2008) point out the issues related to inclusion of potential outcomes as additional controls. Since HS funding expansions significantly affect language proficiency and special education status, these variables would be considered bad controls. Nevertheless, it is reassuring to find that the main effect is smaller, but remains statistically signifant, when these variables are added as controls.

In the last column, I include school-specific linear trends to control for possible improvements in students' neighborhoods. Adding these trends leads to smaller coefficients with similar standard errors, which means that school trends are correlated with the changes in HS funding. Since the overall effect is relatively sensitive to adding school trends, I replicate the main results adding school trends. Online Appendix Table OA.3 presents these estimates which show that the effect for Hispanics remains statistically significant, though the effect size is slightly smaller.

I next analyze whether the results are sensitive to different measures of HS exposure, and find that the results are consistent across all. In Column (1) of Appendix Table A.4, I show the results obtained with my preferred measure, and in the next two columns I present results obtained with two alternatives. Instead of using the age-eligible children in the denominator, I first use total population and then use estimates of age- and income-eligible children. The second measure, "HS per capita," captures the exposure of the program for the whole county population. Since HS is a community-wide program which provides extensive services not only to children but also to their families, the results using this measure could be interpreted as the effect of community-wide exposure to HS. Column (2) reports the results using HS per capita as the main right hand side variable. The third measure, "HS per poor child," captures the size of the HS program, as HS targets children from families living under poverty. As discussed above, the ideal denominator to capture the size of the HS program is age- and income-eligible children. However, the data on counts of children living in poverty at the local level are not available for each year and each age over the period of my study.[30] Due to limited data availability, this measure is subject to measurement error. In Column (3) I present the results using HS per poor child.[31] Regardless of the differences in the measures, this table shows that being exposed to an average-sized HS program leads to similar test score gains.

As a falsification exercise, I analyze the effect of exposure to HS on test scores for the sample of children who are not identified as economically disadvantaged (not certified for

[30]The best available data on the county population of children in poverty are from the Small Area Income and Poverty Estimates (SAIPE) of the U.S. Census Bureau, for years 1989, 1993, 1995, 1997-1999. These data report estimated counts for the number of children 0-17 and children 5-17. I take the difference between these two estimates to construct the population 0-5 living in poverty. The number of poor children age three or four is two-fifth of the number of children under age five in poverty (Frisvold, 2006). However, this measure overestimates the true number of children based on the reported state numbers. More details can be found in Data Appendix A.1.

[31]In Online Appendix Tables OA.4, I show the results using test scores by gender and ethnicity using HS per capita in Panel A and HS per poor child in Panel B. These results strictly follows the patterns of the estimates in Table 2.

free-or-reduced lunch and not living under poverty), who are not likely to benefit from HS. I present these results in Appendix Table A.5 and show that HS exposure does not affect test scores for this group.

To check for potential selection effects, I test whether increases in federal HS spending affect the composition of a particular grade within a school (similar to Carrell and Hoekstra (2010)). If the variation in HS spending is not correlated with selection into the sample, I would expect to find no correlation. Appendix Table A.6 presents regression results in which I regress exogenous student characteristics on HS exposure conditional on school fixed effects. Results suggest that there is no evidence that the composition of a particular cohort (probability of being a specific race or gender, likelihood of being certified for free-or-reduced lunch, and income composition of the county) is correlated with exposure to HS conditional on school fixed effects. In the last column, I test whether HS funding increases are associated with the "predicted test scores" using the observable characteristics of students and find that there is no association between them. Overall, these tests suggest that the main results are not driven by selection bias.

Since HS serves children between the ages of three and five, one would not expect exposure to HS that occurred during other ages to be associated with improvements in outcomes (Thompson, 2018). As a falsification test, I estimate models where I analyze the exposure of HS at different ages. Table 6 reports results from specifications that use the sample of third graders with the assignment of exposure to HS changing from age three through age eight as the right-hand-side variable. These findings suggest that HS exposure at ages three and four matters the most, with the largest impact on test scores coming from exposure at age four. This is expected, considering that four-year-olds make up around 50% of total children served in HS (HHS, 2000).

In sum, this section presents evidence that the results are robust to excluding community-specific linear trends, do not appear to be the result of the expansion of the alternative early childhood programs during this time period, and are not driven by the selection-bias. Also, the results show that HS funding exposure at age four leads to improvements in third grade test scores, but not exposure at other ages.

8 Cost-Benefit Analysis

This section uses the results from Section 5.1 to provide a back-of-the-envelope calculation of cost-benefit analysis of federal HS spending expansions. Several studies have attempted to calculate the social benefits of the HS program and have shown that, in most cases, the program passes a cost-benefit test. However, as stated in Elango et al. (2015), this is a challenging exercise and it requires strong assumptions. Here, I attempt to obtain the costs and benefits associated with a 500 dollar increase in federal HS funding, assuming that the HS program accrues only test score gains in third grade and that these estimates will translate into later earnings.[32] This analysis adopts the cost-benefit formulation that is constructed by Kline and Walters (2016) for one year of HS attendance using the Head Start Impact Study. All monetary values are converted to 2014 dollars.

The marginal cost is $500 per child, since the analysis is based on the test-score impact of a 500 dollar increase in federal funding per child. To calculate the marginal benefit, I need two parameters: (1) the potential link between test scores and earnings, and (2) a prediction of average earnings for my sample (students who were free-or-reduced lunch certified in Texas). Although I cannot directly measure the impact on earnings due to data limitations, other studies that examine the link between test scores and earnings provide estimates. Following Kline and Walters (2016), I use a conservative estimate that earnings rise by 10 percent for each standard deviation increase in test scores.

Chetty et al. (2011) calculate that the present value of earnings at age 12 for the average individual in the U.S. is approximately $566,720 (in 2014$). The average present discounted

[32]Following Kline and Walters (2016), I assume that there are no effects on crime, health, or grade repetition, or no impacts on parents that raise benefits of the return of the program. This is an unrealistic assumption, considering that Carneiro and Ginja (2014) find large and significant health and behavioral effects for cohorts attended in HS in similar years.

value of the predicted earnings at age four corresponds to around \$434,000, with a discount rate of 3%. Adjusting for the fact that the median earnings in Texas are about 94 percent of the median earnings in the U.S., it corresponds to \$407,960.[33] Children who participate in free-or-reduced lunch are economically disadvantaged and likely to earn less than the median earner. As an approximation, the median income for families at or below 150 percent of the federal poverty level is 38 percent of the average in Texas.[34] Using the estimate for intergenerational income elasticity reported by Lee and Solon (2009) of 0.4, the average child in free-or-reduced lunch is expected to earn 75 percent of the average $((1-(1-0.38)*0.4))$. These predictions yield a present value of earnings of approximately \$305,000.

Putting the pieces together, 10% of \$305,000 is \$30,500, and multiplying it with the test score impact of 0.04 yields roughly \$1,220 of projected earnings impact.

These calculations show that a conservative estimate of the benefit-cost ratio is roughly 2, which is above the estimated value of one year of HS attendance reported in Kline and Walters (2016) for the HSIS cohorts (between 1.50 and 1.84). This estimate of 2 is much larger than the estimated rates of return associated with the Earned Income Tax Credit (0.88) and the Food Stamps (0.66) reported in Hendren (2016).

9 Conclusion

This paper provides new evidence on the old debate of whether early childhood investments for economically disadvantaged children narrow the academic achivement gap in elementary school. Many early childhood interventions have been found to be effective in improving children's skill development initially, yet the effects disappear by the time students reach third grade (Bailey et al., 2017). Utilizing previously unexplored variation in the federal funding expansions for the Head Start (HS) program in the 1990s and using student-

[33] Using the Current Population Survey Annual Social and Economic Supplement (CPS ASEC), between 1988 and 1994, the median earnings in the U.S. was \$25,310 (in 2014\$), while in Texas it was approximately \$23,814 (in 2014\$).

[34] Between 1988 and 1994, in Texas the median earnings was about \$25,310, but families with incomes at or below 150 percent of the poverty made \$9,584 using the CPS.

level administrative data from Texas, I find that exposure to more generous HS funding during childhood led to substantial gains in test scores.

Head Start (HS) has served low-income children for more than 50 years to reduce education and health disparities across socio-economic groups. During the 1990s, the federal government quadrupled the funding for HS, with the aim of improving both the program's quality and capacity. The significant federal funding increase within a relatively short time created a natural experiment that resulted in large variation of the adoption of funding expansions across communities and over time. Combining several sources on program characteristics and budgets, I show that additional HS funding led to significant increases in HS participation rates. Moreover, federal funding expansions improved child-teacher ratios, child-staff ratios, full-time enrollment, and education spending in the HS programs. These findings provide new evidence on the ways public funds are spent and suggest that both program capacity and quality improvements are important pathways for the ultimate effects on test scores.

The main finding of test score improvements is driven by the gains of low-income Hispanic students. In particular, I show that being exposed to an average-sized HS program led to a 15% reduction in the test score gap in math and reading combined between Hispanics and whites. For Hispanics, my results suggest that additional funding improved language proficiency and reduced the likelihood of special education status, which could partly explain the test score gains for this group. I also show that increased funding led to large and significant increases in HS participation rates among Hispanics.

Early childhood investments are receiving significant political attention. Therefore, it is important for policymakers to have credible estimates for the benefits of the programs - not only on cognitive outcomes, but also on non-cognitive, health, and labor market outcomes. My findings, which indicate that HS passes the cost-benefit test by a wide margin, provide useful insight for policymakers considering future public investments in early childhood education.

References

Andrews, R. J., Jargowsky, P., and Kuhne, K. (2012). The Effects of Texas's Targeted Pre-Kindergarten Program on Academic Performance. NBER Working Paper No. 18598, National Bureau of Economic Research.

Angrist, J. D. and Pischke, J.-S. (2008). Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.

Bailey, D., Duncan, G. J., Odgers, C. L., and Yu, W. (2017). Persistence and Fadeout in the Impacts of Child and Adolescent Interventions. Journal of Research on Educational Effectiveness, 10(1):7–39.

Bailey, M. J., Sun, S., and Timpe, B. (2018). Prep School for Poor Kids: The Long-Run Impacts of Head Start on Human Capital and Economic Self-Sufficiency. Working paper.

Barnett, W. S. (1995). Long-Term Effects of Early Childhood Programs on Cognitive and School Outcomes. The Future of Children, 5(3):pp. 25–50.

Barnett, W. S. (2011). Effectiveness of Early Educational Intervention. Science, 333(6045):975–978.

Barnett, W. S. and Hustedt, J. T. (2005). Head Start's Lasting Benefits. Infants & Young Children, 18(1):16–24.

Barr, A. and Gibbs, C. (2018). Breaking the Cycle? The Intergenerational Effects of Anti-Poverty Program in Early Childhood. Working paper.

Bauer, L. and Schanzenbach, D. (2016). The Long-Term Impact of the Head Start Program. Technical report, Hamilton Project.

Bitler, M. P., Hoynes, H. W., and Domina, T. (2014). Experimental Evidence on Distributional Effects of Head Start. Technical report, National Bureau of Economic Research.

Black, S. E., Breining, S., Figlio, D. N., Guryan, J., Karbownik, K., Nielsen, H. S., Roth, J., and Simonsen, M. (2017). Sibling Spillovers. Technical report, National Bureau of Economic Research.

Blau, D. and Currie, J. (2006). Pre-school, Day Care, and After-school Care: Who's Minding the Kids? Handbook of the Economics of Education, 2:1163–1278.

Carneiro, P. and Ginja, R. (2014). Long-Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start. American Economic Journal: Economic Policy.

Carnoy, M. and Loeb, S. (2002). Does External Accountability Affect Student Outcomes? A Cross-State Analysis. Educational Evaluation and Policy Analysis, 24(4):305–331.

Carrell, S. E. and Hoekstra, M. L. (2010). Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids. American Economic Journal: Applied Economics, 2(1):211–228.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., and Yagan, D. (2011). How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star*. Quarterly Journal of Economics, 126(4).

Currie, J. (2001). Early Childhood Education Programs. Journal of Economic Perspectives, pages 213–238.

Currie, J. and Neidell, M. (2007). Getting Inside the Black Box of Head Start Quality: What Matters and What Doesn't. Economics of Education Review, 26(1):83–99.

Currie, J. and Thomas, D. (1995). Does Head Start Make a Difference? American Economic Review, 85(3):341–64.

Currie, J. and Thomas, D. (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.

Duncan, G. J. and Magnuson, K. (2013). Investing in Preschool Programs. The Journal of Economic Perspectives, 27(2):109–132.

Duncan, G. J. and Murnane, R. J. (2011). Whither Opportunity?: Rising Inequality, Schools, and Children's Life Chances. Russell Sage Foundation.

Elango, S., García, J. L., Heckman, J. J., and Hojman, A. (2015). Early Childhood Education. Technical report, National Bureau of Economic Research.

Feller, A., Grindal, T., Miratrix, L. W., and Page, L. (2016). Compared to What? Variation in the Impact of Early Childhood Education by Alternative Care-type Settings. Annals of Applied Statistics.

Figlio, D. N. and Ozek, U. (2019). An Extra Year to Learn English? Early Grade Retention and the Human Capital Development of English Learners. Working Paper 25472, National Bureau of Economic Research.

Frisvold, D. E. (2006). Head Start Participation and Childhood Obesity. Vanderbilt University Economics Working Paper No. 06-WG01.

Garces, E., Thomas, D., and Currie, J. (2002). Longer-Term Effects of Head Start. American Economic Review, 92(4):pp. 999–1012.

Gibbs, C., Ludwig, J., and Miller, D. L. (2013). Head Start Origins and Impacts. Legacies of the War on Poverty, pages 39–65.

Gibbs, C. R. (2017). Does Full-day Kindergarten Reduce Achievement Gaps? Technical report, University of Wisconsin-Madison, Institute for Research on Poverty Focus Newsletter Vol.33 no.2.

Hendren, N. (2016). The Policy Elasticity. Tax Policy and the Economy, 30(1):51–89.

Hoynes, H. W. and Schanzenbach, D. W. (2012). Work Incentives and the Food Stamp Program. Journal of Public Economics, 96(1):151–162.

Jackson, C. K. (2018). Does School Spending Matter? The New Literature on an Old Question. Working paper, National Bureau of Economic Research.

Johnson, R. C. and Jackson, C. K. (2019). Reducing Inequality Through Dynamic Complementarity: Evidence from Head Start and Public School Spending. American Economic Journal: Economic Policy, forthcoming.

Kline, P. and Walters, C. (2016). Evaluating Public Programs with Close Substitutes: The Case of Head Start. The Quarterly Journal of Economics, 131(4).

Krueger, A. B. and Whitmore, D. M. (2001). The Effect of Attending a Small Class in the Early Grades on College-test Taking and Middle School Test Results: Evidence from Project STAR. The Economic Journal, 111(468):1–28.

Lee, C.-I. and Solon, G. (2009). Trends in Intergenerational Income Mobility. The Review of Economics and Statistics, 91(4):766–772.

Ludwig, J. and Miller, D. L. (2007). Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design. The Quarterly Journal of Economics, 122(1):159–208.

Ludwig, J. and Phillips, D. A. (2008). Long-Term Effects of Head Start on Low-income Children. Annals of the New York Academy of Sciences, 1136(1):257–268.

Miller, D. L., Shenhav, N., and Grozs, M. Z. (2018). Selection into Identification in Fixed Effects Models, with Application to Head Start. Working paper.

Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., Mashburn, A., and Downer, J. (2012). Third Grade Follow-Up to the Head Start Impact Study: Final Report. OPRE Report 2012-45. Administration for Children & Families.

Reardon, S. F., Robinson, J. P., and Weathers, E. S. (2008). Patterns and trends in racial/ethnic and socioeconomic academic achievement gaps. Handbook of Research in Education Finance and Policy.

Richardson, J. T. (2010). Accountability Incentives and Academic Achievement: The Benefit of Setting Standards Low. University of California, Davis, Job Market Paper.

Sanders, N. J. (2012). What Doesn't Kill You Makes You Weaker: Prenatal Pollution Exposure and Educational Outcomes. Journal of Human Resources, 47(3):826–850.

Shager, H. M., Schindler, H. S., Magnuson, K. A., Duncan, G. J., Yoshikawa, H., and Hart, C. M. (2013). Can Research Design Explain Variation in Head Start Research Results? a Meta-Analysis of Cognitive and Achievement Outcomes. Educational Evaluation and Policy Analysis, 35(1):76–95.

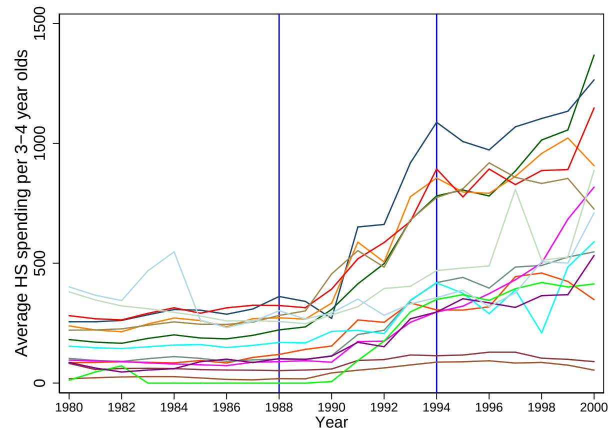
Thompson, O. (2018). Head Start's Long-Run Impact Evidence from the Program's Introduction. Journal of Human Resources, 53(4):1100–1139.

Walters, C. (2015). Inputs in the Production of Early Childhood Human Capital: Evidence from Head Start. American Economic Journal: Applied Economics, 7(4):76–102.

Zhai, F., Brooks-Gunn, J., and Waldfogel, J. (2014). Head Start's Impact is Contingent on Alternative Type of Care in Comparison Group. Developmental Psychology, 50(12):2572.

10 Figures

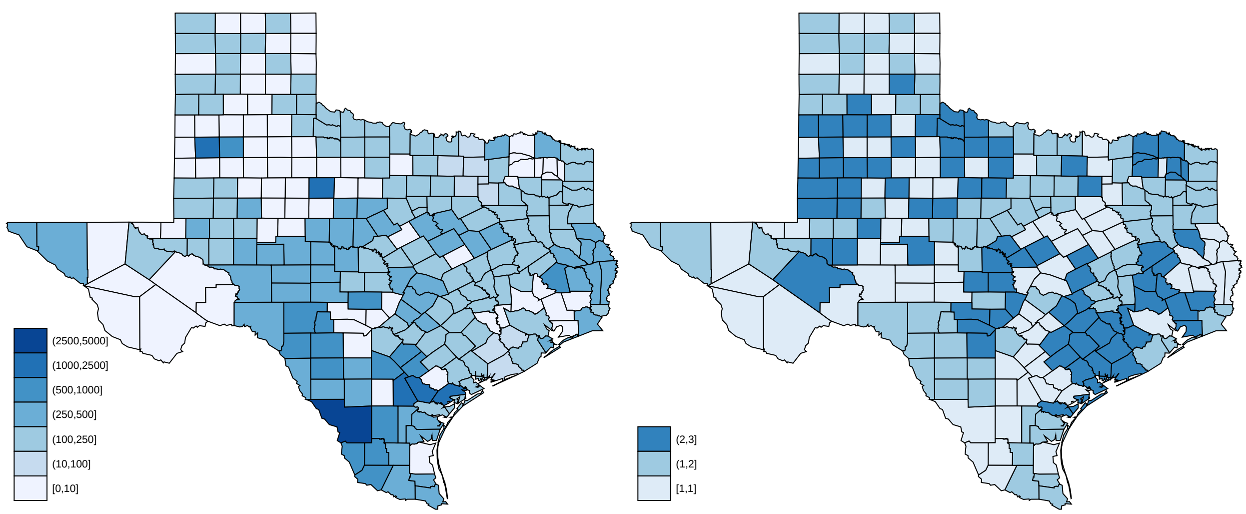
Figure 1: Head Start Funding per Child in the 15 Most Populous Counties in Texas



Notes: Head Start spending (in 2014$) data are from the Consolidated Federal Funds Reports (CFFR), coupled with the population counts for three- and four-year-olds at the county-level from the SEER. For more details about data construction, see Section 3. 15 most populous counties include Bend, Bexar, Brazoria, Cameron, Collin, Dallas, Denton, El Paso, Fort, Harris, Hidalgo, Montgomery, Nueces, Tarrant, Travis, and Williamson. Vertical lines (1988-1994) indicate the period of this study.

Figure 2: Geographic Variation in Head Start Funding per Child

(a) Spending per child in 1988 (b) Growth in Spending per child 1988-1994

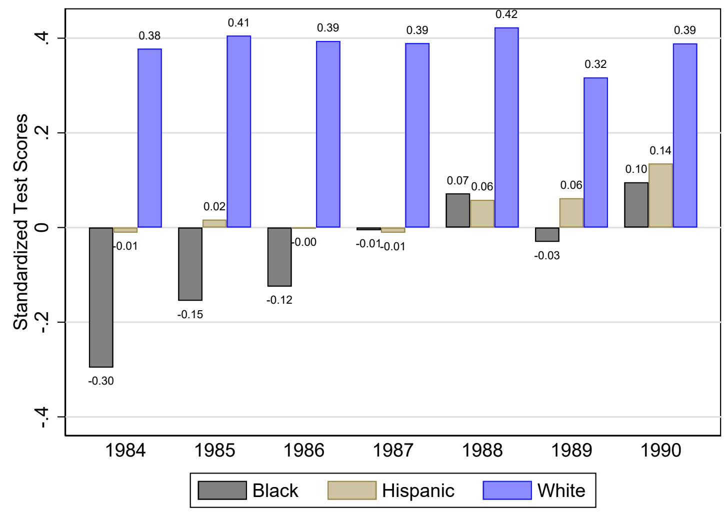


Notes: Head Start spending (in 2014$) data are from the Consolidated Federal Funds Reports (CFFR), coupled with the population counts for three- and four-year-olds at the county-level from the SEER. Growth is calculated using 1988 as the base period. For more details about data construction, see Section 3.

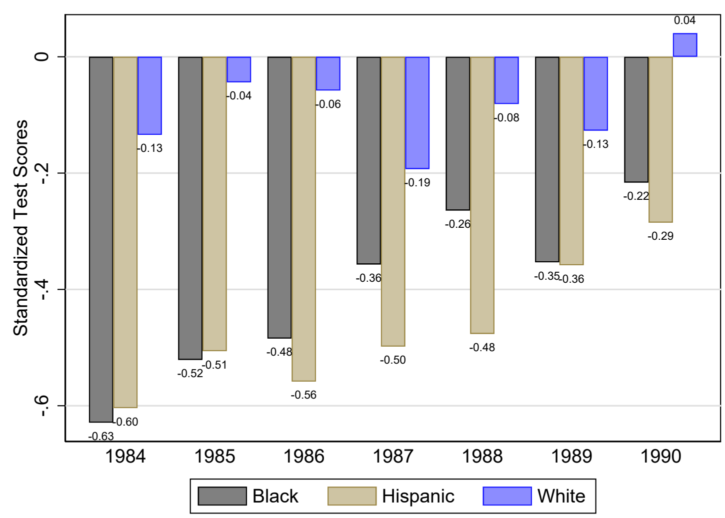
Figure 3: Histogram of Standardized Test Scores across Birth Cohorts,

by Free-or-reduced Lunch Status

(a) Not Free/Reduced Lunch Certified



(b) Free/Reduced Lunch Certified



Notes: Test scores are standardized combined math and reading scores. These data include all third grade students who took the standardized test in Texas between 1994 and 1999, from the Texas Education Agency (TEA). The sample is divided into two groups: (i) students who are not identified as economically disadvantaged and (ii) students who are certified for free-or-reduced lunch or who are identified as economically disadvantaged based on their families' welfare eligibility.

11 Tables

Table 1: Sample Characteristics of Third Grade Students in Texas

	Full Sample	Not Disav. Sample	Free/Reduced Lunch Eligible Sample					
			All	Male	Female	White	Black	Hispanic
Head Start								
Head Start Funding per Child	401.67	312.28	517.11	517.65	516.57	335.27	297.57	685.58
	(688.78)	(431.79)	(907.01)	(903.45)	(910.51)	(362.05)	(210.80)	(1174.54)
Outcomes								
Reading and Math Composite Score	0.02	0.31	-0.36	-0.46	-0.25	-0.07	-0.38	-0.45
	(0.99)	(0.77)	(1.11)	(1.16)	(1.05)	(0.99)	(1.04)	(1.15)
Standardized Reading Score	0.02	0.30	-0.34	-0.47	-0.21	-0.08	-0.33	-0.43
	(0.99)	(0.78)	(1.11)	(1.16)	(1.04)	(1.03)	(1.05)	(1.14)
Standardized Math Score	0.02	0.29	-0.33	-0.40	-0.27	-0.05	-0.39	-0.41
	(0.99)	(0.77)	(1.12)	(1.16)	(1.06)	(0.97)	(1.03)	(1.17)
Limited English Proficiency	0.10	0.02	0.21	0.22	0.20	0.01	0.00	0.37
	(0.30)	(0.14)	(0.41)	(0.41)	(0.40)	(0.09)	(0.07)	(0.48)
Participates in a Special Education Program	0.09	0.07	0.11	0.15	0.07	0.14	0.12	0.10
	(0.29)	(0.26)	(0.31)	(0.36)	(0.26)	(0.35)	(0.32)	(0.30)
Controls								
Year of Birth	1987.18	1987.08	1987.30	1987.31	1987.29	1987.26	1987.32	1987.28
	(2.06)	(2.05)	(2.06)	(2.05)	(2.06)	(2.07)	(2.04)	(2.06)
Female	0.50	0.49	0.50	0.00	1.00	0.50	0.51	0.50
	(0.50)	(0.50)	(0.50)	(0.00)	(0.00)	(0.50)	(0.50)	(0.50)
White, not of Hispanic Origin	0.53	0.78	0.22	0.22	0.22	1.00	0.00	0.00
	(0.50)	(0.41)	(0.41)	(0.41)	(0.41)	(0.00)	(0.00)	(0.00)
African American	0.14	0.07	0.23	0.22	0.23	0.00	1.00	0.00
	(0.35)	(0.26)	(0.42)	(0.42)	(0.42)	(0.00)	(0.00)	(0.00)
Hispanic	0.31	0.12	0.54	0.55	0.54	0.00	0.00	1.00
	(0.46)	(0.33)	(0.50)	(0.50)	(0.50)	(0.00)	(0.00)	(0.00)
Free/Reduced Lunch Eligible	0.44	0.00	1.00	1.00	1.00	1.00	1.00	1.00
	(0.50)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Bilingual	0.06	0.01	0.13	0.14	0.13	0.00	0.00	0.25
	(0.24)	(0.07)	(0.34)	(0.34)	(0.34)	(0.04)	(0.02)	(0.43)
Participates in an ESL program	0.03	0.01	0.06	0.06	0.06	0.01	0.00	0.09
	(0.17)	(0.10)	(0.23)	(0.24)	(0.23)	(0.08)	(0.06)	(0.29)
Participates in a Gifted/Talented program	0.06	0.08	0.03	0.03	0.03	0.02	0.03	0.03
	(0.24)	(0.28)	(0.17)	(0.16)	(0.18)	(0.14)	(0.17)	(0.18)
N	762815	429905	332910	165303	167607	72073	75614	180775

Notes: Student data are from the Texas Education Agency (TEA) which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 and 1999 for students in third grade with non-missing demographic characteristics. HS spending (in 2014$) data are from the Consolidated Federal Files Reports (CFFR) and include years between 1988 and 1994. For detailed description, see Section 3.

Table 2: Baseline Estimates of the Effect of Head Start Funding Exposure on Student Outcomes in Third Grade

	All	Males	Females	Whites	Blacks	Hispanics
	(1)	(2)	(3)	(4)	(5)	(6)
A: Reading and Math Composite Score						
Head Start Funding	0.037**	0.042*	0.036*	-0.020	0.054	0.060***
	(0.015)	(0.025)	(0.020)	(0.048)	(0.045)	(0.014)
Mean Y	-0.356	-0.459	-0.253	-0.071	-0.382	-0.449
B: Language Proficiency						
Head Start Funding	0.009**	0.010**	0.008***	-0.000	-0.004	0.010**
	(0.004)	(0.005)	(0.003)	(0.001)	(0.004)	(0.004)
% Effect (Coef/Mean)	0.92	1.02	0.82	-0.01	-0.74	0.80
Mean Y	0.789	0.783	0.795	0.992	0.995	0.631
C: Special Education						
Head Start Funding	-0.006	-0.009	-0.008*	-0.001	-0.032**	-0.011**
	(0.004)	(0.006)	(0.004)	(0.013)	(0.014)	(0.005)
% Effect (Coef/Mean)	-0.66	-0.90	-0.80	-0.14	-5.86	-0.83
Mean Y	0.112	0.150	0.074	0.145	0.117	0.097
Mean Head Start Funding($)	517					
Obs	332910	165303	167607	72073	75614	180775

Notes: This table contains results obtained when the independent variable is real federal HS spending per child (2014$) when the child was four years old. HS spending per child is scaled by the mean spending ($517), thus the coefficients should be interpreted as the effect of exposure to an average-sized program. All regressions include controls for demographics and county-level characteristics, school, test year and birth year fixed effects, along with community-specific linear trends. Sample consists of students who are certified for free/reduced lunch or in poverty based on the description by the Texas Education Agency (TEA). Student data are from the TEA which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 and 1999. HS spending data are from the Consolidated Federal Files Reports (CFFR) and include years between 1988 and 1994. Standard errors are clustered at the county level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 3: The Effect of Head Start Funding per Child on Head Start Enrollment, 1988-1994

	Total Enrollment	Whites	Blacks	Hispanics
	(1)	(2)	(3)	(4)
Head Start Funding	0.092***	0.020***	0.013***	0.048***
	(0.020)	(0.007)	(0.004)	(0.013)
Mean Y	0.183	0.043	0.039	0.106
Obs	805	805	805	805

Notes: This table reports the estimated coeffients for enrollment per poor population of children 3 and 4 in Texas. The data are at the local community-level and the independent variable is real federal HS spending per child (2014$). HS spending per child is scaled by the mean spending ($517). Local community is defined as one or more counties that each HS grantee serves. Head Start spending data are from the Consolidated Federal Funds Reports (CFFR), coupled with the population counts for three- and four-year-olds at the county-level from the SEER. Poor children 3-4 counts are estimated using the SAIPE data. For more details about the data description, see Section 3. All regressions include local community and year fixed effects. Standard errors are clustered at the county level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 4: Mechanisms: The Effect of Head Start Funding on Program Characteristics and Budget

	Program Characteristics			Director	Program Budgets for Spending on			
	Child per Teacher	Child per Staff	Share of Full-Time Enrollee	Director's Salary	Education	Health	Nutrition	Social
Real HS Spending (in millions)	-0.198***	-0.077**	0.011***	0.306	0.069*	0.011	0.001	0.015
	(0.072)	(0.030)	(0.003)	(0.300)	(0.038)	(0.012)	(0.003)	(0.016)
Mean Y	25.319	12.383	0.540	71.947	2.760	0.366	0.177	0.291
Mean X (in millions)	3.717	3.717	3.717	4.774	3.762	3.762	3.762	3.762
Obs	509	509	509	264	142	142	142	142

Notes: The data are at the local community-level and the independent variable is real federal HS spending in millions (2014$). Local community is defined as one or more counties that each HS grantee serves. HS spending data are from the Consolidated Federal Funds Reports (CFFR). Program characteristics are from the PIRs and budget spending breakdown are from the PCCOST data. Data on budgets are only available for years 1993-1994 for some programs. Director's salary is available for years 1992-1994. The rest of the variables are available for years 1988-1994. For more details about the data description, see Section 3. All the monetary values in outcome variables are converted into 2014 dollars and they are in thousands. All regressions include local community, year fixed effects and 1980 community characteristics interacted with linear trends. Estimates are weighted using the number of three- and four-year-old children in a local community. Standard errors are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.

Table 5: The Effect of Head Start Exposure on Third Grade Standardized Test Scores

Sensitivity of Results to Specifications

	Main Results	Omit County Trends	Omit Pre-K Controls	Omit Income Controls	Omit Safety Net Controls	Add Lang Prof+Spec Ed Controls	Add School Trends
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Head Start Funding	0.037**	0.040**	0.036**	0.034**	0.021	0.023*	0.024*
	(0.015)	(0.016)	(0.015)	(0.015)	(0.014)	(0.013)	(0.014)
Mean Y	-0.356	-0.356	-0.356	-0.356	-0.356	-0.356	-0.356
Obs	332910	332910	332910	332910	332910	332910	332910

Notes: This table reports results obtained when the dependent variable is third grade combined standardized test scores in math and reading, and the independent variable is real federal HS spending per child (in 2014$) when the child was four years old. HS spending per child is scaled by the mean spending ($517), thus the coefficients should be interpreted as the effect of exposure to an average-sized program. All regressions include controls for demographics, school, test year and birth year fixed effects. Sample consists of students who are certified for free/reduced lunch or in poverty based on the description by the Texas Education Agency (TEA). Student data are from the TEA which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 and 1999. HS pending data are from the Consolidated Federal Files Reports (CFFR) and include years between 1988 to 1994. Standard errors are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.

Table 6: Falsification Test: The Effect of Head Start Funding Exposure on Third Grade Combined Standardized Test Scores - Differential Effects by Age of Exposure, 3-8

	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8
Head Start Funding	0.024*	0.037**	0.022	0.010	0.004	0.001
	(0.014)	(0.015)	(0.020)	(0.030)	(0.029)	(0.010)
Mean Y	-0.356	-0.356	-0.356	-0.356	-0.356	-0.356
Obs	332910	332910	332910	332910	332910	332910

Notes: This table reports results obtained when the dependent variable is third grade combined standardized test scores in math and reading, and the independent variable is real federal HS spending per child (in 2014$), assigned when the child was at ages 3 to 8 years old. HS spending per child is scaled by the mean spending ($517), thus the coefficients should be interpreted as the effect of exposure to an average-sized program. All regressions include controls for demographics and county-level characteristics, school, test year and birth year fixed effects, along with local community-specific linear trends. Sample consists of students who are certified for free/reduced lunch or in poverty based on the description by the Texas Education Agency (TEA). Student data are from the TEA which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 and 1999. HS spending data are from the Consolidated Federal Files Reports (CFFR) and include years between 1988 and 1994. Standard errors are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.

A Appendix

A.1 Data Appendix

Public Use Data

- Head Start Spending Data from Consolidated Federal Funds Reports (CFFR):

  Federal government expenditures data which are reported in states and counties of the U.S. It is collected under the authority of Title 13 of the U.S. Code and contains statistics on the geographic distribution of federal program expenditures including HS grants, using data submitted by federal departments and agencies (CFFR, 2010). Thus, the geographic level of analysis is chosen as local community since it is the available unit of observation in the data available.[35] HS is administered by the Department of Health and Human Services (HHS), Administration for Children and Families (ACF), and Office of Head Start (OHS). HHS describes “grantees” as the agencies that receive grant awards directly. “Delegates” are other agencies grantees may contract services.

  These data span from 1983-2010, available through the Census.[36] The program identification code for Head Start expenditures is 93.600. These expenditures are defined as:

  1. ADMINISTRATION FOR CHILDREN & FAMILIES-HEAD START
  2. ADMINISTRATION FOR CHILDREN, YOUTH AND FAMILIES_HEAD START
  3. ADMINISTRATION FOR CHILDREN, YOUTH, AND FAMILIES-HEAD START
  4. ADMINISTRATION FOR CHILDREN, YOUTH, AND FAMILIES-HEADSTART
  5. HEAD START

  Agencies that provide the Head Start grants are listed as:

[35] Source: `https://www.census.gov/prod/2011pubs/cffr-10.pdf`
[36] `http$://$www2.census.gov/pub/outgoing/govs/special60/`

1. IMMEDIATE OFFICE OF THE SECRETARY OF HEALTH AND HUMAN SERVICES
2. ADMINISTRATION FOR CHILDREN AND FAMILIES.

When the federal government announces grant availability for a specific local community, it also announces which counties each grantee is expected to serve. However, the retrospective grant announcements are not available to researchers. To determine the serving counties for each grantee, I use an administrative data set (PCCOST) provided by Currie and Neidell (2007) which includes detailed information on the allocation of total expenditures for health and other services for each grantee and its network. These data cover the years 1990-2000 with full coverage of networks for the year 1994. I assigned the networks based on the 1994 data. Based on my web search for a randomly chosen subset of grantees in Texas, I did not find evidence that the serving counties changed from 1988 to 1994. If the assignment of the networks is wrong for some counties, it will create measurement error in the main right hand side variable in my analysis. I then confirmed these networks of counties using the website of the grantees and an additional data set provided by Frisvold (2006). Frisvold (2006) constructed these networks of counties in 2005 using the website of the state's Head Start Association, the state's Head Start Collaboration Office, or through personal communication with a staff member in these organizations. However, this data set does not take into account the fact that the networks could have changed over time. In my work, I adjusted for that using the administrative PCCOST data. Figure A.5 maps out the raw data in Texas in 1994, with 69 grantees. These grantees serve more than 200 counties in Texas. Each network ranges between one and 14 counties and also 48 counties have zero HS dollars.

Additional data are used to confirm these networks, which are provided by Frisvold (2006). To further check the reliability of the assigned regions, I did web searches using the websites of the state's Head Start Association, the state's Head Start Collaboration Office, or through personal communication with a staff member at the Head Start grantees.

To give a concrete example of the construction of these networks, consider the headquarters of the Brazos Valley Community Action Agency in Texas, which was established in 1967,[37] located in Bryan, Texas (Brazos County). This grantee serves HS programs in eight other counties.[38] Thus, the raw data as shown in Appendix Figure A.6a records $4.3 million in expenditures for Brazos County in 1994 and zero dollars for all the serving counties. Using the network of counties that I constructed, I reallocated dollars for the serving counties in proportion to the number of age-eligible children in each county (see Appendix Figure A.6b for reallocation map).

- Population Counts: I use two separate data sets to construct the three different measures of the size of the HS program: HS spending per age-eligible child, per capita, and per poor child.

  1. County-level population data of children three and four years old are constructed using data from the Surveillance, Epidemiology, and End Results Program (SEER) which include county-level population counts for each age group starting 1969.[39]

  2. County-level population to construct the per capita measure also comes from SEER. Instead of extracting specific age groups, I collapse the entire population at the local community-year level.

  3. The number of poor children is from the Small Area Income and Poverty Estimates (SAIPE) of the U.S. Census Bureau. In the SAIPE data, county-level estimates of children under 17 and children 5-17 are available. Using these two variables, I construct the number of children under age 5 by taking the difference. To create

[37] http://www.bvcaa.org/history-of-bvcaa-inc/

[38] http://www.bvcaa.org/programs/head-startearly-head-start/

[39] Source: http://www.nber.org/data/seer_u.s._county_population_data.html.

age eligible poor child counts, I follow Frisvold (2006) which states that children who are age 3-4 years old are two fifths of children under age five. These data are only available for years 1989, 1993, 1995, 1997-1999, the years in between are determined through linear interpolation.

- Program Information Reports (PIR):[40] Starting in 1988, the Office of Head Start Programs has collected comprehensive data from all grantees and delegates on the services, staff, children, and families served by the program. These data are important for my analysis as they provide information on number of funded enrollees, number of staff, demographic composition of children and staff, qualifications of directors and teachers, and so on. I use this information to show how much the funding expansions translate into enrollment versus the quality of the HS programs. PIR data are not commonly used because the format of these data and variables collected changed over time. Part of these data from 1988 and 1998 were generously provided to me by Currie and Neidell (2007).

- Common Core of Data (CCD):[41] CCD includes the school level information for all public schools. These data are available starting 1986 at the school level and provide information on pupil teacher ratio, a measure used for education quality in the education literature, as well as the demographic composition of students and the grade levels offered in a specific school.

- County-level Demographics: This information is important for my analysis as there could be other confounding factors at the county-level that might affect the estimates, such as other War on Poverty programs that target preschool age children. Thus, I will add to my analysis county-level controls (income per capita and other government transfers including food stamps per capita and cash transfers per capita) at birth and at the time of the survey, which are collected from the Regional Economic Informa-

[40]Source: `https://eclkc.ohs.acf.hhs.gov/hslc/data/pir`
[41]Source: `https://nces.ed.gov/ccd/pubschuniv.asp`

tion Systems (REIS), and other county demographics including the poverty rate from Census. Also, the relationship between business cycles and child outcomes is well-established in the literature. To control for exposure to business cycles at birth, I will include county-level unemployment rates from the Bureau of Labor Statistics (BLS). Finally, to control for the composition of the demographics of the population, population counts at the county-level for racial and age groups, I will use data from the Surveillance, Epidemiology, and End Results Program (SEER).

Texas Education Agency (TEA)

I use student-level data from the TEA, which include information on test scores monitored through the Texas Academic Assessment System (TAAS) for grades 3 to 8 for years between 1994 and 2002. These data also contain information on gender, ethnicity, free-or-reduced lunch status, language proficiency and special education status on each student. The TEA started offering a Spanish version of the test for students with limited language proficiency in 2000. Explanations for some variables:

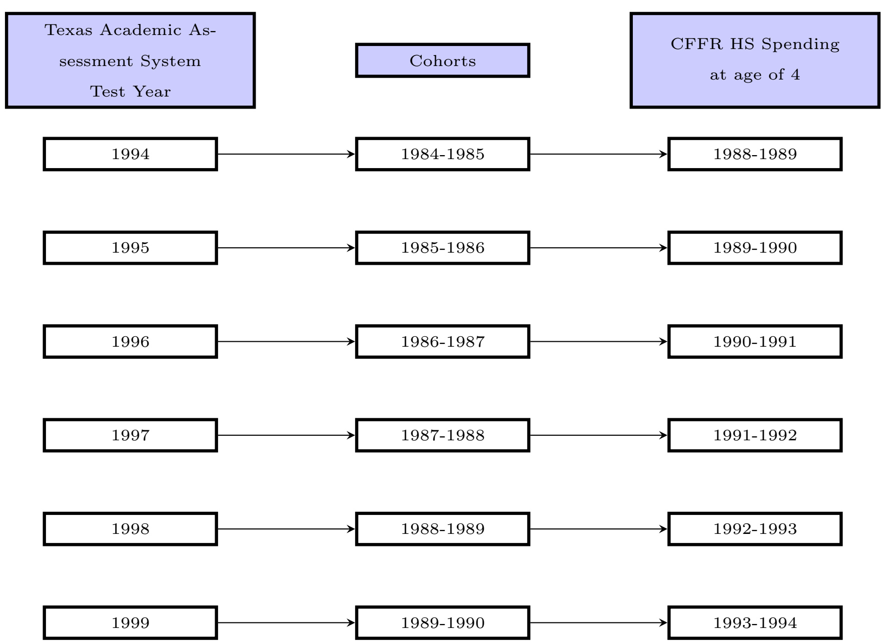
- Outcome variables: Texas Learning Index (TLI) reading and math scores
- Disadvantage variables:
  - 0: Not identified as economically disadvantaged
  - 1: certified for free meals under the National School Lunch and Child Nutrition Program
  - 2: Eligible for reduced-price meals under the National School Lunch and Child Nutrition Program
  - 9: Other economic disadvantage, including:
    * from a family with an annual income at or below the official federal poverty line

* eligible for Temporary Assistance to Needy Families (TANF) or other public assistance
* received a Pell Grant or comparable state program of need-based financial assistance
* eligible for programs assisted under Title II of the Job Training Partnership Act (JTPA)
* eligible for benefits under the Food Stamp Act of 1977

School district to county crosswalk is obtained from the TEA website: `http://mansfield.tea.state.tx.us/TEA.AskTED.Web/Forms/DownloadFile.aspx`. For schools which do not currently operate, I manually entered the county information gathered via web searches.

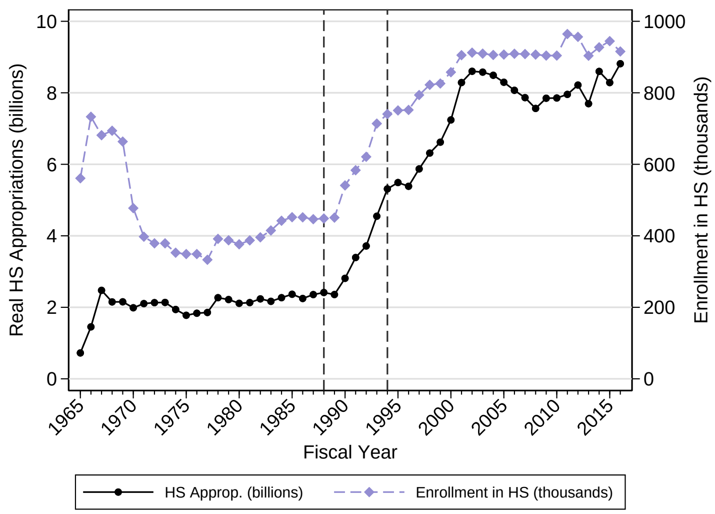
<u>Data Structure</u>

The chart below shows the correspondence of test years with cohorts and Head Start exposure of those cohorts at age four.



A.2 Figures

Figure A.1: Head Start Program Facts

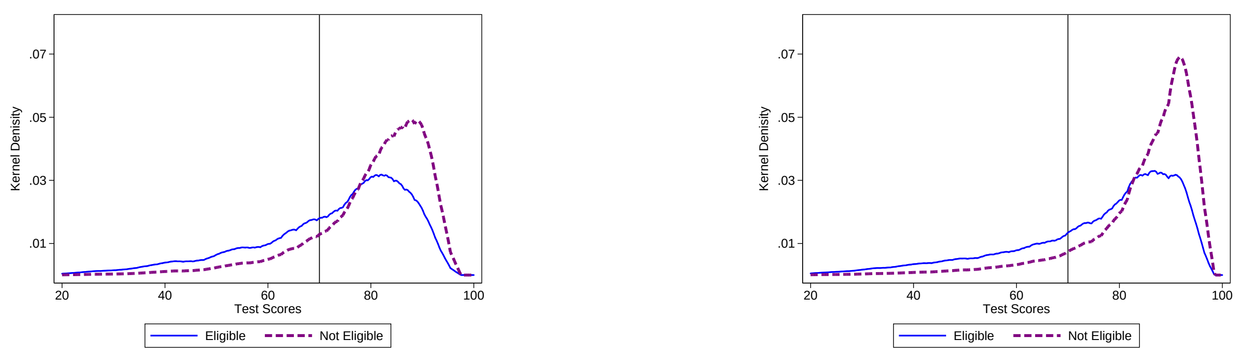


Notes: The data are from the HHS website: `https://eclkc.ohs.acf.hhs.gov/hslc/data/factsheets/2015-hs-program-factsheet.html`. Federal Head Start appropriations are in 2014 dollars. The dashed lines highlight the period of this study, from 1988 to 1994.

Figure A.2: Kernel Density of Test Scores, by free-or-reduced lunch Certification

(a) Math

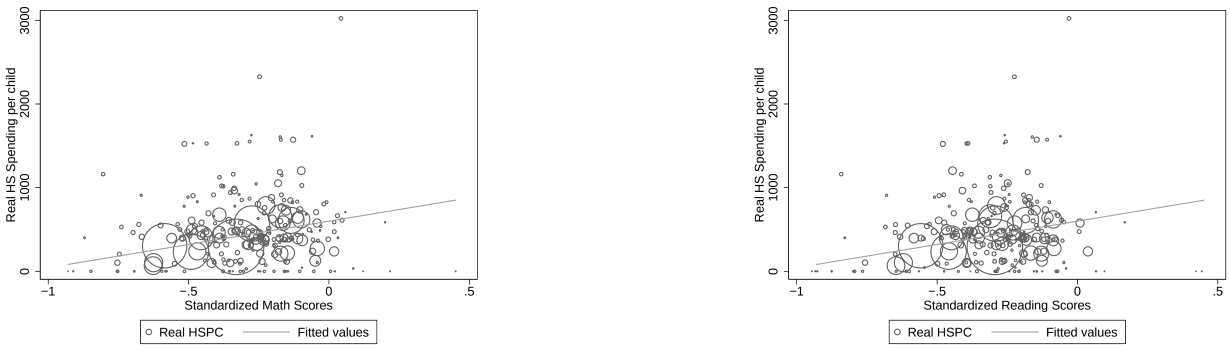
(b) Reading



Notes: Test score data include all third grade students who took the standardized test in Texas between 1994 and 1999, from the Texas Education Agency (TEA). The sample is divided into two groups: (i) students who are not identified as economically disadvantaged and (ii) students who are certified for free-or-reduced lunch or who are identified as economically disadvantaged based on their families' welfare eligibility. The minimum passing score is 70, determined by the TEA. Kernel density calculated using a bandwidth of two.

Figure A.3: Raw Correlations between Head Start Funding per Child and Standardized Test Scores - Free/Reduced Lunch Certified Sample

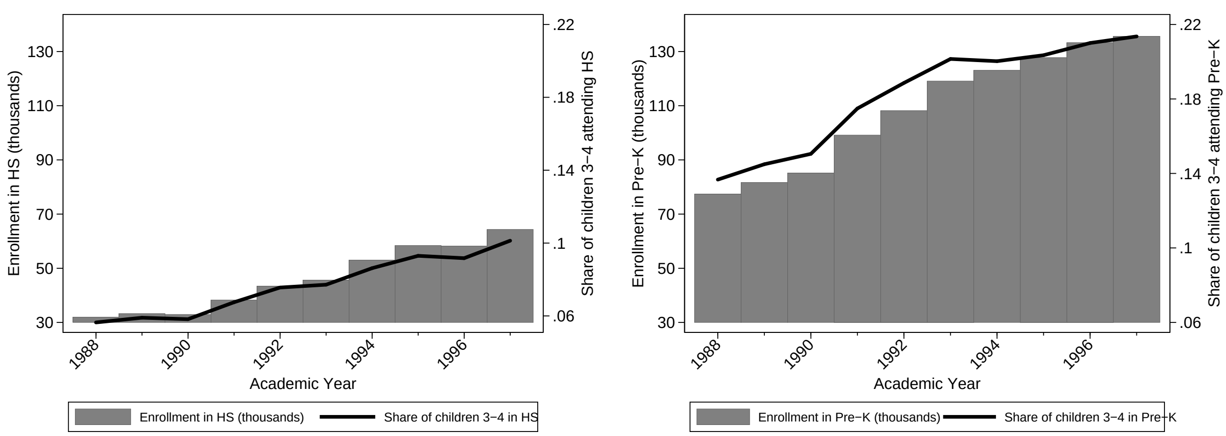
(a) Math (b) Reading



Notes: Head Start spending (in 2014$) data are obtained from the Consolidated Federal Funds Reports and third grade student test score data are from the Texas Education Agency (TEA) between 1994 and 1999. The data are collapsed to the county-level using averages. The bubbles present the local communities, weighted by the population of three- and four-year-olds.

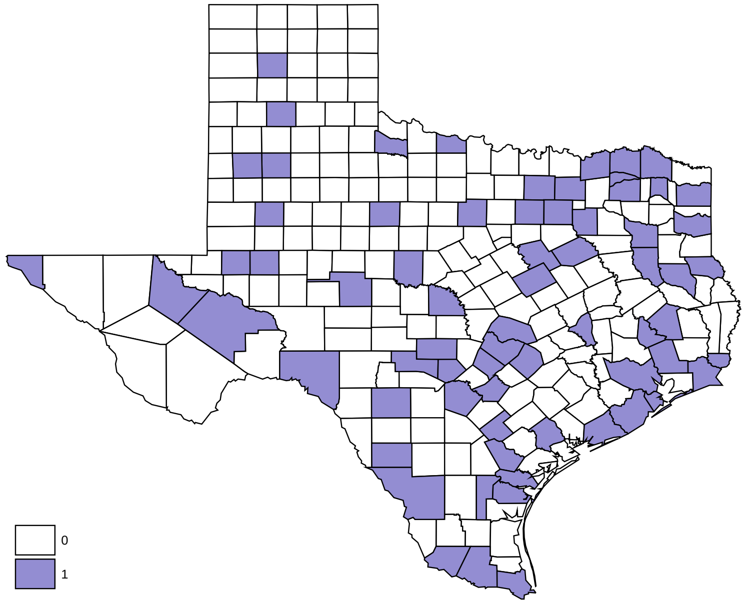
Figure A.4: Early Childhood Education Expansions in Texas

(a) HS expansions (b) Pre-K expansions



Notes: Head Start enrollment data are from the Program Information Reports (PIR). Pre-K enrollment data are from the Common Core Data (CCD). The share measure is calculated using the population counts for three- and four-year-olds from the SEER.

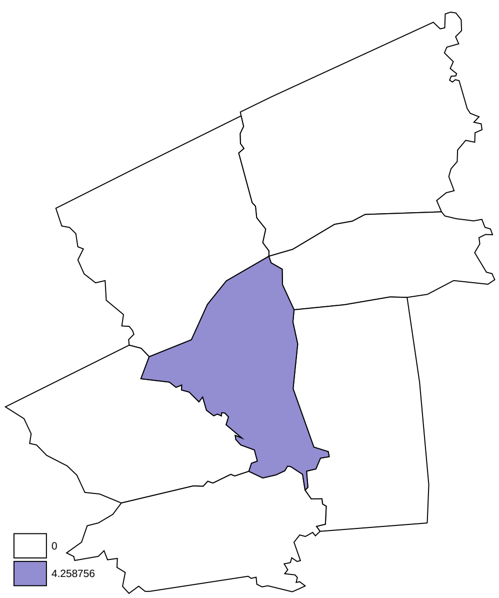
Figure A.5: Raw Data: Indication of Positive Head Start Spending in 1994



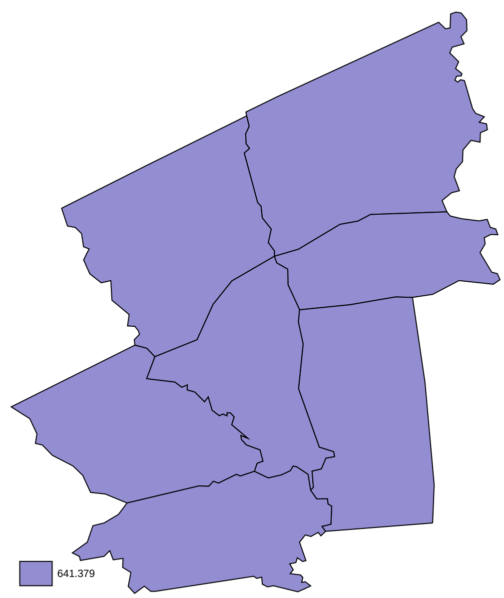
Notes: Raw federal Head Start spending data at the grantee-level, from Consolidated Federal Funds Reports. There are 69 grantees that served more than 200 counties in 1994.

Figure A.6: Brazos Valley Community Action Agency, Before and After Reallocation

(a) Raw Data



(b) After Allocation



Notes: Federal Head Start spending (in 2014$) data at the grantee-level are obtained from the Consolidated Federal Funds Reports, coupled with administrative data on the counties that each grantee serves (PCCOST data) from Currie and Neidell (2007). The figure on the left shows the raw federal funding data forBrazos Valley Community Action Agency in 1994. Using the PCCOST data, I determine the serving counties for this agency and distribute HS dollars at the local level based on the share of total age-eligible children living in a community.

A.3 Tables

Table A.1: Correlations between Community-Level Variables and Average Head Start Funding per Child, 1988-1994

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Percent Urban	0.473											5.419
	(2.102)											(4.953)
Percent Black		-19.700***										-0.411
		(7.414)										(17.044)
Percent Hispanic			9.709**									-2.373
			(4.524)									(6.498)
Percent Farmland				6.218***								7.771
				(2.033)								(4.859)
Education Expenditures per Capita					2.740**							1.970
					(1.294)							(1.437)
Welfare Expenditures per Capita						127.832						60.198
						(109.667)						(69.116)
Income per Capita							-0.064***					-0.136**
							(0.022)					(0.063)
Percent Single Mother								60.255				-61.389
								(39.080)				(155.728)
Percent of Children 0-18 under Poverty									24.562**			19.528
									(9.598)			(14.569)
Fraction of Pop Under 5										119.006		94.262
										(80.893)		(109.341)
Fraction of Pop Older than 65											21.081**	44.310
											(8.174)	(31.462)
Log population												245.163
												(185.484)
Mean Y($)	411	411	411	411	411	411	411	411	411	411	411	493
Mean X	78	12	22	64	273	2	10408	9	18	8	9	
R-Squared	0.000	0.090	0.191	0.082	0.082	0.094	0.103	0.070	0.199	0.053	0.021	0.249
Obs	117	117	117	117	117	117	117	117	117	117	117	117
F-test	0.051	7.061	4.605	9.352	4.485	1.359	8.588	2.377	6.549	2.164	6.652	2.069
p-value	0.822	0.009	0.034	0.003	0.036	0.246	0.004	0.126	0.012	0.144	0.011	0.024

Notes: The data are at the local community-level and the dependent variable is average federal HS spending per child (2014$). Local community is defined as one or more counties that each HS grantee serves. Head Start spending data are from the Consolidated Federal Funds Reports (CFFR), coupled with the population counts for three- and four-year-olds at the county-level from the SEER. The 1980 county controls are from City and County Data Book. Estimates are weighted using the 1980 county population. Standard errors are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.

Table A.2: Correlations between Community-Level Variables and Change in Head Start Funding per Child, 1988-1994

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Percent Urban	-0.700											3.816
	(0.749)											(2.572)
Percent Black		-7.652***										3.674
		(1.925)										(9.403)
Percent Hispanic			2.870***									-4.179
			(0.847)									(3.858)
Percent Farmland				1.686**								2.313
				(0.773)								(3.041)
Education Expenditures per Capita					1.422**							1.264
					(0.557)							(1.034)
Welfare Expenditures per Capita						22.238						-14.385
						(21.962)						(27.851)
Income per Capita							-0.023***					-0.064
							(0.005)					(0.045)
Percent Single Mother								14.222*				-89.023
								(8.529)				(88.389)
Percent of Children 0-18 under Poverty									8.115***			20.691**
									(2.265)			(9.999)
Fraction of Pop Under 5										29.081		70.482
										(19.135)		(73.127)
Fraction of Pop Older than 65											11.374**	25.603
											(4.495)	(20.774)
Log population												164.773
												(125.252)
Mean Y($)	176	176	176	176	176	176	176	176	176	176	176	256
Mean X	78	12	22	64	273	2	10408	9	18	8	9	
R-Squared	0.009	0.118	0.145	0.053	0.191	0.025	0.119	0.034	0.188	0.028	0.052	0.209
Obs	117	117	117	117	117	117	117	117	117	117	117	117
F-test	0.872	15.795	11.478	4.751	6.526	1.025	21.712	2.780	12.834	2.310	6.403	1.691
p-value	0.352	0.000	0.001	0.031	0.012	0.313	0.000	0.098	0.000	0.131	0.013	0.077

Notes: The data are at the local community-level and the dependent variable is long-change in real federal HS spending per child (2014$) from 1988 to 1994. Local community is defined as one or more counties that each HS grantee serves. Head Start spending data are from the Consolidated Federal Funds Reports (CFFR), coupled with the population counts for three- and four-year-olds at the county-level from the SEER. The 1980 county controls are from City and County Data Book. Estimates are weighted using the 1980 county population. Standard errors are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.

Table A.3: The Effect of Head Start Funding per Child on Pre-K Enrollment, 1988-1994

	Enrolled in Pre-K/Pop 3-4	Enrolled in Pre-K/Poor Pop 3-4
Head Start Funding	-0.004	0.008
	(0.003)	(0.016)
% Effect (Coef/Mean)	-0.37	0.84
Mean Y	0.174	0.524
Obs	805	805

Notes: The data are at the local community-level and the independent variable is real federal HS spending per child (2014$). Local community is defined as one or more counties that each HS grantee serves. Head Start spending data are from the Consolidated Federal Funds Reports (CFFR), coupled with the population counts for three- and four-year-olds at the county-level from the SEER. Poor children 3-4 counts are estimated using the SAIPE data. Pre-K enrollment data are from CCD, aggregated up to the local community-level. For more details about the data description, see Section 3. All regressions include local community and year fixed effects. Standard errors are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.

Table A.4: The Effect of Head Start Exposure on Third Grade Standardized Test Scores

Sensitivity of Results to Different Measures of Head Start Exposure

	Per Child (Age 3-4)	Per Capita	Per Poor Child (Age 3-4)
Head Start Funding Per Child (Main)	0.037**		
	(0.015)		
Head Start Funding Per Capita		0.030***	
		(0.009)	
Head Start Funding Per Poor Child			0.029**
			(0.013)
Mean Y	-0.356	-0.356	-0.356
Mean Funding Level ($)	517	18	1198
Obs	332910	332910	332910

Notes: This table contains results obtained when the dependent variable is third grade combined standardized test scores in math and reading. Each column reports results obtained using different independent variables: (1) federal Head Start spending per three- and four-year-old child, (2) federal Head Start spending per capita, and (3) federal Head Start spending per poor three- and four-year-old child. All the dollar values are in 2014 dollars. All HS spending measures are scaled by the mean spending, thus the coefficients should be interpreted as the effect of exposure to an average-sized program. The exposure variables are assigned based on the local community and year the child was four years old. All regressions include controls for demographics and county-level characteristics, school, test year and birth year fixed effects, along with local community-specific linear trends. Sample consists of students who are certified for free/reduced lunch or in poverty based on the description by the Texas Education Agency (TEA). Student data are from the TEA which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 and 1999. HS spending data are from the Consolidated Federal Files Reports (CFFR) and include years between 1988 and 1994. Standard errors are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.

Table A.5: Falsification Tests: The Effect of Head Start Exposure on Third Grade Standardized Test Scores For Students Not Certified for Free/Reduced Lunch

	All	Males	Females	Whites	Blacks	Hispanics
	(1)	(2)	(3)	(4)	(5)	(6)
Head Start Funding	0.001	-0.007	0.002	0.005	0.090	0.008
	(0.008)	(0.014)	(0.008)	(0.011)	(0.076)	(0.016)
Mean Y	0.310	0.254	0.368	0.391	-0.038	0.049
Obs	406517	205843	200674	315498	31391	49659

Notes: This table reports results obtained when the dependent variable is third grade combined standardized test scores in math and reading, and the independent variable is real federal HS spending per child (in 2014$) when the child was four years old. HS spending per child is scaled by the mean spending ($517), thus the coefficients should be interpreted as the effect of exposure to an average-sized program. All regressions include controls for demographics and county-level characteristics, school, test year and birth year fixed effects, along with local community-specific linear trends. Sample consists of students who are not certified for free/reduced lunch or in poverty based on the description by the Texas Education Agency (TEA). Student data are from the TEA which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 and 1999. HS spending data are from the Consolidated Federal Files Reports (CFFR) and include years between 1988 and 1994. Standard errors are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.

Table A.6: Falsification Tests: The Effect of Head Start Exposure on Exogenous Student Characteristics and County-level Income

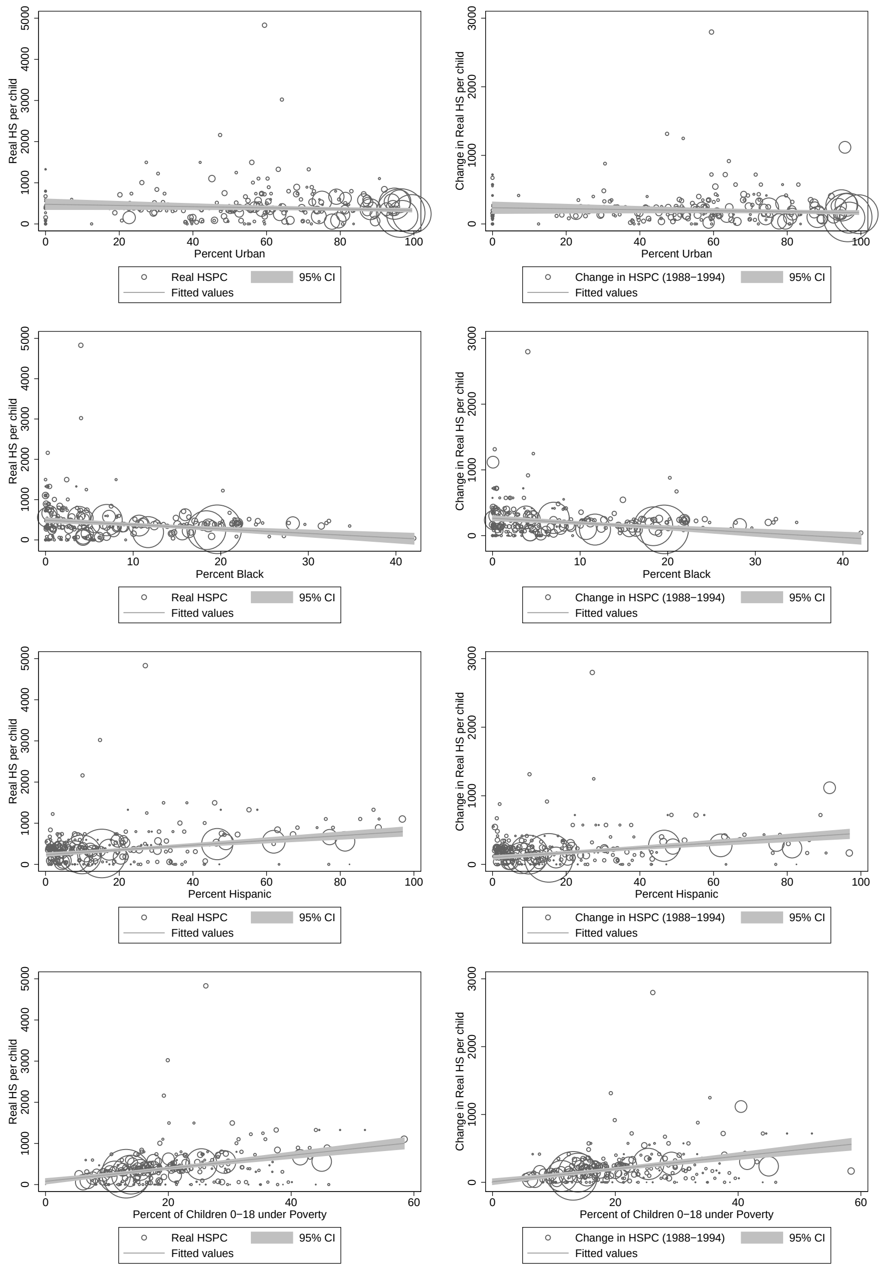
	Black	Hispanic	Female	Free/Reduced Meal	Income PC	Predicted Math Score
Head Start Funding	-0.001	0.002	0.000	0.001	0.000	0.002
	(0.001)	(0.002)	(0.002)	(0.004)	(0.001)	(0.001)
Mean Y	0.146	0.312	0.498	0.397	35613.425	0.994
Obs	762815	762815	762815	762815	762815	762815

Notes: This table reports results obtained when the dependent variable is third grade combined standardized test scores in math and reading, and the independent variable is real federal HS spending per child (in 2014$) when the child was four years old. HS spending per child is scaled by the mean spending ($517), thus the coefficients should be interpreted as the effect of exposure to an average-sized program. All regressions include controls for demographics and county-level characteristics, school, test year and birth year fixed effects, along with local community-specific linear trends. Sample consists of all the third grade students. Student data are from the TEA which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 and 1999. HS spending data are from the Consolidated Federal Files Reports (CFFR) and include years between 1988 and 1994. Standard errors are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.

B Online Appendix

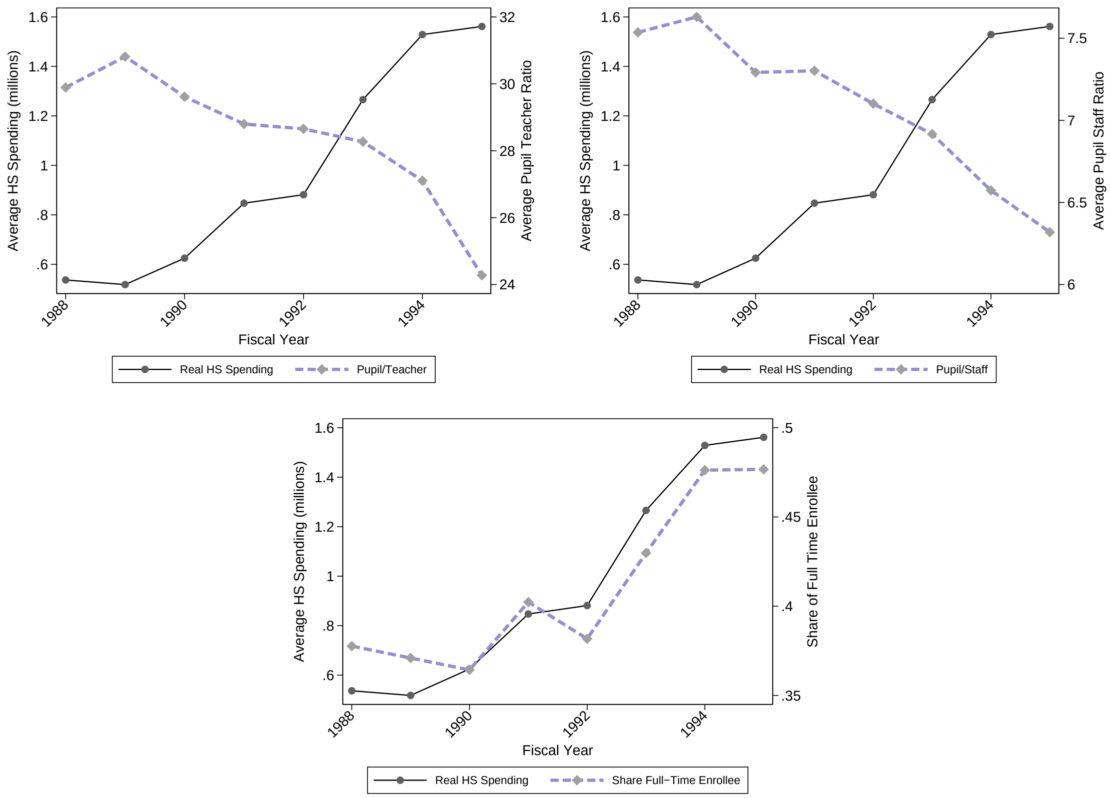
B.1 Figures

Figure OA.1: Correlations between 1980 County Characteristics and Head Start Funding per Child



Notes: Head Start spending (in 2014$) data are from Consolidated Federal Funds Reports (CFFR), coupled with the population counts for three- and four-year-olds at the county-level from the SEER. The 1980 county controls are from City and County Data Book. The bubbles present the counties, weighted by the population of three- and four-year-olds. On the left, the figures present the correlations between the county characteristics and the average HS spending per child for 1988-1994. On the right, the figures show the correlations between the county characteristics and the change in HS spending per child from 1988 to 1994.

Figure OA.2: Head Start Funding and Program Quality Trends



Notes: Federal Head Start spending (in 2014$ in millions) data are from the Consolidated Federal Funds Reports (CFFR). Head Start program data are from the Program Information Reports (PIR) for years 1988 to 1995.

B.2 Tables

Table OA.1: Head Start Funding and Director Quality, 1992-1994

	(1)	(2)	(3)
Director has BA+ degree	0.041		
	(0.481)		
Director's yrs of eduation		0.054	
		(0.113)	
Director's yrs of experience			0.005
			(0.044)
Mean Y (in millions)	4.639	4.639	4.639
Mean X	0.802	5.121	7.579
Adj. R-Squared	0.951	0.951	0.951
Obs	195	195	195

Notes: The data are at the local community-level and the dependent variable is real federal HS spending per child (2014$). Head Start spending data are from Consolidated Federal Funds Reports (CFFR). Head Start program directors' data are from the Program Information Reports (PIR), and include years between 1992 to 1994. For more details about the data description, see Section 3. The 1980 county controls are from City and County Data Book which include variables described in Section 3. All regressions include community, year fixed effects and community characteristics interacted with linear trends. Estimates are weighted with the number of children 3-4. Standard errors are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.

Table OA.2: Baseline Estimates of the Effect of Head Start Funding on Standardized Math and Reading Test Scores in Third Grade

	All	Males	Females	Whites	Blacks	Hispanics
	(1)	(2)	(3)	(4)	(5)	(6)
A: Standardized Math Scores						
Head Start Funding	0.054***	0.059*	0.050***	-0.017	0.053	0.082***
	(0.020)	(0.031)	(0.015)	(0.050)	(0.045)	(0.020)
Mean Y	-0.332	-0.397	-0.269	-0.049	-0.388	-0.412
B: Standardized Reading Scores						
Head Start Funding	0.028*	0.031	0.030	-0.020	0.032	0.045***
	(0.015)	(0.020)	(0.028)	(0.045)	(0.041)	(0.016)
Mean Y	-0.339	-0.471	-0.209	-0.085	-0.333	-0.434
Obs	332910	165303	167607	72073	75614	180775

Notes: This table reports results obtained when the dependent variable is third grade standardized test scores in math and reading separately, and the independent variable is real federal HS spending per child (in 2014$) when the child was four years old. HS spending per child is scaled by the mean spending ($517), thus the coefficients should be interpreted as the effect of exposure to an average-sized program. All regressions include controls for demographics and county-level characteristics, school, test year and birth year fixed effects, along with school-specific linear trends. Sample consists of students who are certified for free/reduced lunch or in poverty based on the description by the Texas Education Agency (TEA). Student data are from the TEA which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 and 1999. HS spending data are from the Consolidated Federal Files Reports (CFFR) and include years between 1988 and 1994. Standard errors are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.

Table OA.3: Baseline Estimates of the Effect of Head Start Funding on Standardized Test Scores - Adding School Trends

	All	Males	Females	Whites	Blacks	Hispanics
	(1)	(2)	(3)	(4)	(5)	(6)
Head Start Funding	0.024*	0.021	0.024	-0.052	-0.098	0.051***
	(0.014)	(0.025)	(0.021)	(0.065)	(0.090)	(0.014)
Mean Y	-0.356	-0.459	-0.253	-0.071	-0.382	-0.449
Obs	332910	165303	167607	72073	75614	180775

Notes: This table reports results obtained when the dependent variable is third grade combined standardized test scores in math and reading, and the independent variable is real federal HS spending per child (in 2014$) when the child was four years old. HS spending per child is scaled by the mean spending ($517), thus the coefficients should be interpreted as the effect of exposure to an average-sized program. All regressions include controls for demographics and county-level characteristics, school, test year and birth year fixed effects, along with school-specific linear trends. Sample consists of students who are certified for free/reduced lunch or in poverty based on the description by the Texas Education Agency (TEA). Student data are from the TEA which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 and 1999. HS spending data are from the Consolidated Federal Files Reports (CFFR) and include years between 1988 and 1994. Standard errors are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.

Table OA.4: The Effect of Alternative Head Start Funding Measures on Third Grade Standardized Test Scores

	All	Males	Females	Whites	Blacks	Hispanics
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A:						
Head Start Funding Per Capita	0.030***	0.039***	0.023	-0.031	0.052	0.045***
	(0.009)	(0.015)	(0.014)	(0.052)	(0.050)	(0.010)
Mean Y	-0.356	-0.459	-0.253	-0.071	-0.382	-0.449
Mean Funding Level ($)	18					
Panel B:						
Head Start Funding Per Poor Child	0.029**	0.034	0.027*	-0.002	0.033	0.053***
	(0.013)	(0.021)	(0.016)	(0.029)	(0.036)	(0.014)
Mean Y	-0.356	-0.459	-0.253	-0.071	-0.382	-0.449
Mean Funding Level($)	1198					
Obs	332910	165303	167607	72073	75614	180775

Notes: This table contains results obtained when the dependent variable is third grade combined standardized test scores in math and reading using the independent variables; real federal HS spending per capita (2014$) when the child was four years old in Panel A and the real federal HS spending per poor child (2014$) when the child was four years old in Panel B. Both measures are scaled by the mean spending, thus the coefficients should be interpreted as the effect of exposure to an average-sized program. All regressions include controls for demographics and county-level characteristics, school, test year and birth year fixed effects, along with community-specific linear trends. Sample consists of students who are certified for free/reduced lunch or in poverty based on the description by the Texas Education Agency (TEA). Student data are from the TEA which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 and 1999. HS spending data are from the Consolidated Federal Files Reports (CFFR) and include years between 1988 and 1994. Standard errors are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.