Local Area Spending Exposure to Head Start and Academic Performance: Evidence from Texas

Esra Kose*

Job Market Paper
December 12, 2016

The latest version can be found at: http://esrakose.weebly.com/research.

Abstract

Head Start (HS) is the largest preschool program in the U.S., yet the effectiveness of the program remains an open question. This paper leverages variation across local communities in the timing and intensity of federal HS spending expansions during the 1990s to estimate the effect of HS on academic achievement in the medium-run. Using student-level data from Texas, I find that exposure to more generous HS funding at age four significantly improves test scores in third grade through fifth grade for low-income children. My results show that HS benefited Hispanics with limited language proficiency the most and that a 500 dollar increase in HS funding per child closes about 15% of the raw test score gap between Hispanics and whites. An investigation of the mechanisms behind the test score gains suggests that federal funding expansions are associated with increases in HS enrollment and improvements in measured program inputs.

Keywords: Early-Childhood Education, Head Start, Program Evaluation

*I thank my advisors 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 available for use to approved researchers. All errors are my own. Esra Kose, Department of Economics, University of California, Davis, 1 Shields Ave., Davis, CA, 95616. Phone: (510) 613-5614. E-mail: ekose@ucdavis.edu.

1 Introduction

Head Start (HS) began in 1965, as part of President Lyndon B. Johnson's "War on Poverty" initiative to provide education, health, and social services to low-income children age three to five and their families. The goal of HS is to reduce education and health disparities between economically disadvantaged children and their more advantaged counterparts. As of 2016, HS serves over 900,000 children, with around $8.3 billion in federal spending (Human and Health Services, 2016). In the 1990s, Congress passed two important acts aimed at improving the quality and capacity of the program, which quadrupled federal funding for HS. The significant expansion within a relatively short time in the history of HS creates a natural experiment to study the effectiveness of the program, and a rare opportunity to learn about the use of federal HS funding in the local communities.

As the largest federal early childhood program, HS has been evaluated extensively over its existence. Most studies have indicated that HS has not fulfilled its mission of closing the achievement gap between low-income children and their more advantaged peers when children are in elementary school. Quasi-experimental studies of HS have mostly examined the first 15 years of the program and shown positive gains on test scores in the short-term that fade out quickly (Currie and Thomas, 1995). Similarly, results from the Head Start Impact Study (HSIS), the first large-scale randomized experiment of HS conducted in 2002, show positive test-score gains during preschool and first grade (Kline and Walters, 2016) that do not persist into third grade (Puma et al., 2012). The mixed findings from the best available empirical evidence have amplified skepticism of the program's effectiveness (Barnett, 2011).

This paper provides new evidence on the effect of HS on academic achievement by exploiting a new source of variation in program funding intensity across local communities over time. Specifically, I ask to what extent did federal HS funding expansions in the 1990s make a difference for student performance in Texas? How does the effectiveness of the HS program relate to the way HS funds are spent? To investigate these issues, I use unique student-level data on academic performance, and student backgrounds, together with grantee-level HS budgets during the period from 1988 to 1995.

This analysis differs from earlier studies of HS in several important aspects. First, it implements a more policy-relevant evaluation to assess whether HS has succeeded in narrowing the achievement gap using a large, demographically and socioeconomically diverse student population in Texas, the second most populous state in the U.S. 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 measurement error in the retrospective report of participation to HS and spillover effects across siblings, 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). I study HS in the context of federal grants by exploiting the differential timing in the program funding adoption in local communities following the substantial expansions in the 1990s. This approach provides a new and more applicable understanding of the program's effectiveness.

Second, this study improves upon previous research by exploring the mechanisms through which federal funding affects student performance. Combining several sources on HS enrollment, program characteristics and budgets, I show that improvements in program capacity and program inputs contributed to test score improvements. In particular, federal funding expansions led to program quality improvements including reductions in child-teacher ratios, and increases in full-time enrollment and education spending. This analysis suggests that federal dollars were spent to the targeted goals.

To examine the impact of HS on academic achievement, I employ several data sources. Student-level administrative data on student characteristics and achievement from 1994 to 1999 are provided by the Texas Education Agency (TEA). I construct a community-by-year dataset on HS spending per age-eligible child by combining three sources of data: (1) grantee-level HS spending data from the Consolidated Federal Funds Reports (CFFR); (2)

administrative HS program data by Currie and Neidell (2007), which describes the serving counties for each grantee;[1] and (3) county-level child counts from the Surveillance, Epidemiology, and End Results Program (SEER). Combining student-level data with community HS funding generosity, I estimate the effect of exposure to HS funding at age four on third grade standardized test scores for children born between 1984 and 1991. I focus my analysis on math scores because many studies find that relative to reading, math scores are stronger predictors of future outcomes (e.g., Duncan et al., 2007).[2]

My empirical strategy utilizes a panel fixed effects model, which leverages variation across local communities in the timing and intensity of HS spending expansions in the 1990s. Using this variation, I find that a 500 dollar increase in HS funding at age four increases average test scores in third grade in math by 0.04 standard deviations. Further analysis indicates that HS funding exposure improves test scores for both males and females, with larger point estimates for males but more precise estimates for females. Given that low-income male students are lower-achieving compared to females, these findings are consistent with previous studies documenting higher returns to public investments for groups in the lower end of the skill distribution (e.g., Bitler et al., 2014).

Estimates by race and ethnicity show that improvements among Hispanics are the main driver of the results, in particular for students with limited language proficiency. While the impacts for Hispanics are large and statistically significant, the analogous estimates for whites suggest that these cohorts did not experience significant improvements in test scores. On the other hand, black students benefited from the funding expansions but the results are imprecise. Overall, my findings suggest that a 500 dollar increase in HS funding exposure would close about 15% of the third-grade test score gap between Hispanics and whites. Additional results suggest that HS improves language proficiency for Hispanics which could potentially explain higher test score gains for this group. Other potential explanations in-

[1] Each grantee can serve one county or a group of counties.

[2] The results using reading scores follow the same pattern as the impacts on the math scores. I show the results for reading scores in the Appendix.

clude higher participation rates among Hispanics (especially in Texas) and increased cultural assimilation (Currie and Thomas, 1999; Bitler et al., 2014).

Fade-out of the test score gains in the HS literature has been in the heart of the policy debate. Critics have pointed out the null findings from the quasi-experimental evidence as well as the findings from the recent evaluation of the HSIS as evidence against the effectiveness of the program. To shed new light on the HS policy debate, I next explore whether the test score gains in third grade persist over time. My estimates provide suggestive evidence that the test score gains in math persist through fifth grade.

An investigation into the mechanisms behind the test score gains reveals that both program capacity and quality improvements are responsible. I show that funding expansions translated into large and significant increases in HS enrollment. Moreover, I find that federal funding improved child-teacher ratios, child-staff ratios, full-time enrollment, and education spending increases in the HS programs. These findings are consistent with findings from other interventions showing that reductions in student-teacher ratios benefit students' educational outcomes (e.g., Krueger and Whitmore, 2001; Jackson et al., 2015).

To compare my estimates to the rest of the HS literature, I convert the reduced form third-grade test score impact using the estimate of funding expansions on the program-level HS enrollment. This exercise implies a local average treatment effect equal to approximately 0.29 standard deviations of third-grade test score gains.[3] In contrast to previous studies that report smaller or no measurable test score impacts over the medium-run (Deming, 2009; Puma et al., 2012; Carneiro and Ginja, 2014), my results show a large and significant impact. This discrepancy may be explained by the fact that my analysis includes cohorts who would have attended HS in the 1990s, the period that HS experienced structural changes in terms of quality and capacity. Also, my sample consists of low-income children in Texas which includes a large concentration of Hispanics, different from other studies. Unlike these studies,

[3]This is likely an overestimate of the effect of HS attendance because increased funding may also boost scores for always takers. Also, this result is based on the strong assumption that HS attendance is the only channel through which funding has an effect on test scores.

I use federal funding expansions to identify the impact of HS on academic performance. Ultimately, the estimates using the HSIS (Puma et al., 2012) are within a 95% confidence interval of the estimates in my paper after adjusting for the fact that at least 40% of the control group attended another type of preschool.

I employ various analyses to probe the robustness of the estimates and show that my results are not driven by common shocks, selection bias, or other endogenous factors. First, I show that communities that experienced high growth in HS spending have similar pre-expansion trends in test scores compared to the communities that experienced low growth, which suggest that common shocks are not likely to drive the results. Second, to address the concern that HS may have been introduced or expanded with other county-level social programs that affected children's outcomes, I add detailed controls for county-by-year federal spending on a variety of social programs. I show that the results are robust adding these controls. Third, I control for the share of the age-eligible children enrolled in pre-kindergarten (pre-K) at age four and show that the results are not sensitive to inclusion of the control for pre-K. Moreover, I show that directors' education and experience are not associated with HS funding expansions. Next, I conduct placebo tests on groups not likely to participate in HS and find no evidence of an impact. In addition, I test whether HS funding generosity changes composition of students within school and grade and show that the results do not appear to be driven by endogenous composition effects. Lastly, I demonstrate that HS funding exposure at ages younger than three and older than five is not associated with improved test scores.

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

2 Background and Prior Literature

2.1 The Head Start Program

HS is a federally-funded early childhood education program that provides education, cognitive development, 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 be between ages three and five and live in 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) (HHS, 2016).[4] 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. Initially a summer-only program, HS soon grew to become a nine-month program that operated part-day. Then, in the 1990s, HS expanded substantially and shifted to more of a full-day program. Figure 1 shows that between 1990 and 2000, enrollment increased by about 60% and the federal funding per enrolled child doubled (HHS, 2015). The 1990s expansion was a result of a significant effort by the Bush and Clinton Administrations to improve upon 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 over-

[4]However, eligibility does not imply attendance. Priority is given to those most in need of HS services. Currie (2006) reports that about 65% of eligible children participated in HS in 2000. Kline and Walters (2016) report a 49% participation rate in the HSIS sample in 2002.

all 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 (the 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).[5] Grants are issued through a competitive process with priority given to agencies that are able to demonstrate the most cost-effective operation, but existing programs have priority when re-applying.[6] Grantees must provide at least 20% of the funding, which may include in-kind contributions, such as facilities to hold classes, through community partnerships.[7] Most HS grantees operate through community action agencies, local school systems, private/public non-profits, government agencies, and Indian Tribes.[8] 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).[9] As a result, there is a great deal of geographic variation in funding levels, which provides part of the identification in this paper.

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 (Currie and Neidell, 2007). Because part of the variation in funding stems from the qualifications of the grantees

[5]The Community Services Act of 1974 recognizes HS's transfer from the Office of Economic Opportunity to the Health and Human Services, extends the program's authority, and establishes a mandatory formula to allocate funds among states (HHS, 2015).

[6]The cycle period extended from three years to five years in 2007. With a preference on incumbents, many HS centers have been in the same location for several years (Currie and Neidell, 2007).

[7]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).

[8]During 1988-1995, in Texas the distribution of the programs was around 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).

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

(not due to exogenous policy changes), careful consideration is needed to isolate the exogenous variation necessary 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, HS has been evaluated extensively over its existence.[10] There are two sources of evidence on HS: (1) the Head Start Impact Study (HSIS), the large scale randomized experiment; and (2) studies based on nationally representative panel data such as the PSID and NLSY, which report participation in HS and have long-term follow-up data (Elango et al., 2015). The punchline of previous research is that HS is effective in improving test scores in the short-term, when children are in kindergarten; but that these test score gains fade out in elementary school. At the same time, quasi-experimental studies of HS show that children enrolled in HS have better academic, labor market, and health outcomes in adulthood.

The best available evidence on the short-term impacts of HS on children comes from the HSIS, which took place in 2002.[11] The results of this experiment suggest that short-term average math score gains during HS attendance are around 0.15 standard deviations (Ludwig and Phillips, 2008).[12] However, Puma et al. (2012) show that short-term positive gains for most measures of cognition disappear by age 9 (third grade).[13] It is important to note that in the HSIS 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) should be interpreted with

[10] A number of studies has 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).

[11] This experiment was carried out by Westat for the U.S. Department of Health and Human Services. The longest follow-up for this study is on the outcomes at age 9, in third grade.

[12] The effect sizes range from 0.182 and 0.147 standard deviations for 3 and 4 year olds, respectively (treatment effect on the treated). Early math test is measured by the Woodcock-Johnson applied problems test.

[13] They find some evidence of impacts on the non-cognitive skills that persist through third grade.

caution. Nevertheless, 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 that 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 and find larger short-term impacts at low quantiles of the test score distribution and persistent effects 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 timing, 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. Similar to challenges faced when evaluating the effectiveness of other federal programs, credibly estimating the causal effects of HS has been difficult due to data and methodological limitations. As a practical issue, most big datasets do not allow researchers to identify the participants of HS. As a result, previous research has relied on datasets that report HS participation but have small sample sizes. The data limitation problem leads to statistical inference issue. In addition, participating children are not randomly assigned to the program which leads to the problem of selection bias. 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).[14] Admittedly, this approach suffers from measurement error in the retrospective report in participation to HS and spillover effects across siblings, which may bias estimates toward zero. Other papers that exploit discontinuities due to program funding and eligibility rules significantly

[14] 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 that persist through elementary school, then the gains fade out. Deming (2009) 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).

improve upon earlier studies but still suffer from limited sample sizes (Ludwig and Miller, 2007; Carneiro and Ginja, 2014). Unlike these papers, I use a different source of variation, federal funding expansions to identify the impact of HS on academic performance. This is a more policy-relevant source relative to the effect of participation. Also, I utilize student-level administrative data in my study which is advantageous to overcome statistical inference issues.

Similar to my paper, a few papers exploit variation in HS program expansions to analyze the long-term effects of HS (Frisvold, 2006; Johnson, 2010; Thompson, 2016).[15] For example, Frisvold (2006) uses the relative availability of HS in a county when a child was four years old as a measure of the availability and intensity of the program as an instrument to identify the effects of HS on obesity in adolescence.[16] Thompson (2016) examines the long-term socio-economic well-being of participants by exploiting differences in HS funding levels across counties during the 1966-1968 period of the program's introduction. Unlike these papers, I exploit variation from a much more recent period and use the 1990s funding to analyze the impact on test scores.

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 effects of HS spending exposure at age four on student test scores.

Administrative student-level data include the universe of exam takers in Texas from

[15] Frisvold (2006) and Johnson (2010) use the PSID; and Thompson (2016) explore the NLSY as their main data set.

[16] He defines the relative availability of Head Start as the number of children who attend Head Start as a proportion of the number of eligible children in the local community. Also, Johnson (2010) proposes to use variation in county spending for the first 15 years of the program to analyze the long-term outcomes. He finds suggestive evidence of positive health and education improvements in the long-term.

third to eighth 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).[17] 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).[18] 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 proficiency and special education status on each student at each school district.[19] From school district information, I determine the county of residence, which serves as a proxy for the child's county of residence at age four and jointly with year of birth determines each student's exposure to HS funding.[20]

I conduct the following sample restrictions.[21] First, for my core analysis, I keep only third grade students because data on HS program characteristics are available for this group (1,000,524 observations remain). Second, I drop observations with missing demographic information, those with 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 eligible for free or reduced lunch or who are identified as economically disadvantaged based on their families' welfare eligibility because

[17] The TAAS is administered for years 1994-2002, 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.

[18] This is accomplished through a rating system based on annual student test scores. The fraction of students who achieve passing scores on standardized exams determines a school's rating system which then determines the funding stream that a school acquires from the state government. If schools do not meet the required standards, they face increased state monitoring and loss of operation rights. Therefore, school administrators and teachers have strong incentives to maximize their school's fraction passing (Richardson, 2010).

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

[20] 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.

[21] Raw data includes 7,171,396 records for students from third to eighth grade.

they are more likely to be eligible for HS (332,910 observations remain).[22] From here, I will refer to this sample as free or reduced lunch eligible.

I develop a measure of HS funding per age-eligible child in a local community using a number of sources.[23] HS spending data are from the Consolidated Federal Funds Reports (CFFR), which include information on local appropriations for federally funded programs for each grantee every year starting in 1983 (for more details, see Section A.1).[24,25] I restrict to the years of HS spending between 1988 and 1995, when the analysis sample would be age-eligible to attend HS. Similar to other programs that are financed by federal grants, a HS grantee could oversee one county or a group of counties.[26] When the federal government announces grant availability for a specific local community, it also announces which counties each grantee is expected to serve. However, these 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.[27] I then confirmed these networks of counties using the website of the grantees

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

[23]Ideally, the most accurate measure of the size of local HS program is HS spending per age- and income-eligible child. Due to data limitations in poor child counts at the local level, my main analysis uses HS spending per age-eligible child. However, I also perform analyses using two different measures: (1) HS spending per capita; and (2) HS spending per poor child.

[24]Federal government expenditures data which are reported in state and county 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.

[25]Source: https://www.census.gov/prod/2011pubs/cffr-10.pdf

[26]HS is administered by the Department of Health and Human Services (HHS), Administration for Children and Families (ACF), Office of Head Start (OHS). HHS describes "grantees" as the agencies that receive grant awards directly. "delegates" are other agencies that grantees may contract services.

[27]These data cover the years 1990-2000. I assigned the networks based on 1990 networks for years earlier based on the assumption that the networks did not changed drastically from 1988-1990. 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 1995. 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. However, the magnitude of the measurement error is expected to be small.

and an additional data set provided by Frisvold (2006).[28,29] To construct the population denominator of children ages three and four years old, I use data from the Surveillance, Epidemiology, and End Results Program (SEER) which include county-level population counts for each age group. Together with these two data sources, I construct the "HS funding per child" variable at the local community-year level.[30,31]

I augment these data with additional data sources to bring in information on HS enrollment and to control for county-year economic conditions. First, I compiled data from the Program Information Reports (PIR).[32] 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 are important for my analysis as they provide information on number of funded enrollees, 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 pre-K enrollment from the Common Core of Data (CCD).[33,34] Next, I compiled data on county-level economic conditions. I use the Regional Economic Information Systems (REIS) to construct a county-year data on per capita income, per capita transfers payments for cash income support (AFDC and SSI), medical care (Medicare), food assistance (Food Stamps), retirement and disability programs. Other county demographics including the 1980 population living in an urban area, black, Hispanic, single parent, less than age 5, age 65 or older, percent of the 0-18 year-olds living in poverty,

[28] 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.

[29] Appendix Figure A.1 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.

[30] I discuss the construction of this variable in further detail in Section A.1.

[31] Grantees from urban counties generally serve only one county.

[32] Source: https://eclkc.ohs.acf.hhs.gov/hslc/data/pir

[33] Source: https://nces.ed.gov/ccd/pubschuniv.asp

[34] CCD includes school-level information for all public schools. These data are available starting in 1986 at the school level and provides information on pupil teacher ratio, a measure used for education quality in education literature, as well as the demographic composition of students and the level of grades offered in a specific school.

as well as income, education, welfare spending per capita (in 2014$) are constructed using the 1980 City and County Data Book (before HS spending roll-out period). 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 the SEER.

3.2 Summary Statistics

My analysis centers on children exposed to HS funding in Texas between 1988 and 1995.[35] Figure 2 presents the average HS spending per child in the 15 most populous counties in Texas from 1980 to 2000.[36] The blue horizontal lines in Figure 2 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 3 and Figure 4 are maps of: (i) the levels of funding in 1988 and (ii) growth in the HS spending per child from 1988 to 1995, 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.

Figure 5 shows the distribution of student test scores in math and reading by free or reduced lunch eligibility. The solid line shows the passing score of 70 that is determined by the TEA. This figure shows that free or reduced lunch eligible students were academically behind compared to the rest of the sample, with a lower passing rate.

Table 1 presents summary statistics for third grade students, first for the full sample and economically advantaged students, then for students who are eligible for free or reduced

[35] My analysis starts in 1988 because information on HS program characteristics is not available from the Program Information Reports (PIRs) prior to 1988. Without such data, it is not possible to quantify the effect of funding on the provision of the HS program and substitution across different types of early childhood education activities among four-year-olds.

[36] 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 around 60% of the student population in Texas.

lunch. The first three columns show a significant discrepancy in average test scores by economic status. Relative to free or reduced lunch eligible students, economically advantaged ones have substantially higher test scores. Not surprisingly, the free or reduced lunch eligible sample have higher exposure to HS. Blacks and Hispanics make up more than 75% of students who are eligible for free or reduced lunch, and within this sample, there is significant variation in test scores across subgroups. For example, Hispanics tend to live in counties with higher funding for HS per child but have the lowest test scores. On the other hand, blacks have the lowest HS funding exposure, possibly because they tend to live in more densely populated areas.

To demonstrate the relationship between average test scores and HS funding, I restrict the sample to children eligible for free or reduced lunch and collapse the data on test scores to county-level. Figure 6 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 eligible children.

4 Identification Strategy and Methodology

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 the county of residence and year of birth.[37] 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 eligible students

[37] 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 a bias in my estimates. I test this directly and find no evidence that the HS spending affects the composition of students within school and grade over time.

using a newly assembled dataset on community-level HS spending per child:

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

where Y_{isgbt} denotes the outcome variable (standardized test scores) for student i in school s in community g in birth year b and in test year t. $HSfunding_{g\tau}$ represents HS funding per child in local community g when student i was four years old ($\tau = b+4$). X_{isgbt} is a vector of individual-level demographic controls including gender, ethnicity, an indicator for bilingual, English as a second language, gifted/talented, and special education status. Z_{ct} has three sets of county-level controls including county-level per capita income transfers, controls from the main data collapsed to the county-level, and county-level controls including fraction of 0-18 living in poverty, percent urban, share of 0-5 year-olds by ethnicity, log population, and unemployment rate. $W_{c\tau}$ includes county-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 $\pi_g * b$ is a community-specific linear trend. Standard errors are clustered at the local community-level. 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 a valid approach to examine the effect of HS on test scores, 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 two key identifying assumptions.

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 1995 (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, county population, percent of the 0-18 year-olds living in poverty and income, education, and welfare spending per capita (in 2014$).[38] The results are presented in Table 2 and Table 3 (see Appendix Figure OA.2 for visual presentation of correlations). Simple correlations imply that communities that have a larger Hispanic population and a higher share of single mother, poor, very young, or elderly have more HS funding. In contrast, communities with larger black population tend to have less funding for HS, which could be explained by the fact that blacks in Texas live in more urban areas with high population density. Consistent with the political history, 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 1995 are also similar (see Table 3). After controlling for characteristics that are described above, last columns of Table 2 and Table 3 show that only income per capita and percent children living under poverty are significant determinants of HS expansions and together all variables explain a small fraction of the variation overall (R-squared of around 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 and obtain additional funds, and these directors also operate higher quality programs. Then funding levels could be correlated with child outcomes (Frisvold, 2006; Currie and Neidell, 2007). To rule out this possibility, I examine the effects of directors' education, experience and salary on test scores in Appendix Table A.2 and show that directors' qualifications are not predictive of test scores.[39] A bias

[38] These data are constructed using the 1980 County and City Data Book: `http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/8256`.

[39] 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

could also occur if some communities devote local resources towards 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.

Third, 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 policy changes that targeted four year olds, I directly control for county-level spending for other social programs.[40] Moreover, inclusion of local community-specific linear time trends also accounts for the fact that some communities may be improving over time.

In addition, I show that communities that experienced high growth in HS spending have similar trends in test scores compared to the communities that experienced low growth. Figure 7 presents the evolution of the average test scores (ranging between 0-100 with a passing score of 70) across the low and high growth communities (defined as low/high relative to the median growth in spending per child with the base level of 1983) over the year of exposure to HS. The dashed line shows the evolution of the difference in average test scores between high and low growth communities, which allows for a comparison of the relative preexisting trend. If anything, there is a decreasing trend in test scores from 1983 to 1984 in the high growth communities relative to the low growth ones.

Finally, important to my identification strategy is that HS spending expansions should not systematically change the composition of a particular grade 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

and 1995. I also show in Online Appendix Table OA.3 that federal funding increases are not significantly associated with directors' qualifications.

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

nonrandom selection would misattribute higher student performance to exposure to HS. 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-grade fixed effects. I find that variation in HS funding does not predict changes in the composition within school-grade, suggesting that the estimates are not biased by self-selection of families to particular cohorts within a school.

It is important to note that I do not control for language proficiency in my analysis because of the potential concern with "bad controls" (Angrist and Pischke, 2008). Since one of the main goals of the HS program is to promote language development and literacy, HS could directly improve language proficiency for immigrant children who are learning English as a second language. I formally test whether HS spending exposure for Hispanics reduces the likelihood of having limited language proficiency in Section 5.2 and find that HS significantly improves language proficiency. Moreover, students with difficulty in language tend to be underachieving, therefore, not controlling for limited language proficiency leads to overestimation of the main effect of HS on test scores. Appendix Table A.1 shows that the main results including limited language proficiency controls indeed reveal smaller point estimates, but the bias is negligible.

An important caveat in this analysis is that childhood and family background are not observed in the main data. Previous research has shown that low birth weight, having a mother who is an immigrant, minority, low-educated, or single are predictive of lower test scores in third grade and these characteristics are correlated with the generosity of social spending at the county-level (Ladd et al., 2014). Failing to control for these characteristics would bias the estimates reported here towards zero. As a result, the estimates that are reported here should be interpreted as a lower-bound impact of exposure to HS funding.

4.1 Early Childhood Education Alternatives in Texas

A potential threat to my identification strategy is the existence of other preschool alternatives to HS because 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, which provides early childhood education for four-year-old children to improve the academic performance of at risk children (TASB, 2010).[41,42] Although it is mandatory for any district that has at least 15 eligible children to offer half-day education-based program for four-year-old children, attendance is voluntary.[43,44] 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).[45,46] This raises the possibility that HS and pre-K might have operated jointly to some degree during the 1990s.

Figure 8 plots the number and the share of children enrolled in HS and pre-K in Texas between 1988 and 1995. Relative to HS, pre-K was larger with the share of age-eligible enrollment of 14% compared to 6% in HS in 1988. This figure also shows that the timing of the expansions in both programs coincide, which is concerning since the existence of a large-scale pre-K is a potential threat to identification as a confounding factor. It is important to control for the availability of other preschool alternatives because failing to control for it could misattribute the effects of other preschool to HS.

[41] https://www.tasb.org/Legislative/Issue-Based-Resources/documents/prek2010.aspx

[42] At risk population include children unable to speak and comprehend the English language; children eligible 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).

[43] Texas Education Code 29.1532

[44] 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.

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

[46] State funds eligible children through the Foundation School Program based on Average Daily Attendance (TEA, 2014).

To address this concern, I implement two additional analyses. First, I directly analyze the relationship between HS spending and pre-K expansions between 1988 and 1995. Although at the aggregate state level pre-K expansions coincide with the timing of HS expansions, Table 4 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 reassures that variation in HS per child does not predict pre-K enrollment, there is still a possibility that pre-K and HS cooperate in terms of facility usage and program operations in the 1990s.[47] To take this possibility into account, I control for the share of age-eligible 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

Since the education literature emphasizes test scores in math as a better predictor of the future outcomes compared to test scores in reading, I show all my results for math scores (Duncan et al., 2007). The results using reading scores are smaller and estimated with less precision, however, the direction of the estimates follow the same pattern as the impacts on the math scores.

5.1 Baseline Results: Federal Head Start Funding and Test Scores

Having documented the plausible exogeneity of the HS funding variation, I now present estimates of Equation 1, the effect of HS spending per child on third grade standardized math scores.[48] As noted above, the main sample is restricted to free or reduced lunch

[47]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.

[48]The results for the third grade reading test scores are presented in Appendix Table A.3 and Table A.4.

eligible students and HS exposure is assigned at the time when a child was four years old. To simplify the interpretation of the coefficient of interest, HS spending per child is divided by 1000, thus the coefficient should be interpreted as the effect of a 1000 dollar increase. As a point of reference, during the period of the study from 1988 to 1995 in Texas, average funding per child increased by about 500 dollars (from 230$/child in 1988 to 736$/child in 1995).

Column (1) of Table 5 reports the results for the main sample of interest. The estimated coefficient indicates that being exposed to a 1000 dollar additional federal HS spending per child at age four leads to a statistically significant 0.081 standard deviation increase in test scores. This implies that a 500 dollar increase in federal funding improves test scores by 0.04 standard deviations. Here, the estimated coefficient represent the reduced form effects, which can be interpreted as the average effect of HS funding exposure on economically disadvantaged children.

In the next two columns of Table 5, I analyze whether exposure to HS funding improves third grade test scores differentially by gender. 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, HS may yield larger returns for males compared to females. Consistent with this prediction, in Columns (2) and (3), I find that Head Start exposure is associated with improvements in test scores for both males and females, with larger point estimates for males but more precise estimates for females. For males, I find that a 1000 dollar increase in HS funding improves third-grade test scores by 0.087 standard deviations while for females the effect is 0.066 standard deviations.[49]

In the rest of the table, I present the estimates for non-Hispanic whites, blacks and Hispanics, respectively. While the results for Hispanics are positive, large, and statistically significant, the analogous estimates for whites suggest that these cohorts did not experience

[49]The estimates are not statistically different from one another.

significant improvements in test scores with exposure to additional HS spending. The impact of HS on blacks on the other hand is large but not statistically significant. Considering the historical presence of blacks in HS and positive findings for blacks in the previous literature (e.g., Deming, 2009), it is expected that HS would improve their academic achievement. One possible explanation for the statistical imprecision in the estimates for whites and blacks could be that the variation in funding per child is limited (average funding exposure for HS for blacks is \$306 with a standard deviation of 227, whereas for Hispanics average funding is \$683 with a standard deviation of 1137). The reason for this could be that communities with high fraction of blacks experienced less funding expansion relative to the ones with high fraction of Hispanics, as shown in Table 3.

The race and ethnicity breakdown reveals that improvements for Hispanics are the main driver of the results. In particular, I find that a 500 dollar increase in federal funding for HS leads to a 0.057 standard deviation increase in test scores in math. 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 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 in math (= 0.057/0.352).[50]

5.2 Treatment Effect Heterogeneity

This section explores treatment effect heterogeneity within racial and ethnic groups, and provides possible explanations for these results. Table 6 presents differential effects by gender within racial and ethnic groups. In the first two columns, I find that smaller average effects on whites are driven by the negative (but insignificant) estimates on males. For blacks and whites, I find similar pattern in the direction of the estimates, positive impact on females relative to males. The results for whites are smaller and statistically indistinguishable from zero. On the other hand, exposure to HS improves test scores significantly for black females.

[50] Using Table 5, the average standardized test score for Hispanics is -0.41 and -0.058 for whites. The raw difference is 0.352.

For Hispanics, I find larger effects on males compared to females, with twice as big of an effect size for males compared to females (0.143 compared to 0.075).

As shown in Table 1, Hispanics lag behind both blacks and whites in terms of third grade test scores. Different from black children who are historically under-privileged compared to whites, Hispanics often live in immigrant, Spanish-speaking families and communities (Currie and Thomas, 1999). In Texas during my study period, Hispanics make more than 50% of free or reduced lunch eligible third grade students and 36% of Hispanics have limited language proficiency. To test whether HS differentially affects Hispanics with limited language proficiency, I split the sample of Hispanics by language proficiency and estimate the impact of funding exposure by gender within this group. Table 7 shows that indeed Hispanics with limited language proficiency benefit significantly from the program compared to the language proficient ones. Within this group, the effects are larger for males compared to females.

There are at least three channels by which HS could be beneficial for Hispanics according to previous literature. 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. I test this empirically in Table 8 and find that HS spending exposure increases the likelihood of becoming language proficient. These results are larger for males compared to females, which could also explain higher test score gains for males. Second, HS could promote cultural assimilation which in turn helps children adapt to school more easily. Third, compared to whites and blacks, Hispanics have higher participation in HS in Texas. For example, in 2000, 64% of HS participants were Hispanic.

To sum, this section provides evidence that the main effects on test scores are mostly driven by improvements among Hispanic males with limited language proficiency. This finding is consistent with the evidence from the HSIS documenting that HS is more effective for dual language learners and children with low levels of baseline test scores (Puma et al., 2010; Bitler et al., 2014).

5.3 Test Score Fade Out

Test score fade out has been at the heart of the HS policy debate. As evidence against the effectiveness of HS, critics of the program point out that test score gains at the time of school entry dissipate quickly. In this section, I provide suggestive evidence on the effectiveness of the HS program on test scores from third through eighth grade. Unfortunately, I do not have access to panel data that follows the same students across the test years (my sample is not balanced across cohort-test year). I observe different cohorts in each grade, which limits this analysis. As a result, I am not able to distinguish between the non-linearity of the effect in test scores over the years (the effect in third grade is expected to decrease as students age) and the reduction in exposure to generous funding across the years (older cohorts experience less exposure to HS relative to younger cohorts).

With these caveats in mind, Table 9 presents estimated coefficients from separate regressions with the standardized math scores in each grade as the outcome variable and exposure to HS funding per child at age four as the right-hand-side variable. Specifically, each column includes a sample of students from each grade who took the test in years between 1994 and 1999.[51] Overall, the results suggest that test score gains persist through grade five (Panel (A) of Table 9).

Panel (A) reports the estimates for free or reduced lunch sample for each grade and suggests some evidence of test score fade out. The effect is 0.081 standard deviations in third grade (for ages 8-9), goes down to 0.062 standard deviations in grade four, decreases to 0.058 standard deviations in grade five, and disappears beyond fifth grade. Although not conclusive, this fade out pattern is consistent with the results of previous literature (Currie and Thomas, 1995; Krueger and Whitmore, 2001; Deming, 2009).

[51] It is challenging to repeat this analysis in a balanced sample because the main data set provides each test taker from the third grade to eighth grade for a fix number of years. My best attempt is to restrict the sample to cohorts born in 1984-1988 whom appear in all the exam years. In Appendix Table A.6 I present the estimates for this sample. These findings suggest that most of the effect signs are similar to the main estimates with smaller magnitudes and lack of statistical power. This pattern is expected considering that there is less identifying variation with only five years of exposure.

Panels (B)-(F) of Table 9 contain estimates by gender and race/ethnicity. Panel (B) and Panel (C) show the results for females and males, respectively. The effect on third grade test scores is stronger for females compared with males, however, the effects fade out quickly for females. By comparison, test score gains for males persist strongly until the fifth grade. Deming (2009) finds similar results that show that the effect on test score is more persistent for male students compared to females from age five through 14.

Panels (D) and (E) report the results separately for whites and blacks, statistically imprecise for each grade. Unlike whites and blacks, the results for Hispanics in Panel (F) suggest that test scores gains decrease as students reach fifth grade. The estimated effect decreases from 0.114 standard deviations in third grade to 0.061 standard deviations in fifth grade. However, the point estimates are still sizable but marginally significant in fifth grade.

Overall, this section provides suggestive evidence that test score effects fade out over time. Figure 9 summarizes my findings, which suggest that the effects on test scores appear to be positive and significant through fifth grade, then the gains disappear. Importantly, the findings beyond fifth grade should be cautiously interpreted for two reasons. First, HS exposure decreases as the sample changes from third graders to eighth graders. Mechanically, older cohorts experience less exposure to HS and the identifying variation significantly shrinks for the sample. Second, one would expect that the gains to decrease over time but it is hard to guess by how much the impacts would shrink by eighth grade.

5.4 Discussion of Mechanisms and Magnitude of Estimates

One channel by which HS funding increases could improve test scores is by serving more children. A second potential channel is through improvements in existing program characteristics that could then lead to better academic outcomes. In this section, I explore the extent to which HS expansions translate into capacity growth and improvements in the quality of the program.

5.4.1 Federal Funding and Head Start Enrollment

With the data compiled from the PIRs, in Table 10, I examine the relationship between HS spending per child and aggregate HS enrollment as a share of age-eligible (Column (1)), and age- and income-eligible (Column (2)) children from 1988 to 1995 (unit of observation is community-year). This table shows that after controlling for community and year fixed effects, a 1000 dollar increase in funding per child is associated with a 0.102 percentage point increase in enrollment in HS for all children and a 0.277 percentage point increase in HS enrollment for poor children. These results suggest that a 500 dollar increase in HS funding per child is associated with around a 67% increase in enrollment.[52] These findings suggest that enrollment response to the federal funding increase is substantial.

5.4.2 Federal Funding and Program Inputs

I next examine the effects of federal funding increases on child-teacher ratios, child-staff ratios, share of full-time enrollment, and director's salary.[53] I also estimate effects of federal funding increases on spending for education, health, nutrition, and social services. I employ data on child-teacher ratios and full-time enrollment from the PIRs (available for 1988-1995), director's salary from the PIRs (available for 1992-1995), and program budgets for various types of spending available in the PCCOST data (available for 1993-1995 for some programs).

Prior research has shown that reductions in student-teacher ratios benefit students, particularly children from disadvantaged backgrounds (e.g., Krueger and Whitmore, 2001). However, in the HS literature, Walters (2015) studies the impact of different inputs in HS centers on cognitive and non-cognitive skills using the HSIS and finds that teacher education, teacher certification and class size are not associated with improvements in test scores. In Table 11, I examine the relationship between HS funding and various of program inputs. The

[52]A 500 dollar increase in HS funding is associated with a 0.051 percentage point increase in enrollment $= 0.102 * 500/1000$, which leads to a 67% increase in enrollment at the mean level of enrollment for all children and a 69% increase in enrollment for poor children.

[53]Online Appendix Table OA.1 shows trends for average of some of these inputs from 1988 to 1995.

first two columns of Table 11 show that federal funding increases are associated with significant reductions in child-teacher and child-staff ratios. These results suggest that reductions in child-teacher ratios are partially responsible for test score gains in the medium-run at the period of my study.

Next, Walters (2015) shows that the key input that improves children's cognitive skills is the provision of full-time services at the center level. Consistent with his finding, in Column (3) of Table 11, I show that there is a significant relationship between funding expansions and the share of full-time enrollment at the local community-level.

I find no evidence that funding increases are significantly associated with directors' salary increases. Improvements in test scores could also occur if federal funding is used to improve teachers' salary or training, nutrition or health services which would benefit children's education or health development. In the last four columns, I show that when the federal government doubles HS funding, spending on education increases by around 259 thousand, spending on health services increases by 43 thousand, spending on nutrition increases by 4 thousand, and spending on social services increases by 59 thousand dollars. 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.[54]

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

5.4.3 Magnitude of Estimates

To make accurate comparisons with the literature, I attempt to convert the reduced form impact of test scores in third grade to treatment effect on the treated. 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. However, HS participation is not observed in the main dataset. Nevertheless, I use the estimate from Column (2) of Table 10 that reports a 0.277 increase in the share of poor children enrolled in HS for a 1000 dollar increase in funding at the community level during the period of study. Using the main test score impact of 0.081 from Table 5 and scaling up by 0.277, the implied effect of HS enrollment at the local community level corresponds to 0.292 standard deviations (= 0.081/0.277). This provides a comparison point with the literature; however, one should be cautious about interpreting these estimates as the "true effect" of participating in HS. This is important because HS funding increases also affect quality of the program, which suggests that 0.292 could be interpreted as an upper bound of the effect of HS participation.

Figure 10 compares the test score impact of HS by age from this paper with the impact from Deming (2009), the impact for the four year old cohort from the HSIS (Puma et al., 2012), and the impact estimated by Carneiro and Ginja (2014).[55] Overall, the current study reports a larger impact on third grade test scores compared to the estimated effects reported in Deming (2009), who finds 0.133 standard deviations gain in test scores in elementary school for children who participated in HS compared to their siblings who did not. Unlike my paper, Deming (2009) studies the HS program for children who attended in the 1980s (pre-1990). This is important since HS evolved significantly in the 1990s in terms of quality and capacity.

Furthermore, my findings contrast with two other recent studies. The first study is the follow-up of the HSIS that reports zero effects for third grade test scores by comparing mean

[55] Similar to Gibbs et al. (2013), I use the four year old cohort as the comparison group because they only get one year treatment which simplifies the interpretation (the 3 year old cohort had a higher likelihood of getting additional year of HS).

differences between treatment and control groups (Puma et al., 2012).[56] 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. Scaling up the reported estimates, a test-score gain of 0.2 standard deviations would be included in the 95% confidence interval. Second, Carneiro and Ginja (2014) show significant health and behavioral effects of HS for males but find no evidence of the impact of HS participation on cognitive test scores for males in ages 12-13 who attended HS from 1986 to 2000. A potential explanation for the discrepancy would be based on the nature of the identification strategy utilized in Carneiro and Ginja (2014), which exploits the eligibility cutoffs for HS using the regression discontinuity framework.[57] Their sample size is limited to 1,197 males which could also explain statistical imprecision in their estimate. However, I should note that the average effect in my study is within the 95% confidence interval of the estimates from Carneiro and Ginja (2014).

6 Robustness

In this section, I conduct a variety of robustness exercises to address potential concerns with the estimates. Table 12 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 income per capita at the county 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 next column, I omit controls for the county-level measures of per-capita transfer payments for cash income support, Food Stamps,

[56] As noted above, the longest follow-up using the HSIS is the impacts on the third-graders so there is only one comparison point.

[57] They find the eligibility cut-off predicts HS participation only for males and show main test-score effects for males only. They estimate a local average treatment effects for income eligible children around the eligibility cutoffs.

medical care, retirement, and disability programs and find that the magnitude of the estimate goes down without these controls. In the last column, I included 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 estimate the main specification by gender and race/ethnicity adding school trends. Appendix Table A.5 presents these estimates which show that the effects for females and Hispanics are still statistically significant but slightly smaller.

I next analyze whether the results are sensitive to different measures of HS exposure.[58] In Column (1) of Table 13, I show the results obtained with my preferred measure, and in the next two columns I present results obtained with two alternative measures. 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 because 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.[59] Due to limited data availability, this measure is subject to measurement error. In Column (3) I present the results using HS per poor child,

[58]The construction of these measures is explained in detail in Data Appendix A.1.

[59]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, taking the difference between these two estimates, I construct the population 0-5 living in poverty. The number of poor children age three or four are the 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.

and I find smaller and less precise estimates compared to Column (1).[60]

Next, I analyze the effect of exposure to HS on test scores for the sample of children who are *not* identified as economically disadvantaged, who are not likely to benefit from HS. I present these results in Table 14 and show that HS exposure does not affect test scores for this group.

Additionally, I test whether increases in federal HS spending changes 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. Table 15 presents regression results in which I regress exogenous student characteristics on HS exposure conditional on school-grade 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 eligible for free or reduced lunch and income composition of the county) is correlated with exposure to HS conditional on school-grade 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 outcome measures (Thompson, 2016). As a falsification test, I estimate models where I analyze the exposure of HS at different ages. Table 16 reports results from specifications that use the sample of third graders with the assignment of exposure to HS changing from age zero through age eight as the right-hand-side variable. These findings suggest that HS exposure at ages three and five 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%

[60]In Online Appendix Tables OA.1 and OA.2, I show the results by gender and ethnicity using HS per capita and HS per poor child measures. The results using per capita measure are stronger for males and females as well as Hispanics. The per poor child measure reveals smaller point estimates with less statistical precision.

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 other program expansions 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.

7 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 1000 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.[61] 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 $1,000 per child since this paper estimates the test-score impact of a 1000 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 are free/reduced lunch eligible 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.[62]

[61] 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.

[62] In Appendix Table A.IV., Kline and Walters (2016) list several studies that estimate test scores and

Following Kline and Walters (2016), I use a conservative estimate of 10 percent of earnings impact per standard deviation of 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 $566,720 (in 2014$).[63] The average present discounted value of the predicted earnings at age 4 corresponds to around $434,000 with a discount rate of 3%. Adjusting for the fact that the median earnings in Texas is about 94 percent of the median earnings in the U.S., it corresponds to $407,960.[64] Children who participate in free/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.[65] Using the estimate for intergenerational income elasticity reported by Lee and Solon (2009) of 0.4, the average child in free/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 0.08 yields roughly $2,440 of projected earnings impact.

These calculations show that a conservative estimate of the benefit-cost ratio is roughly 2.4, which is in the range of the estimates for one year of HS attendance reported in Kline and Walters (2016). This estimate of 2.4 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).

earnings impacts.

[63]The reported value is $522,000 in 2010 dollars.

[64]Using the Current Population Survey Annual Social and Economic Supplement (CPS ASEC), between 1988 and 1995, the median earnings in the U.S. was $25,310 (in 2014$), while in Texas it was about $23,814 (in 2014$) using the CPS ASEC samples.

[65]Between 1988 and 1995, 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.

8 Conclusion

Head Start has served low-income children for more than 50 years, yet the effectiveness of the program is still an open question. To shed new light on the ongoing debate, this paper presents new evidence on the effect of exposure to HS on economically disadvantaged children's third grade test scores. In this paper, I utilize previously unexplored variation, local community federal funding expansions in the 1990s, to identify the effect of HS. During the 1990s, the federal government doubled the funding for HS with the aim of improving both quality and capacity. The significant funding expansion within a relatively short time created a natural experiment that resulted in large variation in the program funding expansions across communities and over time.

Using student-level administrative data from the Texas Education Agency, I find that exposure to generous HS funding during childhood leads to substantial gains in third grade test scores for low-income students, in particular students from academically disadvantaged backgrounds. I show that HS benefits Hispanics with limited language proficiency the most. I provide quantitative evidence that HS significantly improves language proficiency among Hispanics, which could partly explain the test score gains in this group. These results are consistent with previous studies' findings that HS is more effective for dual language learners and children with low levels of baseline scores (e.g., Puma et al., 2010; Bitler et al., 2014).

I also show that the gains in test scores can be explained by an increase in both program quality and capacity. In particular, I find that HS funding expansions are associated with reductions in child-teacher ratios, increases in full time enrollment and education spending. In addition, I provide suggestive evidence that test score gains in third grade persist through fifth grade. These findings contrast to the previous literature, which reports smaller or no detectable impacts on children's test scores over the medium-run (Deming, 2009; Puma et al., 2012; Carneiro and Ginja, 2014). A possible explanation for the discrepancy is that the present paper provides evidence on children who would have participated in HS in the 1990s. Also, my sample consists of low-income children in Texas which includes a large concentration of Hispanics, different from other studies.

This study provides new evidence that furthers our understanding of the long-standing question of lasting benefits of the HS program. Additional analyses of the efficacy of the program provide extra insight for policymakers considering future public investments in early childhood education. As the new administration takes place, early childhood investments are receiving significant political attention, and it is important for policymakers to be able to measure the benefits of the program, not only on cognitive outcomes, but also on non-cognitive, health and labor market outcomes. My findings suggest that additional federal funding exposure to HS significantly improves test scores for low-income children. In addition, a conservative analysis indicates that HS passes the cost-benefit test by a big margin.

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.

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 ntervention. *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.

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.

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. (2006). The Take-up of Social Benefits. In Alan Auerbach, D. C. and Quigley, J., editors, *Public Policy and the Income Distribution*. Russell Sage Foundation.

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., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., et al. (2007). School readiness and later achievement. *Developmental psychology*, 43(6):1428.

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

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*.

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. *The 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.

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., Johnson, R. C., and Persico, C. (2015). The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms. Working Paper 20847, National Bureau of Economic Research.

Johnson, R. C. (2010). School quality and the Long-Run Effects of Head Start. *Goldman School of Public Policy, University of California, Berkeley, Working Paper*.

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.

Ladd, H. F., Muschkin, C. G., and Dodge, K. A. (2014). From Birth to School: Early Childhood Initiatives and Third-Grade Outcomes in North Carolina. *Journal of Policy Analysis and Management*, 33(1):162–187.

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.

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.*

Puma, M., Bell, S., Cook, R., Heid, C., Shapiro, G., Broene, P., Jenkins, F., Fletcher, P., Quinn, L., Friedman, J., et al. (2010). Head Start Impact Study: Final Report. *Administration for Children & Families.*

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. (2016). Head Start's Impact in the Very Long-Run. *College of Letters & Science Economics, University of Wisconsin-Milwaukee, Working Paper.*

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.

9 Tables

Table 1: Sample Characteristics of Third Grade Students in Texas

	Full Sample	Not Disadv. Sample	Free/Reduced Lunch Eligible Sample					
			All	Male	Female	White	Black	Hispanic
Head Start								
Real HS per child	406	312	516	516	515	346	306	683
	(686)	(431)	(885)	(882)	(888)	(401)	(227)	(1137)
Outcomes								
Standardized Reading Score	0.00	0.29	-0.34	-0.47	-0.21	-0.09	-0.32	-0.43
	(1.00)	(0.78)	(1.11)	(1.16)	(1.04)	(1.03)	(1.05)	(1.14)
Standardized Math Score	0.00	0.29	-0.33	-0.40	-0.27	-0.06	-0.37	-0.41
	(0.99)	(0.77)	(1.11)	(1.16)	(1.06)	(0.98)	(1.03)	(1.17)
Controls								
Year of Birth	1987.16	1987.06	1987.28	1987.29	1987.28	1987.27	1987.29	1987.26
	(2.07)	(2.06)	(2.09)	(2.08)	(2.09)	(2.09)	(2.07)	(2.09)
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)	(0.42)	(0.42)	(0.42)	(0.00)	(0.00)	(0.00)
African-American	0.14	0.07	0.22	0.22	0.22	0.00	1.00	0.00
	(0.35)	(0.26)	(0.41)	(0.41)	(0.42)	(0.00)	(0.00)	(0.00)
Hispanic	0.32	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	0.46	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)
Limited English Proficiency	0.11	0.02	0.21	0.21	0.20	0.01	0.00	0.36
	(0.31)	(0.14)	(0.41)	(0.41)	(0.40)	(0.09)	(0.07)	(0.48)
Bilingual	0.06	0.01	0.13	0.13	0.13	0.00	0.00	0.24
	(0.24)	(0.08)	(0.34)	(0.34)	(0.34)	(0.04)	(0.02)	(0.43)
Participates ESL Program	0.03	0.01	0.06	0.06	0.06	0.01	0.00	0.09
	(0.18)	(0.10)	(0.23)	(0.24)	(0.23)	(0.08)	(0.06)	(0.29)
Participates in a Special Education Program	0.09	0.07	0.11	0.15	0.08	0.15	0.12	0.10
	(0.29)	(0.26)	(0.32)	(0.36)	(0.26)	(0.35)	(0.32)	(0.30)
Participates in a Gifted/Talented Program	0.06	0.08	0.03	0.03	0.03	0.02	0.03	0.03
	(0.23)	(0.28)	(0.16)	(0.16)	(0.17)	(0.14)	(0.17)	(0.17)
N	739427	407730	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 1995. For detailed description, see Section 3.

Table 2: Community-Level Correlates of Average Real HS Spending per Child, 1988-1995

	(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 local community-level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 3: Community-Level Correlates of Change in Real HS Spending per Child, 1988-1995

	(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 1995. 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 local community-level. * p<0.10, ** p<0.05, *** p<0.01.

Table 4: Effect of HS per child on Pre-K Enrollment, 1988-1995

	Enrolled in Pre-K/Pop 3-4	Enrolled in Pre-K/Poor Pop 3-4
	(1)	(2)
Real HSPC/1000	-0.004	-0.010
	(0.006)	(0.027)
Mean Y	0.179	0.550
Mean X ($)	431	431
Adj. R-Squared	0.706	0.375
Obs	911	911

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, year fixed effects and 1980 community characteristics interacted with linear trends. Standard errors are clustered at the local community-level. * p<0.10, ** p<0.05, *** p<0.01.

Table 5: Baseline Estimates of the Effect of HS Exposure on Third Grade Standardized Math Scores

	All	Males	Females	Whites	Blacks	Hispanics
Real HSPC/1000	0.081**	0.087	0.066***	-0.004	0.109	0.114***
	(0.034)	(0.056)	(0.025)	(0.055)	(0.154)	(0.032)
Mean Y	-0.331	-0.395	-0.267	-0.058	-0.371	-0.410
Mean X($)	516	516	515	346	306	683
Adj. R-Squared	0.388	0.399	0.373	0.433	0.103	0.384
Obs	332910	165303	167607	72073	75614	180775

Notes: This table contains results obtained when the dependent variable is third grade standardized test scores in math and the independent variable is real federal HS spending per child (2014$) when 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 community-specific linear trends. Sample consists of students who are eligible 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 1995. Standard errors are clustered at the local community-level. * p<0.10, ** p<0.05, *** p<0.01.

Table 6: Effect of HS Exposure on Third Grade Standardized Math Scores
Differential Effects by Gender

	Whites		Blacks		Hispanics	
	Males	Females	Males	Females	Males	Females
Real HSPC/1000	-0.062	0.051	-0.055	0.312*	0.143***	0.075***
	(0.066)	(0.088)	(0.214)	(0.170)	(0.048)	(0.024)
Mean Y	-0.116	0.000	-0.470	-0.275	-0.465	-0.355
Mean X($)	347	346	307	306	680	685
Adj. R-Squared	0.451	0.416	0.097	0.104	0.392	0.375
Obs	35747	36326	36984	38630	90259	90516

Notes: This table contains results obtained when the dependent variable is third grade standardized test scores in math and the independent variable is real federal HS spending per child (2014$) when 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 community-specific linear trends. Sample consists of students who are eligible 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 1995. Standard errors are clustered at the local community-level. * p<0.10, ** p<0.05, *** p<0.01.

Table 7: Effect of HS Exposure on Third Grade Standardized Math Scores for Hispanics
Differential Effects by Language Proficiency

	Limited Lang.			Proficient Lang.		
	All	Males	Females	All	Males	Females
Real HSPC/1000	0.153***	0.188***	0.105	0.035	0.036	0.027
	(0.055)	(0.065)	(0.064)	(0.049)	(0.038)	(0.078)
Mean Y	-0.913	-0.966	-0.857	-0.120	-0.166	-0.075
Mean X($)	932	920	944	542	539	544
Adj. R-Squared	0.390	0.386	0.397	0.367	0.384	0.347
Obs	66688	34027	32661	114087	56232	57855

Notes: This table contains results obtained when the dependent variable is third grade standardized test scores in math and the independent variable is real federal HS spending per child (2014$) when 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 community-specific linear trends. Sample consists of students who are eligible 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 1995. Standard errors are clustered at the local community-level. * p<0.10, ** p<0.05, *** p<0.01.

Table 8: Effect of HS Exposure on Language Proficiency in Third Grade for Hispanics
Differential Effects by Gender

	Hispanics		
	All	Males	Females
Real HSPC/1000	0.019**	0.021**	0.016**
	(0.008)	(0.011)	(0.007)
Mean Y	0.631	0.623	0.639
Mean X($)	686	680	685
Adj. R-Squared	0.847	0.843	0.852
Obs	180775	90259	90516

Notes: This table contains results obtained when the dependent variable is an indicator for language proficiency and the independent variable is real federal HS spending per child (2014$) when 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 Hispanic students who are eligible 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 1995. Standard errors are clustered at the county-level. * p<0.10, ** p<0.05, *** p<0.01.

Table 9: Effect of Head Start Exposure on Test Scores for Each Grade

	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
A: All						
Real HSPC/1000	0.081**	0.062**	0.058*	-0.007	-0.001	-0.032
	(0.034)	(0.029)	(0.034)	(0.035)	(0.029)	(0.074)
Mean Y	-0.331	-0.278	-0.280	-0.261	-0.285	-0.256
Mean X ($)	516	510	502	461	413	388
Adj. R-Squared	0.388	0.361	0.475	0.523	0.531	0.487
Obs	332910	502225	473779	458568	445483	422951
B: Females						
Real HSPC/1000	0.066***	0.051**	0.029	-0.002	0.002	-0.028
	(0.025)	(0.023)	(0.025)	(0.042)	(0.027)	(0.069)
Mean Y	-0.267	-0.204	-0.190	-0.151	-0.173	-0.152
Mean X ($)	516	512	505	467	419	391
Adj. R-Squared	0.373	0.336	0.451	0.503	0.514	0.463
Obs	167607	252748	239755	230452	222586	210098
C: Males						
Real HSPC/1000	0.087	0.096***	0.093*	-0.008	-0.025	-0.026
	(0.056)	(0.033)	(0.048)	(0.039)	(0.046)	(0.090)
Mean Y	-0.395	-0.352	-0.373	-0.373	-0.396	-0.359
Mean X ($)	515	510	499	456	408	386
Adj. R-Squared	0.399	0.380	0.492	0.532	0.537	0.501
Obs	165303	249477	234024	228116	222897	212853
D: Whites						
Real HSPC/1000	-0.004	0.014	-0.002	0.025	-0.025	-0.138
	(0.055)	(0.028)	(0.038)	(0.037)	(0.069)	(0.091)
Mean Y	-0.058	-0.118	-0.145	-0.135	-0.162	-0.223
Mean X ($)	346	334	325	301	265	240
Adj. R-Squared	0.433	0.480	0.635	0.608	0.613	0.531
Obs	72073	88817	81254	79671	73550	64430
E: Blacks						
Real HSPC/1000	0.109	0.091	0.050	-0.056	0.161	0.413
	(0.154)	(0.080)	(0.063)	(0.154)	(0.109)	(0.275)
Mean Y	-0.371	-0.328	-0.383	-0.420	-0.423	-0.375
Mean X ($)	306	299	287	263	233	211
Adj. R-Squared	0.103	0.490	0.634	0.644	0.646	0.606
Obs	75614	107615	98825	91542	85873	79788
F: Hispanics						
Real HSPC/1000	0.114***	0.071**	0.061*	0.010	-0.012	-0.057
	(0.032)	(0.033)	(0.035)	(0.039)	(0.032)	(0.084)
Mean Y	-0.410	-0.298	-0.272	-0.246	-0.277	-0.232
Mean X ($)	683	672	658	574	511	481
Adj. R-Squared	0.389	0.315	0.409	0.471	0.483	0.450
Obs	180775	299488	289133	280935	279617	272675

Notes: This table contains results obtained when the dependent variable is standardized test score in math for each grade and the independent variable is real federal Head Start spending per child (2014$) when 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 community-specific linear trends. Sample consists of students who are eligible 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 1983 and 1995. Standard errors are clustered at the local community-level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 10: Effect of HS per child on HS Enrollment, 1988-1995

	Enrolled in HS/Pop 3-4	Enrolled in HS/Poor Pop 3-4
	(1)	(2)
Real HSPC/1000	0.102***	0.277***
	(0.012)	(0.037)
Mean Y	0.076	0.200
Mean X ($)	431	431
Adj. R-Squared	0.778	0.721
Obs	911	911

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. For more details about the data description, see Section 3. All regressions include local community, year fixed effects and 1980 community characteristics interacted with linear trends. Standard errors are clustered at the local community-level. * p<0.10, ** p<0.05, *** p<0.01.

Table 11: Mechanisms: 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
Adj. R-Squared	0.900	0.889	0.616	0.632	0.749	0.532	0.488	0.339
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 1993-1995 for some programs. Director's salary is available for years 1992-1995. The rest of the variables are available for years 1988-1995. 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 thousand dollars. 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 local community-level. * p<0.10, ** p<0.05, *** p<0.01.

Table 12: Effect of HS Exposure on Third Grade Standardized Math Scores
Sensitivity of Results to Specifications

	Main Results	Omit Community Trends	Omit Pre-K Controls	Omit Income Controls	Omit Safety Net Controls	Add School Trends
Real HSPC/1000	0.081**	0.090**	0.080**	0.076**	0.067**	0.057*
	(0.034)	(0.039)	(0.035)	(0.034)	(0.033)	(0.033)
Mean Y	-0.331	-0.331	-0.331	-0.331	-0.331	-0.331
Mean X ($)	516	516	516	516	516	516
Adj. R-Squared	0.388	0.386	0.388	0.388	0.388	0.408
Obs	332910	332910	332910	332910	332910	332910

Notes: This table contains results obtained when the dependent variable is third grade standardized test scores in math and the independent variable is real federal HS spending per child (in 2014$) when the child was four years old. All regressions include controls for demographics, school, test year and birth year fixed effects. Sample consists of students who are eligible 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 1995. Standard errors are clustered at the local community-level. * p<0.10, ** p<0.05, *** p<0.01.

Table 13: Effect of HS Exposure on Third Grade Standardized Math Scores
Sensitivity of Results to Different Measures of HS Exposure

	Per Child (Age 3-4)	Per Capita	Per Poor Child (Age 3-4)
	(1)	(2)	(3)
Real HSPC/1000 (Main)	0.081**		
	(0.034)		
Real HS Per Capita/1000		1.955***	
		(0.528)	
Real HS Per Poor Child/1000			0.023
			(0.015)
Mean Y	-0.331	-0.331	-0.331
Mean X($)	516	19	1236
Adj. R-Squared	0.388	0.388	0.388
Obs	332910	332910	332910

Notes: This table contains results obtained when the dependent variable is third grade test scores in math. 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. 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 eligible 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 1995. Standard errors are clustered at the county-level. * p<0.10, ** p<0.05, *** p<0.01.

Table 14: Falsification Tests: Effect of HS Exposure on Third Grade
Standardized Math Scores For Sample Identified as Not Disadvantaged

	All	Males	Females	Whites	Blacks	Hispanics
Real HSPC/1000	-0.0001	0.006	-0.019	-0.008	0.132	0.037
	(0.030)	(0.041)	(0.030)	(0.041)	(0.233)	(0.036)
Mean Y	0.288	0.268	0.310	0.372	-0.083	0.019
Mean X($)	376	375	376	342	321	636
Adj. R-Squared	0.324	0.334	0.319	0.252	0.353	0.414
Obs	407730	206649	201081	316346	31539	49871

Notes: This table contains results obtained when the dependent variable is third grade standardized test scores in math and the independent variable is real federal HS spending per child (in 2014$) when 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 *not* eligible 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 1995. Standard errors are clustered at the county-level. * p<0.10, ** p<0.05, *** p<0.01.

Table 15: Falsification Tests: Effect of HS Exposure on Exogenous Student Characteristics and County-level Income

	Black	Hispanic	Female	Free/Reduced Meal	Income PC	Predicted Math Score
Real HSPC/1000	-0.000	0.002	0.002	0.025	-136.423	-0.001
	(0.001)	(0.003)	(0.001)	(0.017)	(130.470)	(0.001)
Mean Y	0.140	0.352	0.495	0.380	36379.401	-0.002
Adj. R-Squared	0.378	0.492	0.008	0.329	0.985	0.461
Obs	11220926	11220926	11220926	11220926	11220926	11220926

Notes: This table contains results obtained when the dependent variable is student characteristics and county-level income per capita and the independent variable is real federal HS spending per child (in 2014$) when 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 all the students from third to eighth grade. 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 1995. Standard errors are clustered at the county-level. * p<0.10, ** p<0.05, *** p<0.01.

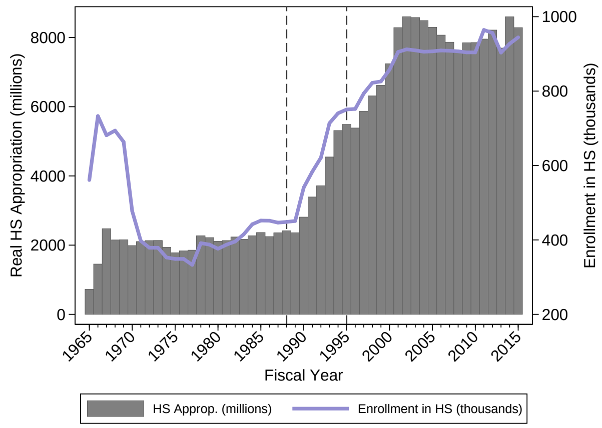
Table 16: Effect of HS Exposure on Third Grade Standardized Math Scores Differential Effects by Age of Exposure, 0-8

	Age 0	Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8
Real HSPC/1000	0.007	0.013	-0.068	0.025	0.081**	0.072*	-0.014	-0.004	0.008
	(0.054)	(0.037)	(0.048)	(0.036)	(0.034)	(0.039)	(0.020)	(0.018)	(0.013)
Mean Y	-0.331	-0.331	-0.331	-0.331	-0.331	-0.331	-0.331	-0.331	-0.331
Adj. R-Squared	0.396	0.395	0.394	0.390	0.388	0.393	0.395	0.392	0.391
Obs	332910	332910	332910	332910	332910	332910	332910	332910	332910

Notes: This table contains results obtained when the dependent variable is third grade standardized test scores in math and the independent variable is real federal Head Start spending per child (in 2014$), assigned when the child was ages of 0 to 8. 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 eligible 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 1995. Standard errors are clustered at the county-level. * p<0.10, ** p<0.05, *** p<0.01.

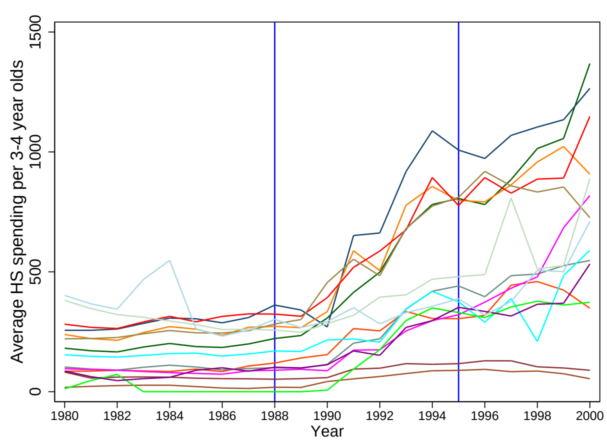
10 Figures

Figure 1: Head Start Program Facts



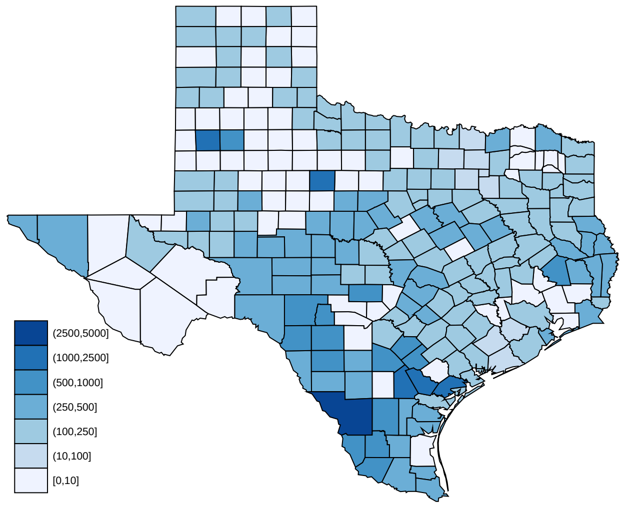
Notes: The data are from the HHS website: https://eclkc.ohs.acf.hhs.gov/hslc/data/factsheets/2015-hs-program-factsheet.html. Federal Head Start appropriations are in 2014 dollars. The dashed lines highlight the period of this study, from 1988 to 1995.

Figure 2: Real HS Spending per 3-4 year old in the 15 Most Populous Counties in Texas



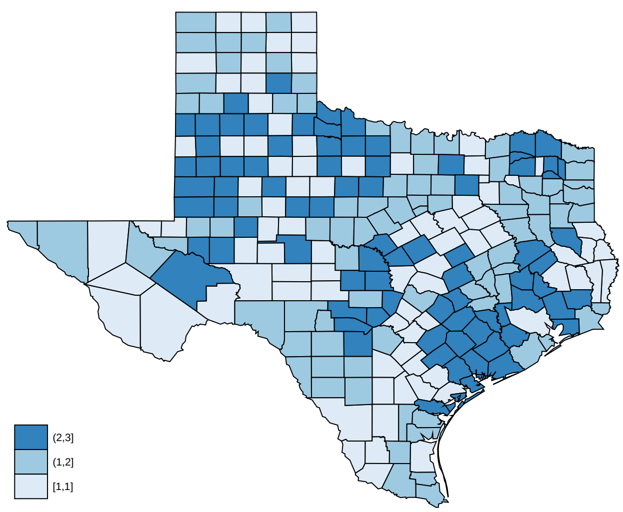
Notes: Head Start spending (in 2014$) data are from the Consolidated Federal Funds Reports (CFFR), coupled with the population counts for three- and four-year-olds at the county-level from the SEER. For more details about data construction, see Section 3. 15 most populous counties include Bend, Bexar,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

Figure 3: Real HS Spending per 3-4 year old (1988)



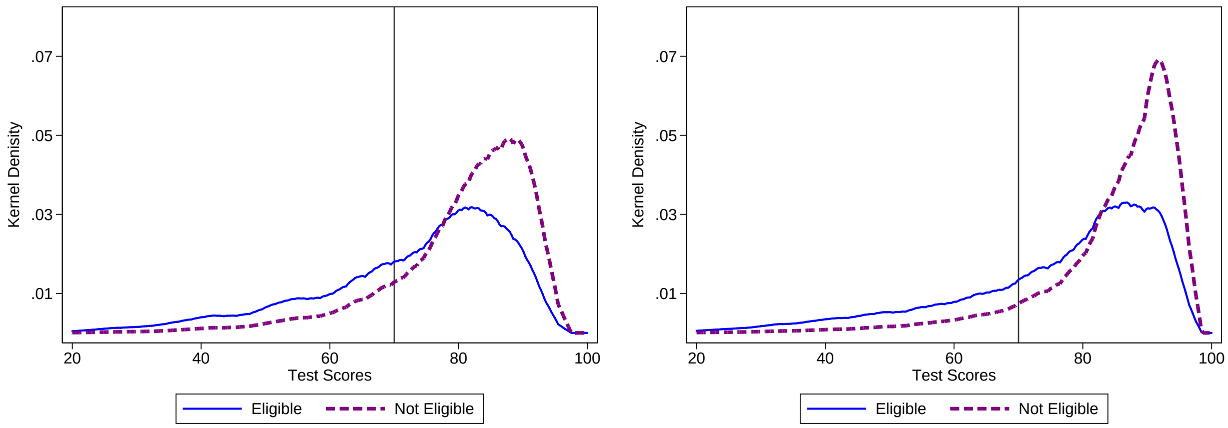
Notes: Head Start spending (in 2014$) data are from the Consolidated Federal Funds Reports (CFFR), coupled with the population counts for three- and four-year-olds at the county-level from the SEER. For more details about data construction, see Section 3.

Figure 4: Growth in Real HS Spending per child in Texas (1988-1995)



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.

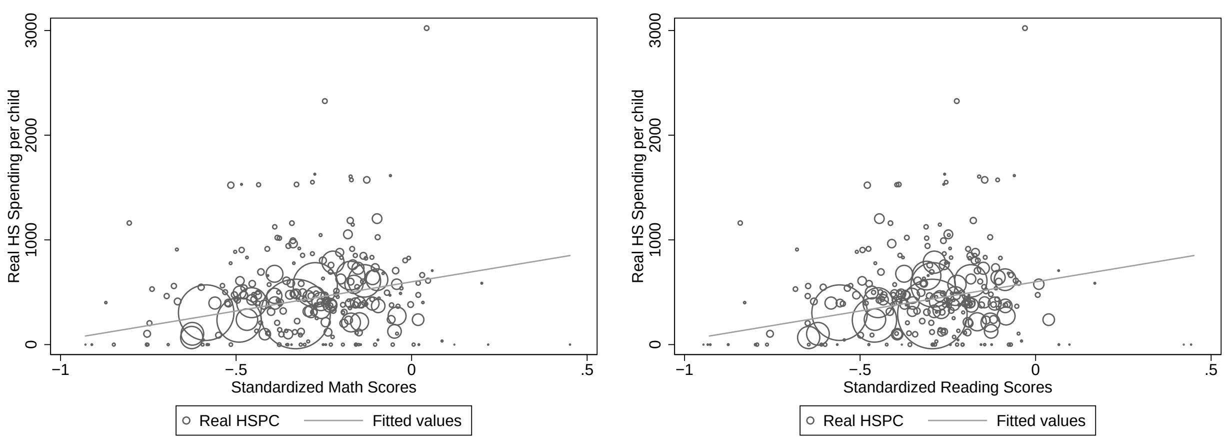
Figure 5: Kernel Density, by Free or Reduced Lunch Eligibility



(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 eligible 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.

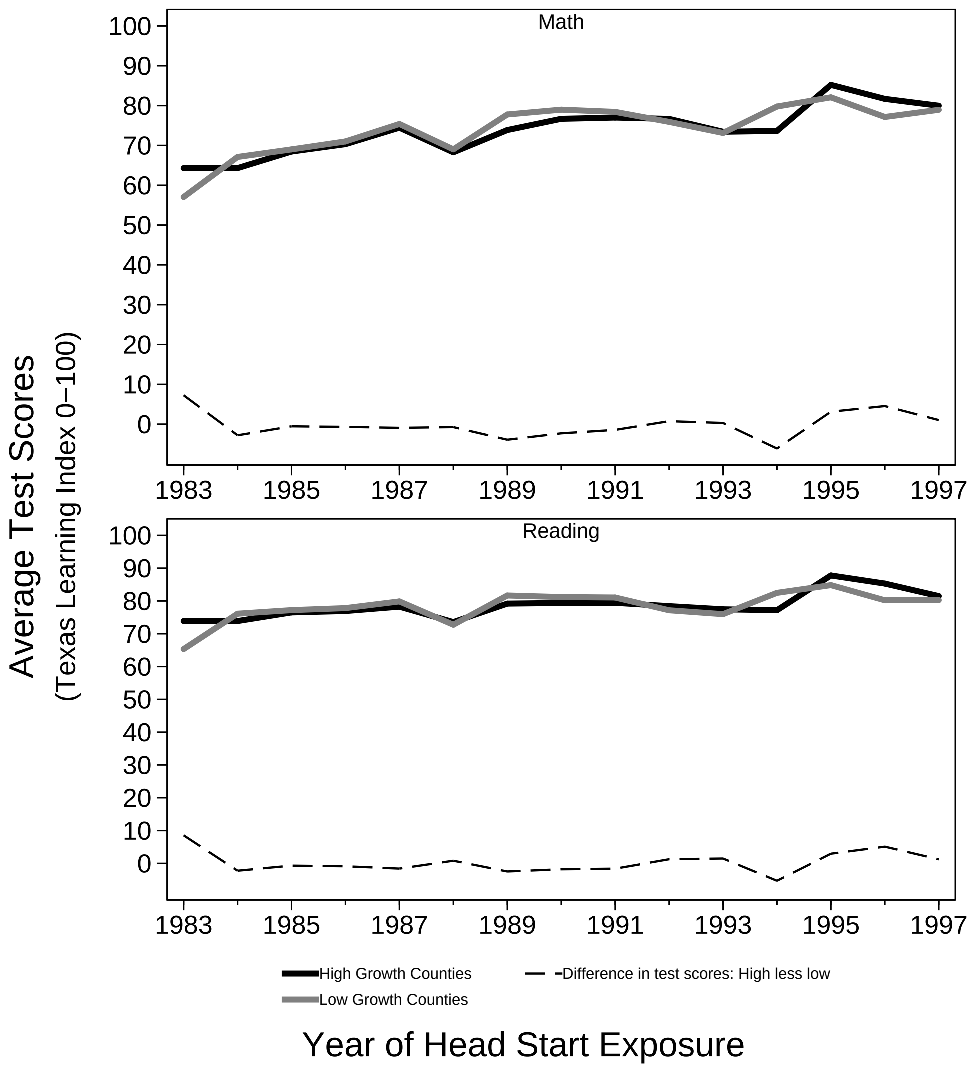
Figure 6: Raw Correlations between HS Spending per child and Standardized Test Scores
Free or Reduced Lunch Eligible Sample



(a) Math (b) Reading

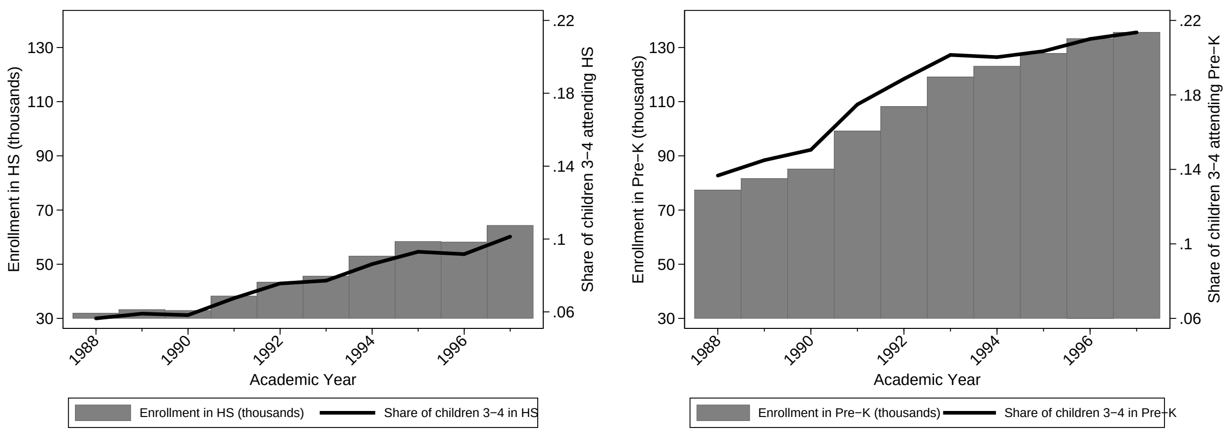
Notes: Head Start spending (in 2014$) data are obtained from the Consolidated Federal Funds Reports and third grade student test score data are from the Texas Education Agency (TEA) between 1994 and 1999. The data are collapsed to the county-level using averages. The bubbles present the local communities,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

Figure 7: Average Test Scores and the Year of Exposure
High and Low Growth Counties in HS Spending



Notes: This figure plots the average test scores estimated using a community-year panel model separately for high and low growth counties over the years of federal HS funding exposure, defined as low/high relative to the median growth in spending per child with the base level of 1983. The dashed line plots the difference in the average test scores between high and low growth counties.Test score data include all students from third to eighth grade who took the standardized test in Texas between 1994 and 1999, from the Texas Education Agency (TEA). Texas Learning Index is a difficulty-weighted scoring metric designed to maintain consistency in test scoring across testing cohorts which ranges between 0 and 100.

Figure 8: Early Childhood Education Expansions in Texas

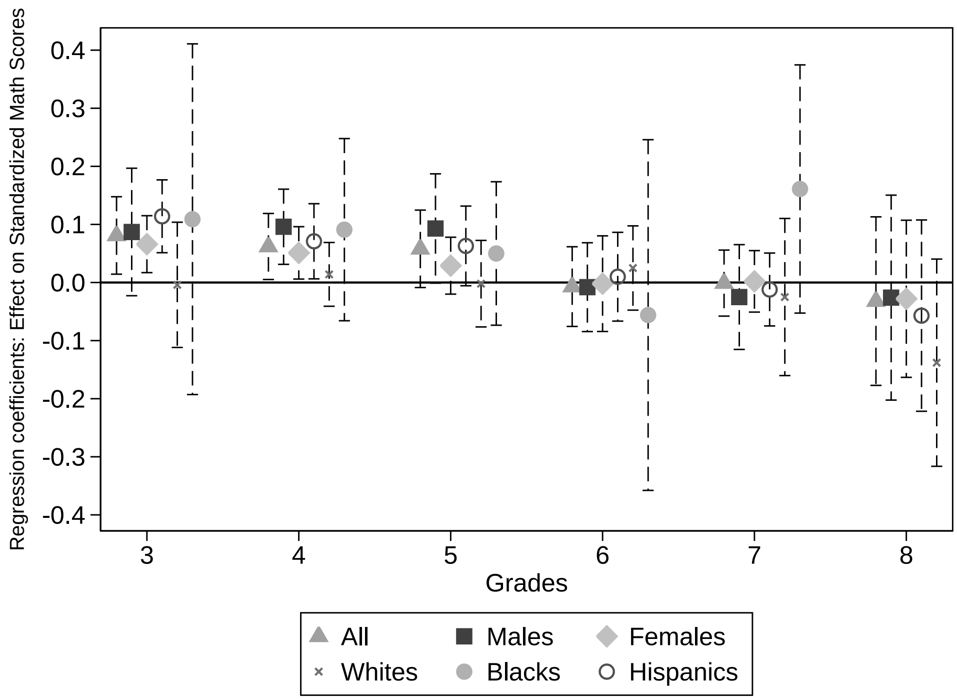


(a) HS expansions

(b) Pre-K expansions

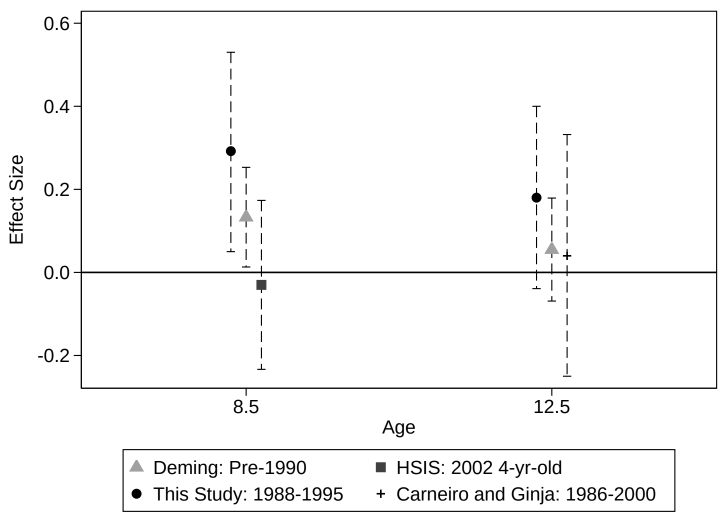
Notes: Head Start enrollment data are from the Program Information Reports (PIR). Pre-K enrollment data are from the Common Core Data (CCD). The share measure is calculated using the population counts for three- and four-year-olds from the SEER.

Figure 9: Effect of Head Start Exposure on Test Scores for Each Grade



Notes: This figure plots coefficients obtained when the dependent variable is standardized test scores in math for each grade and the independent variable is real federal Head Start spending per child (in 2014$) when the child was four years old and their 95% confidence intervals. The effect in eighth grade for blacks is dropped for clarity of visualization. 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 eligible 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. Head Start spending data are from Consolidated Federal Files Reports (CFFR) and include years between 1983 to 1995. Standard errors are clustered at the local community-level.

Figure 10: Comparison of Effect Sizes for Pre-1990, 1990s, and 2002 Head Start Cohorts



Notes: This figure plots effect sizes and their 95% confidence intervals from three studies: (1) the estimated effects by age from Deming (2009) for pre-1990 HS cohorts, (2) the impacts by age for the 4 year old cohort in the Head Start Impact Study for 2002 cohort from Puma et al. (2012), (3) the impacts by age using current study which analyzes HS years from 1988 and 1995, and (4) the estimated PIAT-Math effect for males at ages 12-13 for HS participants from 1986 to 2000 from Carneiro and Ginja (2014).

A Appendix

A.1 Data Appendix

<u>**Public Use Data**</u>

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

  These data span from 1983-2010, available through Census.[66] 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:

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

  Restricting the years to 1983-1995, I construct the spending data at the grantee-year level. One grantee could serve one county or a group of counties. The regions that each grantee served are determined using administrative dataset PCCOST provided by Currie and Neidell (2007) which groups the counties served by each grantee for years 1990-2000.[67] Regions for 1983-1989 are assigned based on the 1990 data assuming that serving counties did not change between 1983-1990.[68] Additional data are used to confirm these regions, 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,[69] locate in Bryan, Texas (Brazos County). This grantee serves HS programs in eight other counties.[70] Thus, the raw data as shown in Appendix Figure A.3a records $4.3 million in expenditures for Brazos County in 1994 and zero

[66] http$://$www2.census.gov/pub/outgoing/govs/special60/

[67] These data are downloaded from Professor Matthew Neidell's website.

[68] This assumption is not as worrisome for my main analysis because the core period of interest is from 1988 to 1995.

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

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

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.3b 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.[71]
  2. County-level population to construct the per capita measure also comes from the SEER. Instead of extracting specific age groups, I collapse 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 age eligible poor child counts, I follow Frisvold (2006) which states that children who are age 3-4 years old are two fifth of children under age five. These data are only available for years 1989, 1993, 1995, 1997-1999, the years in between is determined through linear interpolation.

- **Program Information Reports (PIR)**:[72] 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)**:[73] CCD includes the school level information for all public schools. These data are available starting 1986 at the school level and provides 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

[71]Source: http://www.nber.org/data/seer_u.s._county_population_data.html.
[72]Source: https://eclkc.ohs.acf.hhs.gov/hslc/data/pir
[73]Source: https://nces.ed.gov/ccd/pubschuniv.asp

estimates, such as other War on Poverty program 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 Information 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*: Eligible 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

A.2 Tables

Table A.1: Effect of HS Exposure on Third Grade Standardized Math Scores
Sensitivity of Results to Adding Limited Language Proficiency Control

	All	Males	Females	Whites	Blacks	Hispanics
Real HSPC/1000	0.073**	0.077	0.058**	-0.004	0.109	0.102***
	(0.032)	(0.054)	(0.024)	(0.055)	(0.154)	(0.029)
Mean Y	-0.332	-0.397	-0.269	-0.058	-0.371	-0.412
Mean X($)	517	518	517	346	306	686
Adj. R-Squared	0.392	0.403	0.378	0.433	0.103	0.389
Obs	332910	165303	167607	72073	75614	180775

Notes: This table contains results obtained when the dependent variable is third grade standardized test scores in math and the independent variable is real federal HS spending per child (2014$) when 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 community-specific linear trends. Sample consists of students who are eligible 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 1995. Standard errors are clustered at the local community-level. * p<0.10, ** p<0.05, *** p<0.01.

Table A.2: Effect of HS Director's Qualifications on Third Grade Standardized Math Scores

	Director's Qualifications			
	Has BA+	Yrs of Educ.	Yrs of Exper.	Salary
Dependent variable: z-score Math Test	0.007	-0.005	0.0004	-0.001
	(0.024)	(0.006)	(0.001)	(0.000)
Mean Y	-0.278	-0.278	-0.278	-0.278
Mean X	0.8	4.6	10.1	44.028
Adj. R-Squared	0.106	0.106	0.106	0.106
Obs	168285	168285	168285	168285

Notes: This table contains results obtained when the dependent variable is third grade standardized math test scores and the independent variables include Head Start program directors' characteristics when 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 community-specific linear trends. Sample consists of students who are eligible for free/reduced lunch or in poverty based on the description by the Texas Education Agency (TEA). Student data are from the Texas Education Agency which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 and 1999. Head Start program directors' data are from the Program Information Reports (PIR), include years between 1992 to 1995. Standard errors are clustered at the local community-level. * p<0.10, ** p<0.05, *** p<0.01.

Table A.3: Effect of HS Exposure on Third Grade Standardized Reading Scores

	All	Males	Females	Whites	Blacks	Hispanics
Real HSPC/1000	0.033	0.033	0.028	-0.012	0.036	0.046
	(0.026)	(0.036)	(0.050)	(0.054)	(0.118)	(0.031)
Mean Y	-0.338	-0.470	-0.208	-0.091	-0.316	-0.432
Mean X($)	516	516	515	346	306	683
Adj. R-Squared	0.421	0.430	0.395	0.483	0.114	0.409
Obs	332910	165303	167607	72073	75614	180775

Notes: This table contains results obtained when the dependent variable is third grade standardized test scores in reading and the independent variable is real federal HS spending per child (2014$) when 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 community-specific linear trends. Sample consists of students who are eligible 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 1995. Standard errors are clustered at the local community-level. * p<0.10, ** p<0.05, *** p<0.01.

Table A.4: Effect of HS Exposure on Third Grade Standardized Reading Scores
Differential Effects by Gender

	Whites		Blacks		Hispanics	
	Males	Females	Males	Females	Males	Females
Real HSPC/1000	-0.071	0.064	-0.082	0.182	0.061**	0.029
	(0.056)	(0.085)	(0.156)	(0.144)	(0.029)	(0.055)
Mean Y	-0.224	0.041	-0.485	-0.154	-0.547	-0.317
Mean X($)	347	346	307	306	680	685
Adj. R-Squared	0.501	0.445	0.096	0.095	0.414	0.392
Obs	35747	36326	36984	38630	90259	90516

Notes: This table contains results obtained when the dependent variable is third grade standardized test scores in reading and the independent variable is real federal HS spending per child (2014$) when 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 community-specific linear trends. Sample consists of students who are eligible 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 1995. Standard errors are clustered at the local community-level. * p<0.10, ** p<0.05, *** p<0.01.

Table A.5: Baseline Estimates of HS Exposure on Third Grade Standardized Math Scores Adding School Trends

	All	Males	Females	Whites	Blacks	Hispanics
Real HSPC/1000	0.057*	0.048	0.047*	-0.043	0.093	0.103***
	(0.033)	(0.053)	(0.024)	(0.062)	(0.154)	(0.034)
Mean Y	-0.331	-0.395	-0.267	-0.058	-0.371	-0.410
Mean X($)	516	516	515	346	306	683
Adj. R-Squared	0.408	0.420	0.394	0.446	0.126	0.411
Obs	332910	165303	167607	72073	75614	180775

Notes: This table contains results obtained when the dependent variable is third grade standardized test scores in math and the independent variable is real federal HS spending per child (2014$) when 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 school-specific linear trends. Sample consists of students who are eligible 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 1995. Standard errors are clustered at the local community-level. * p<0.10, ** p<0.05, *** p<0.01.

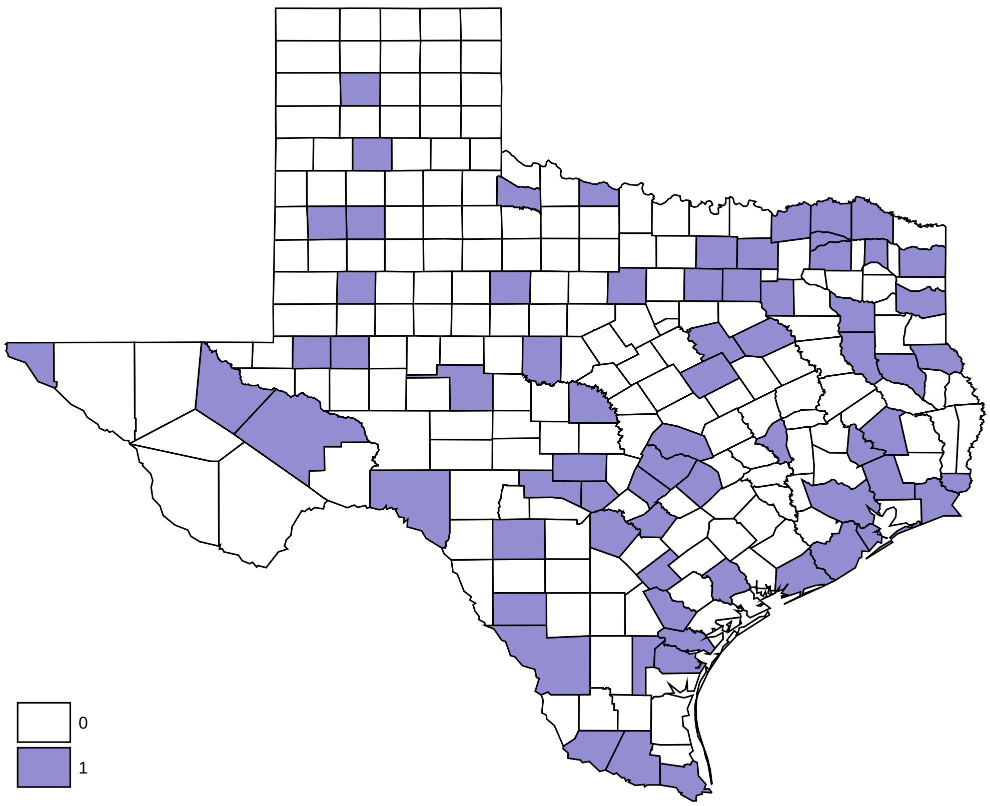
Table A.6: Effect of Head Start Exposure on Test Scores for Each Grade Balanced for Cohorts Born 1984-1988

	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
A: All						
Real HSPC/1000	0.009	0.010	0.066**	0.054	0.066**	-0.014
	(0.056)	(0.032)	(0.027)	(0.069)	(0.045)	(0.061)
Mean Y	-0.389	-0.355	-0.343	-0.268	-0.296	-0.243
Mean X	421	399	396	392	387	388
SD X	0.769	0.740	0.725	0.722	0.712	0.728
Adj. R-Squared	0.355	0.486	0.612	0.467	0.523	0.487
Obs	300251	329068	266877	318287	294271	313199
B: Females						
Real HSPC/1000	0.080	0.013	0.008	0.046	-0.061	-0.074
	(0.062)	(0.066)	(0.059)	(0.091)	(0.057)	(0.070)
Mean Y	-0.327	-0.252	-0.228	-0.151	-0.186	-0.143
Mean X	419	400	396	393	390	390
Adj. R-Squared	0.337	0.458	0.599	0.450	0.509	0.467
Obs	151156	161801	131455	157478	146789	155941
C: Males						
Real HSPC/1000	-0.082	0.007	0.144**	0.062	-0.034	-0.059
	(0.073)	(0.045)	(0.063)	(0.066)	(0.061)	(0.071)
Mean Y	-0.452	-0.454	-0.455	-0.383	-0.405	-0.342
Mean X	422	399	396	391	388	386
Adj. R-Squared	0.371	0.503	0.619	0.472	0.528	0.497
Obs	149095	167267	135422	160809	147482	157258
D: Whites						
Real HSPC/1000	-0.005	-0.014	0.011	0.081	-0.026	-0.278
	(0.118)	(0.058)	(0.077)	(0.132)	(0.134)	(0.266)
Mean Y	-0.084	-0.156	-0.180	-0.148	-0.178	-0.192
Mean X	265	263	261	254	247	240
SD X	0.284	0.279	0.292	0.291	0.271	0.269
Adj. R-Squared	0.370	0.615	0.682	0.495	0.615	0.542
Obs	63176	61860	50619	57886	56587	69248
E: Blacks						
Real HSPC/1000	-0.091	-0.009	0.116	0.447*	-0.294	-0.223
	(0.162)	(0.114)	(0.144)	(0.266)	(0.318)	(0.242)
Mean Y	-0.438	-0.373	-0.461	-0.430	-0.415	-0.387
Mean X	228	222	222	218	215	211
Adj. R-Squared	0.365	0.629	0.713	0.546	0.636	0.603
Obs	69680	72198	57515	63552	56554	62076
F: Hispanics						
Real HSPC/1000	0.024	0.021	0.053**	0.020	-0.067	-0.086
	(0.074)	(0.041)	(0.025)	(0.062)	(0.067)	(0.061)
Mean Y	-0.473	-0.400	-0.340	-0.251	-0.297	-0.221
Mean X	565	518	508	494	4801	481
Adj. R-Squared	0.355	0.436	0.577	0.442	0.479	0.443
Obs	164703	191449	155943	191890	160008	198868

Notes: This table contains results obtained when the dependent variable is standardized test scores in math for each grade and the independent variable is Head Start spending per child when 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 eligible for free/reduced lunch or in poverty based on the description by the Texas Education Agency (TEA) and who are born between 1984 and 1988 (balanced in the cohort-test year level). Student data are from the TEA which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 to 1999. Head Start spending data are from Consolidated Federal Files Reports (CFFR) and include years between 1988 to 1992. Standard errors are clustered at the local community-level. * p<0.10, ** p<0.05, *** p<0.01.

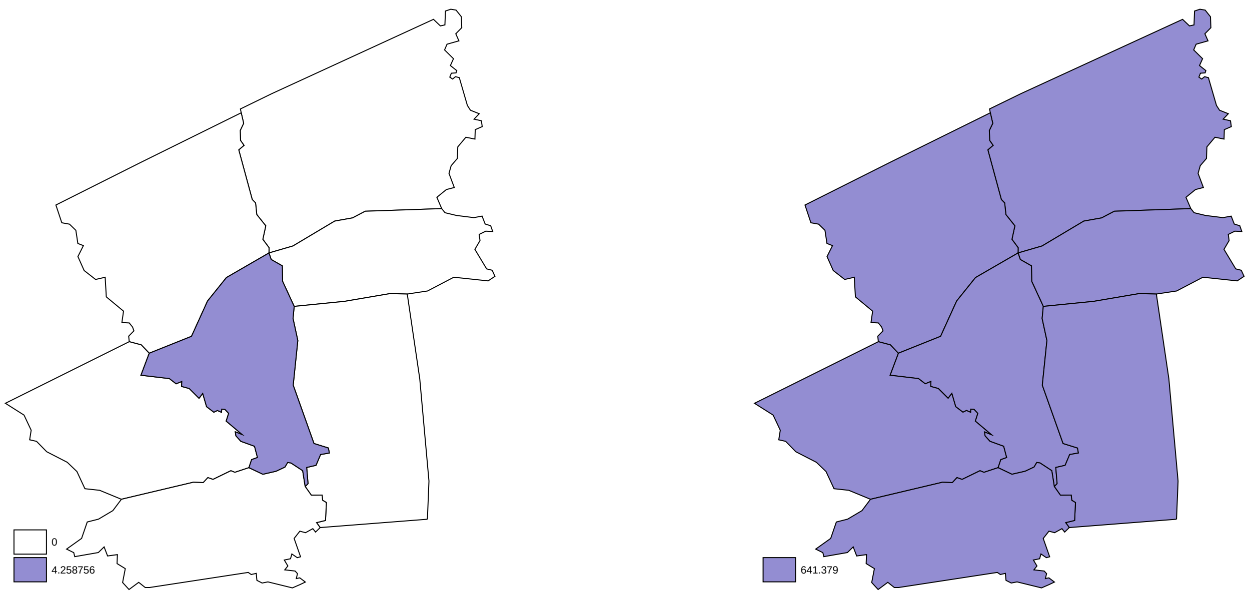
A.3 Figures

Figure A.1: 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 serves more than 200 counties in 1994.

Figure A.2: 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 an 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 that Brazos 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.

B Online Appendix

B.1 Tables

Table OA.1: Estimates of the Effect of HS Exposure using HS per Capita on Third Grade Standardized Math Scores

	All	Males	Females	Whites	Blacks	Hispanics
Real HS Per Capita/1000	1.955***	2.377**	1.345***	0.070	2.231	2.456***
	(0.528)	(0.912)	(0.434)	(1.656)	(4.950)	(0.510)
Mean Y	-0.331	-0.395	-0.267	-0.058	-0.371	-0.410
Mean X($)	18	18	18	11	10	25
Adj. R-Squared	0.388	0.399	0.373	0.433	0.103	0.384
Obs	332910	165303	167607	72073	75614	180775

Notes: This table contains results obtained when the dependent variable is third grade standardized test scores in math and the independent variable is real federal HS spending per capita (2014$) when 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 community-specific linear trends. Sample consists of students who are eligible 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 1995. Standard errors are clustered at the local community-level. * p<0.10, ** p<0.05, *** p<0.01.

Table OA.2: Estimates of the Effect of HS Exposure using HS per Poor Child on Third Grade Standardized Math Scores

	All	Males	Females	Whites	Blacks	Hispanics
Real HS Per Poor Child/1000	0.023	0.025	0.017	-0.001	0.041	0.043***
	(0.015)	(0.021)	(0.012)	(0.017)	(0.055)	(0.013)
Mean Y	-0.331	-0.395	-0.267	-0.058	-0.371	-0.410
Mean X($)	1195	1197	1194	1059	914	1384
Adj. R-Squared	0.388	0.399	0.373	0.433	0.103	0.384
Obs	332910	165303	167607	72073	75614	180775

Notes: This table contains results obtained when the dependent variable is third grade standardized test scores in math and the independent variable is real federal HS spending per poor child (2014$) when 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 community-specific linear trends. Sample consists of students who are eligible 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 1995. Standard errors are clustered at the local community-level. * p<0.10, ** p<0.05, *** p<0.01.

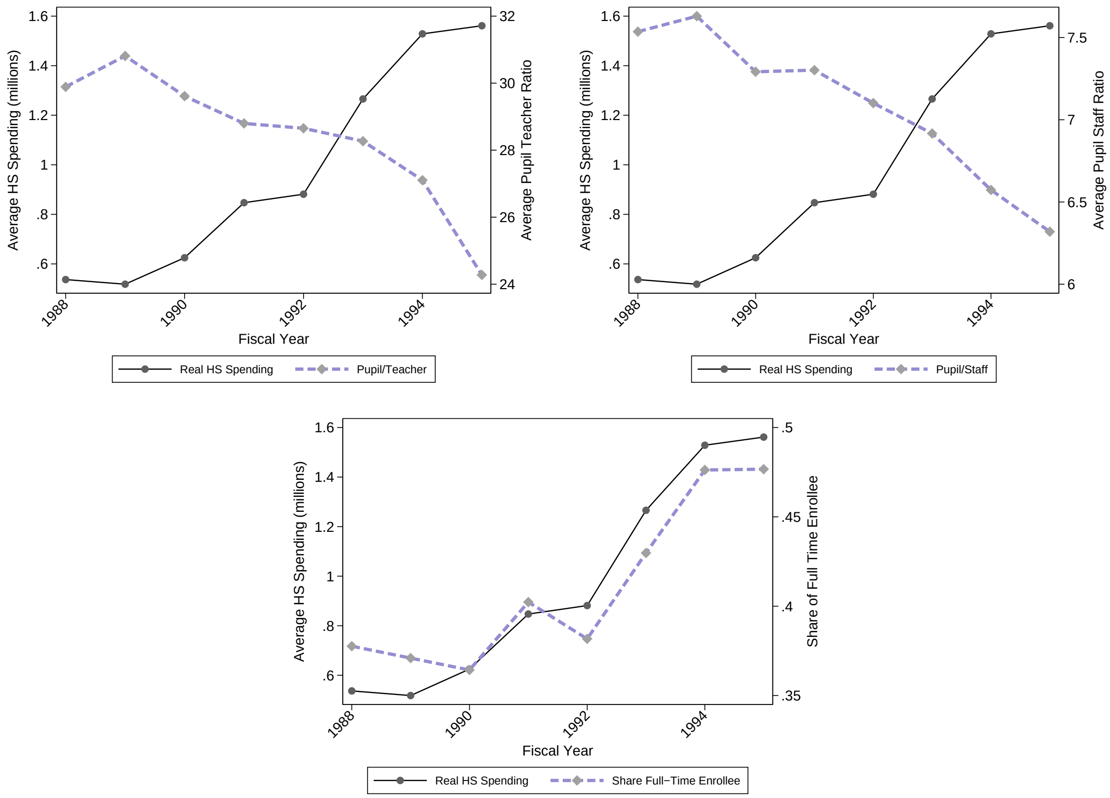
Table OA.3: Director Quality and Head Start Funding, 1992-1995

	(1)	(2)	(3)
Director has BA+ degree	-0.098		
	(1.502)		
Director's yrs of eduation		0.167	
		(0.277)	
Director's yrs of experience			-0.023
			(0.091)
Mean Y (in millions)	4.774	4.774	4.774
Mean X	0.750	4.763	7.547
Adj. R-Squared	0.749	0.750	0.749
Obs	264	264	264

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), include years between 1992 to 1995. 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 local community-level. * p<0.10, ** p<0.05, *** p<0.01.

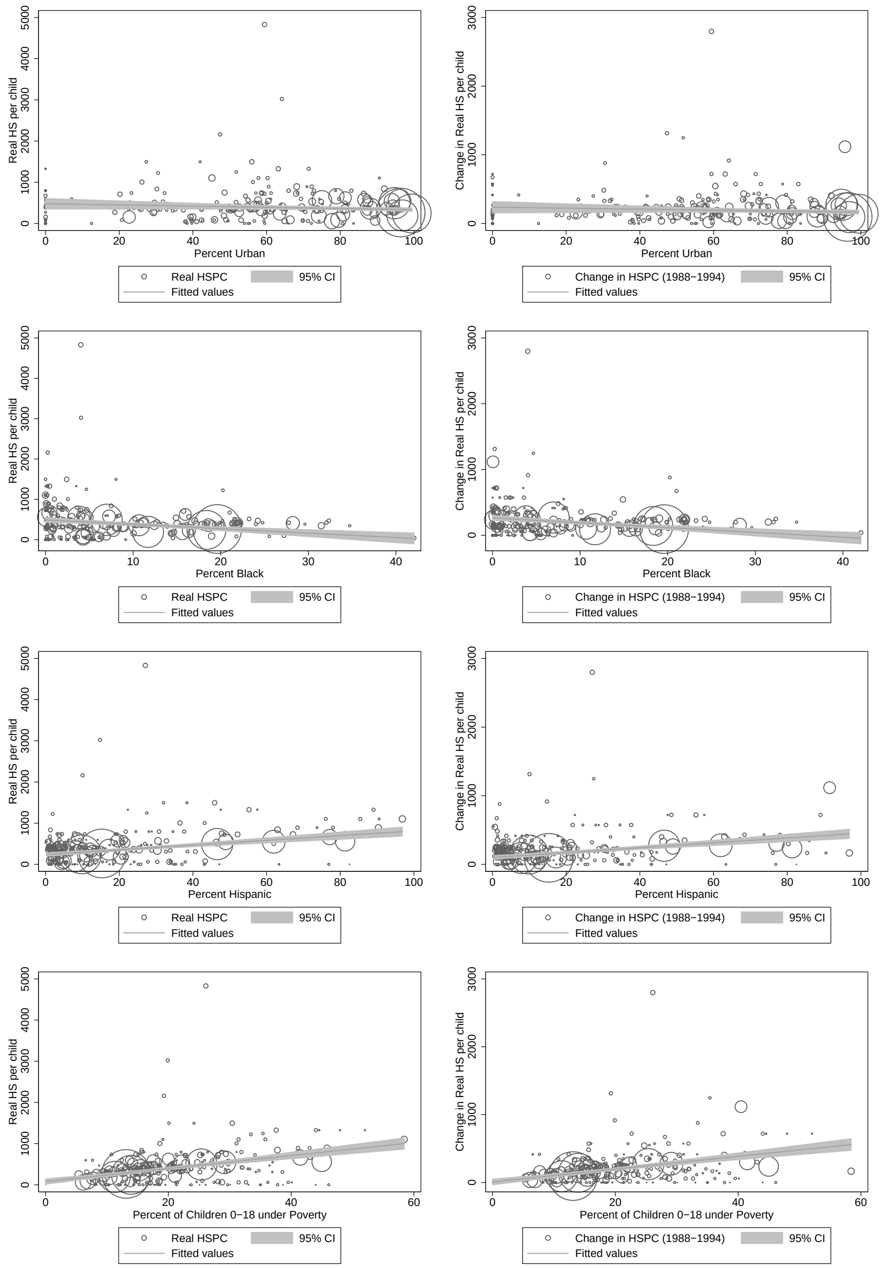
B.2 Figures

Figure OA.1: HS 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.

Figure OA.2: Correlations between 1980 County Characteristics and HS 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 1988-1995. On the right, the figures show the correlations between the county characteristics and the *change* in HS spending per child from 1988 to 1995.