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

Esra Kose*

May 1, 2021

Abstract

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

JEL classification codes: H53, I38, J13

Keywords: Early Childhood Education, Head Start, Program Evaluation

---

*Esra Kose, Assistant Professor, Department of Economics, Bucknell University, One Dent Drive Lewisburg, PA 17837 (email: ek016@bucknell.edu). I am thankful to Doug Miller, Marianne Page, Marianne Bitler, and Scott Carrell for their help and support. I also thank Nicholas J. Sanders, Elira Kuka, Na'ama Shenhav, Hilary Hoynes, Chris Walters, Aline Buetikofer, Gaetano Basso, Cassie Hart, Vasco Yasenov, Michel Grosz, Ariel Pihl, and brown bag participants at UC Davis for providing input. I am grateful to Matthew Neidell and David Frisvold for generously providing data. Student-level data are from the Texas Education Agency. These data are currently administered by the University of Texas Education Research Centers in Austin and Dallas and are available for use to approved researchers. An earlier version of this paper circulated under the title "Local Area Spending Exposure to Head Start and Academic Performance: Evidence from Texas." All errors are my own.

I. Introduction

Racial, ethnic, and socioeconomic disparities in academic achievement have been ongoing issues in the U.S. education system (Reardon, Robinson and Weathers, 2008). These disparities persist over time, raising serious concerns about the children's life prospects and the state of social mobility in the U.S. (Heckman, 2006; Duncan and Murnane, 2011).

Do early childhood investments play a role in narrowing the academic achievement gap? Early childhood is a period of physical, cognitive, socioemotional, and language development that a child's environment can influence. There is strong theoretical support in the fields of economics, neuroscience, and child development that points to early childhood as a critical time to invest. Despite the strong theoretical support, regardless of the program's scale, many early childhood interventions targeting skill development demonstrate initially promising results that fade out when students reach the third grade (Bailey et al., 2017). On the other hand, prior research provides evidence on the long-term effects of such programs on children's socioeconomic and health outcomes. Given the promising initial results and long-term gains, the fade-out pattern in medium-run outcomes is puzzling.

I study whether early childhood investments close the achievement gap in elementary school in the context of the largest federal early childhood program in the U.S., Head Start. The program began in 1965 to provide education, health, and social services to low-income children. Although Head Start is shown to improve children's long-term outcomes,[1] critiques have indicated that the program has not fulfilled its mission of closing the achievement gap across socioeconomic groups when children are in elementary school (Barnett, 2011). The best available evidence is based on a random assignment of 4,442 children to a national sample of Head Start centers in the early 2000s. It reports small impacts on test scores at the end of the program participation year that fade out by the third grade (Puma et al., 2012). Studying the period of the 1990s when federal Head Start

[1] The following studies show positive long-term effects of Head Start: Currie and Thomas (1995, 1999); Garces, Thomas and Currie (2002); Deming (2009); Bauer and Schanzenbach (2016); Thompson (2018); Bailey, Sun and Timpe (2020); Barr and Gibbs (2019); Johnson and Jackson (2019).

funding grew substantially, I show that the program reduced the achievement gap in the third grade between Hispanic and white students in Texas.

The 1990s were a golden age for Head Start due to immense expansion and the passing of two important acts to improve the program's capacity and quality. This paper investigates two research questions related to this time period. First, to what extent did federal Head Start funding expansions during the 1990s affect student performance in Texas? Second, how does the effect of public investments relate to the way funds are spent? The investigation of the first question provides new insight into the old debate. The second question presents new evidence of whether public investments were spent as planned by linking administrative data on Head Start program characteristics and budgets to funding expansions. The evidence concerning the second question has been limited, primarily due to data availability and quality (Currie and Neidell, 2007). Furthermore, which program quality measures in early childhood education matter most for child outcomes is still an open question (Blau and Currie, 2006; Walters, 2015).

My empirical strategy uses a panel fixed effects model that leverages variation of cohort exposure across counties in the timing and intensity of Head Start funding expansions in the 1990s. The main identification assumption is that funding expansions are exogenous to other underlying geographic-level trends in test scores. I employ various analyses to check the reliability and robustness of the estimates and show the results are not driven by the expansion of alternative early childhood programs during this time, selection bias, or other endogenous factors.

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

My main finding is that exposure to Head Start funding expansions significantly improved academic performance in the third grade. For free/reduced lunch (FRL) certified students, a \$500 increase in Head Start funding per child led to a 0.03 standard deviation *(σ)* increase in average test scores in math and reading combined.[2] Using several sources on program characteristics and budgets, I show that additional Head Start funding led to significant increases in program participation rates, number of teachers hired, and full-time enrollment. Moreover, federal funding expansions are associated with more generous spending on education, nutrition, social services, and parent involvement in the Head Start program. In total, this analysis suggests that both program capacity and quality improvements are essential pathways for the ultimate effect on test scores.

Estimates by race and ethnicity show that improvements among Hispanic students are the primary driver of these results. Overall, among low-income students, the funding increase during the 1990s led to a 13 percent reduction of the test score gap in math and reading combined between Hispanics and whites. These gains persist through the fifth grade. There are at least three testable channels through which Head Start could be beneficial for Hispanics. First, Hispanics have higher participation rates in Head Start in Texas compared to whites and blacks combined. I find that additional funding induced more Hispanic children to participate in the program. Second, the program participation could improve their language skills by exposing them to English at an early age, which could affect their academic performance. My analysis shows that Head Start funding exposure significantly improved their language proficiency. Third, the program could help special needs children during childhood; Head Start centers must reserve 10 percent of their enrollment for children with disabilities. I show that additional funding led to a reduction in the likelihood of having special education status for Hispanics.

This paper contributes to the economics literature in several ways. First, my analyses provide new evidence on the effects of Head Start on academic achievement by taking advantage of a different policy lever (spending, not participation), a different time period, and a unique

[2] This intent-to-treat (ITT) estimate of 0.03σ corresponds to a treatment effect on the treated (TOT) of an approximately 0.3σ increase in test scores.

population. Using spending increases in the 1990s in Texas, I find that Head Start expansions led to improvements in medium-run outcomes, especially for low-income Hispanic students.

Second, my findings align well with a growing number of papers that find public investments and education policies in early childhood, or elementary school, are more effective for Hispanics and children with low levels of baseline scores (for example, Currie and Thomas, 1999; Bitler, Hoynes and Domina, 2014; Gibbs, 2014; Figlio and Ozek, 2019). My study explores the use of public funds in early childhood for low-income children during the 1990s, which has been understudied.

Third, I contribute to the growing literature that examines the effect of public spending on academic success. The most recent causal evidence shows that public spending during the formal school year improves student outcomes (Jackson, 2018). What is less known is whether public spending during early childhood makes a difference in academic performance. My paper provides new evidence on this question in the context of Head Start.

The rest of the paper is organized as follows. I provide background information on the Head Start expansions and review the prior research in Section II. Section III describes data sources, followed by an overview of the methodology in Section IV. I report my results in Section V and discuss potential mechanisms in Section VI. I then present robustness checks in Section VII and conclude in Section VIII.

II. Background and Prior Literature

A. The Head Start Program

Head Start is a federally funded early childhood education program that provides education, health, nutrition, and other services to economically disadvantaged children and their families. The program is designed to reduce disparities in school readiness, health, and other social aspects between low-income children and their more advantaged peers. The main eligibility criteria are that children must be between three and five and come from families with income at or below the poverty level. Also, at least 10 percent of the children served in each center must have some type of disability, regardless of the income eligibility.

Head Start began as a part of the "War on Poverty" initiative in 1965. Between 1990 and 2000, enrollment increased by about 60 percent, and federal funding per enrolled child doubled (see Appendix Figure A.1). The 1990s expansion resulted from a significant effort by the Bush and Clinton Administrations to improve capacity and quality constraints. Additional funding was used to increase enrollment, improve teacher salaries and training, expand services for families of children attending the program, and help local Head Start agencies purchase facilities. As a result of these policy efforts, there was a substantial ramp-up in federal funding per child, providing a natural experiment to study the program.

Head Start is a federal-local matching grant program. The federal government determines the program's funding annually, as a component of the federal budget, and allocates it to states based on the relative number of public assistance recipients, unemployed persons, and children from families below the poverty line (Community Services Act of 1974). To receive funding, local agencies must write grant proposals directly to the Head Start Bureau.[3] Grantees must provide at least 20 percent of the financing, which may include in-kind contributions through community partnerships. These 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).[4] As a result, there is geographic variation in funding levels, which provides part of the identification in this paper.

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

[3] Grants are issued through a competitive process, with priority given to agencies that can demonstrate the most cost-effective operation, and existing programs have priority when reapplying (Currie and Neidell, 2007).

[4] However, each center must comply with publicly known standards, which are described in the Head Start Act of 2007.

identify the program's effectiveness. Section IV addresses some of these issues in detail.

B. Prior Literature

Head Start has been evaluated extensively over its existence.[5] Prior research reports initial effects on cognitive outcomes that fade out when students reach the third grade. In the long-run, Head Start is shown to be effective in improving socioeconomic and health outcomes.

The best available evidence on the short-term impacts of Head Start on children comes from the Head Start Impact Study (HSIS), an independent national random assignment study of Head Start that took place in the early 2000s and followed the children up to the third grade. The reports by Puma et al. (2005, 2010, 2012) show that (1) one year of Head Start improved cognitive skills by a small amount; (2) by the end of the first grade, the effects mostly faded out; and (3) on the third-grade follow-up, the impact of Head Start disappeared. More recent papers, which address the flaws in the HSIS, find significant positive effects of Head Start for short-term test score impacts up to the first grade (Feller et al., 2016; Zhai, Brooks-Gunn and Waldfogel, 2014; Kline and Walters, 2016). Additionally, Bitler, Hoynes and Domina (2014) find larger short-term impacts at low quantiles of the test score distribution and persistent effects through the first grade for Spanish speakers at the bottom of the test score distribution. However, these recent papers do not provide evidence on test score effects in the third grade.

The best available evidence on the longer-term effects of Head Start comes from quasi-experimental studies. Previous literature has exploited within-family comparisons of siblings who have and have not participated in the program (for example, Currie and Thomas, 1995, 1999; Garces, Thomas and Currie, 2002; Deming, 2009; Bauer and Schanzenbach, 2016). Other papers use discontinuities due to program funding and eligibility rules (Ludwig and Miller, 2007; Carneiro and Ginja, 2014). A few recent papers use variation in Head Start funding expansions in its early introduction (Thompson, 2018; Barr and Gibbs, 2019; Johnson and Jackson, 2019; Bailey, Sun and Timpe, 2020). These papers show that Head Start has been effective in improving

[5] Several studies have reviewed the literature on Head Start's effectiveness (Barnett, 1995; Currie, 2001; Barnett and Hustedt, 2005; Ludwig and Phillips, 2008; Shager et al., 2013; Duncan and Magnuson, 2013; Gibbs, Ludwig and Miller, 2013).

the long-term socioeconomic and health outcomes.

Similar to Thompson (2018) and Johnson and Jackson (2019) that use Head Start funding expansions across counties over time in the 1960s, my paper uses a new variation that stems from the 1990s program expansion to analyze the program's impact on academic performance. Additionally, I use student-level administrative data in my study, which is advantageous to overcome statistical inference issues. Exploring Head Start's impact on the low-income student population in Texas, which is predominantly Hispanic, I show that the program led to significant gains in academic achievement in third grade through fifth grade.

III. Data Construction and Summary Statistics

A. Data Construction

I combine several datasets on student test scores and demographics, Head Start spending, school quality, and economic conditions to analyze the effect of the program's expansion on academic performance. This section provides an overview of the data sources; Appendix II includes detailed descriptions.

Administrative student-level data include the universe of exam takers in Texas from the third grade through the eighth grade between 1994 and 1999. The majority of the analysis focuses on the third-grade students. These data are from the TEA and include test scores monitored through the Texas Academic Assessment System (TAAS). Relevant to my paper, these data contain information on the year of birth, sex, ethnicity, FRL status, language proficiency, and special education status for each student at each school district. From the school district information, I determine the county of residence, which serves as a proxy for the child's county of residence at age four. Along with the year of birth, this determines each student's exposure to Head Start funding.

I conduct the following sample restrictions. Starting with the universe of students who took the TAAS test between 1994 and 1999, I first drop observations with missing demographic information, missing test scores, exempt testing status, or nonstandard test administration. Next, using the administrative data description, I restrict my analysis to students who are certified for FRL or are identified as economically disadvantaged based on their families' welfare eligibility because they are more likely to be eligible for Head Start. From here, I will refer to this sample as FRL certified.

I develop a measure of Head Start funding per age-eligible child in a county using several sources. Head Start spending data are from the CFFR, which include information on local appropriations for federally funded programs for each grantee. A Head Start grantee could oversee one county or a group of counties. This factor is considered when constructing the county-level spending by incorporating information from the Head Start budgets (PCCOST). Head Start spending years are restricted to 1988 and 1994, when the third-grade students would be age-eligible to attend Head Start. To construct the population denominator of children ages three and four years old,[6] county-level population counts for each age group are extracted from the Surveillance, Epidemiology, and End Results Program (SEER). Together with these three data sources, I construct the Head Start funding per child variable at the county-year level.

I augment these data with additional data sources to bring information on (1) Head Start enrollment and program budget items, (2) school quality and other pre-kindergarten (pre-K) alternatives in Texas, and (3) economic conditions. First, I compiled data from the Program Information Reports (PIR). Starting in 1988, the Office of Head Start has collected comprehensive data from all grantees and delegates on the services, staff, children, and families served by the program. These data provide information on the number of funded enrollment, staff, demographic composition of children and staff, and director qualifications. Second, I source the school-level pupil-teacher ratio and information on school-level state-funded pre-K (outside of Head Start) enrollment from the Common Core of Data (CCD).[7]

Next, I compiled data on county-level economic conditions as controls. These include per capita income, per capita transfer payments for cash income support, medical benefits, food assistance,

[6] Three- and four-year-old children make up to approximately 90 percent of Head Start enrollment (Head Start Fact Sheet, 1990).

[7] One caveat of using the CCD is that it only reports the enrollment in public pre-K programs that take place in public elementary schools, excluding public pre-K programs that take place in private preschools or daycare. Therefore, it does not provide the complete picture of the preschool scene.

retirement, and disability programs from the Regional Economic Information System (REIS) data. I also control for percentage urban population; percentage black; percentage Hispanic; percentage single parent; percentage of 0- to 18-year-olds living in poverty from the 1983 City and County Data Book; and unemployment rate from the Bureau of Labor Statistics.

B. Summary Statistics

My analysis centers on children exposed to Head Start funding in Texas between 1988 and 1994. Figure 1 presents the average Head Start spending per child in the 15 most populous counties in Texas from 1980 to 1995.[8] This figure shows substantial variation in Head Start spending per child across counties and within a county over time. For more insight into the geographic variation in Head Start funding, Figure 2 presents maps of (i) funding levels in 1988 and (ii) the growth from 1988 to 1994, respectively.[9] These maps show a great deal of variation across counties in both levels and growth in funding. My basic identification strategy uses this geographic and time variation to identify the effect of Head Start on academic achievement.

Table 1 presents summary statistics for third-grade students, first for the full sample and economically advantaged students and then for FRL-certified students. The first three columns show a large discrepancy in average test scores by economic status.[10] Economically advantaged students have substantially higher test scores relative to FRL-certified students. Importantly, Hispanics make up more than 50 percent of these FRL-certified students. They have the lowest test scores, on average, relative to other economically disadvantaged students and tend to live in counties with higher funding for Head Start per child.

Figure 3 shows the evolution of standardized test scores across birth cohorts by their FRL status. The most striking observation is that there is a considerable achievement gap among racial and ethnic groups, even among the noncertified students. For students who were noncertified (Figure 3.a.) and certified (Figure 3.b.) for FRL, the academic performance of Hispanic and

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

[9] The growth measure is calculated using 1988 as the base period, and the shades are determined based on the terciles of the growth distribution.

[10] The distribution of student test scores in math and in reading by FRL status is shown in Appendix Figure A.2.

black students seems to be improving across cohorts. Overall, these observations suggest there is value in analyzing the potential effects of early childhood investments separately by race and ethnicity.

To demonstrate the relationship between average test scores and Head Start funding, I restrict the sample to FRL-certified children and collapse the data on test scores to county-level averages. Appendix Figure A.3 shows a positive correlation between average test scores and Head Start spending per child for FRL-certified children.

IV. Empirical Strategy

To study Head Start's effect on academic performance, I exploit variation in county-level Head Start funding per child in the first half of the 1990s. Following Ludwig and Miller (2007) and Sanders (2012), I assign Head Start funding exposure based on each student's county of residence and birth year.[11] My empirical strategy uses a county and birth year panel fixed effects approach, which relies on variation in Head Start funding per child within counties and over time, conditional on observables. Formally, I estimate the following equation for the sample of FRL-certified students using a newly assembled dataset on county-level Head Start spending per child:

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

where Y_{iscbt} denotes the outcome variable (standardized test scores) for student i in school s in county c in birth year b, and in test year t. $HSfunding_{c(b+4)}$ represents Head Start funding per child in county c when student i was four years old ($= b + 4$). X_{iscbt} is a vector of individual-level demographic controls, including sex, ethnicity, and an indicator for bilingual and English as a second language. Z_{ct} has three sets of county-level controls, including per capita income transfers, average characteristics of students collapsed from the main data, and additional controls, which include share of zero- to five-year-old population by ethnicity, log population, and

[11] Because individuals are not linked across the years, I rely on the county of residence at each grade to assign the treatment. If low-income families make migration decisions based on the availability of the services provided by Head Start, the assignment using county of residence would create bias in my estimates. I test this directly in Appendix Table A.1 and find no evidence that the Head Start spending affects the student composition within school over time.

unemployment rate. $W_{c(b+4)}$ includes county-level controls at the time of Head Start, 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. $\pi_c * b$ is a county-specific linear trend. Standard errors are clustered at the county level. The coefficient of interest is β, which is interpreted as the conditional change in the outcome variable from a unit (500$) increase in exposure to federal Head Start funding per child at age four.

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

First, I take county-level characteristics measured in 1980, before the expansions occurred, and use them to predict the levels and changes in Head Start spending per child from 1988 to 1994 (similar to Hoynes and Schanzenbach, 2012). For this analysis, I collapse the data at the county level. The independent variables include the percentage of the 1980 population living in an urban area, black, Hispanic, single parent, less than age 5, age 65 or older, total population, percentage of 0- to 18-year-olds living in poverty as well as income, education, and welfare spending per capita (in 2014 dollars).

The results are presented in Appendix Tables A.2 and A.3.[12] Simple correlations imply that counties with a larger Hispanic population, a more extensive farmland, and a higher share of single mothers, poor, very young, or elderly have more Head Start funding. In contrast, counties with a larger black population tend to have less funding, which could be because blacks in Texas live in more urban areas with high population density. Additionally, counties with more social spending have higher funding for Head Start, and counties with higher per capita income have lower funding. The determinants of the change in Head Start funding from 1988 to 1994 are also similar (see Appendix Table A.3).

After controlling for the characteristics described above, the last columns of Appendix Tables A.2 and A.3 show that the percentage of blacks and percentage of 0- to 18-year- old population

[12] See Appendix Figure A.4 for a visual presentation of the correlations.

living under poverty are significant determinants of Head Start expansions, and together all variables explain around 20 percent of the overall variation (R-squared ≈ 0.20). Nevertheless, to control for possible differences in trends across counties, I include county-specific time trends in my analyses. The results are robust to the exclusion of these trends.

Second, director quality could be a possible confounding factor if directors who can obtain more funds may also run better programs in other aspects. For example, a bias will occur if better quality directors write better grants to receive additional funds and then operate higher-quality programs. In this case, funding levels could be correlated with child outcomes (Frisvold, 2006; Currie and Neidell, 2007). To rule out this possibility, I show in Appendix Table A.4 that federal funding increases are not significantly associated with directors' qualifications.[13] A bias could also occur if some communities devote local resources toward children and if these resources during childhood also help children succeed in their school years. Including school fixed effects controls for this possibility as well as for other neighborhood characteristics that could be correlated with student success.

Next, Head Start 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 Head Start's introduction corresponds to the introduction or expansion of 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-old children, I directly control for county-level spending for other social programs.[14] Moreover, including county-specific linear time trends also accounts for the fact that some communities may be improving over time.

Finally, critical to my identification strategy is that Head Start spending expansions should not systematically change a particular cohort's composition within a school. For example, if low-income families decided to move based on Head Start's funding generosity and more generous communities had better schools, such nonrandom selection would misattribute higher student

[13] 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 the years between 1992 and 1994.

[14] See Section III for a detailed discussion about which controls are added.

performance to Head Start exposure. In Appendix Table A.1, I formally test this and other types of self-selection by examining whether student characteristics such as sex, race/ethnicity, and county-level income per capita are correlated with funding after conditioning on school fixed effects. I show that Head Start funding variation does not predict changes in the composition within school, suggesting the estimates are not biased by families' self-selection into a particular cohort within a school.

A. State-Provided Early Childhood Education Alternatives in Texas

A potential threat to my identification strategy is the existence of state-provided preschool alternatives to Head Start, as all preschools provide similar services to improve children's school readiness. Texas has offered a half-day public pre-K program since the 1985-86 academic year. State-provided pre-K aims to enhance children's academic performance by providing early childhood education for four-year-old children identified as at risk.[15] Although it is mandatory for any district with at least 15 eligible children to offer a half-day education-based program for four-year-old children, attendance is voluntary (TEC §29.1532).[16] Thus, starting in 1985, an eligible child in Texas could attend a public pre-K as an alternative to Head Start. Although funding for pre-K is allocated directly to school districts by the state of Texas, districts are encouraged to partner with licensed childcare centers and Head Start programs to provide preschool services (Barnett et al., 2011).[17] This raises the possibility that Head Start and pre-K might have operated jointly to some degree during the 1990s.

Appendix Figure A.5 plots the number and share of children enrolled in Head Start and pre-K in Texas between 1988 and 1994. Relative to Head Start, which had a 6 percent share of age-eligible enrollment in 1988, pre-K was more extensive, with a 14 percent share. This figure also shows that

---

[15] The at-risk population includes the following: children unable to speak and comprehend the English language; children certified for the FRL program; children who 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 (TEC) §29.1531.).

[16] Andrews, Jargowsky and Kuhne (2012) evaluate this program and show it has been effective at improving math and reading test scores, reducing the likelihood of being retained in a grade, and decreasing the probability of receiving special education services.

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

the timing of the expansions of both programs coincide. This is a concern, as the existence of a large-scale pre-K is a potential threat to identification as a confounding factor. Hence, it is essential to control for the availability of other preschool alternatives, as failing to control for it could misattribute the effects of other preschools to Head Start.

To diagnose and address this concern, I directly analyze the relationship between Head Start spending and pre-K expansions between 1988 and 1994. Table 2, Column 5 shows that, after controlling for county and year fixed effects, there is no significant relationship between a share of age- and income-eligible children enrolled in pre-K and Head Start funding per child. If anything, this analysis suggests that Head Start program expansions may have crowded out the enrollment for pre-K for low-income children in Texas.

Although this analysis confirms that the variation in Head Start per child does not significantly predict pre-K enrollment, there is still a possibility that, in terms of facility usage and program operations, pre-K and Head Start cooperated in the 1990s.[18] 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 Head Start and show that the results are not sensitive to adding this control.

V. Results

Having documented the plausible exogeneity of the Head Start funding variation, I now present the results of Equation 1. As noted above, the main sample is restricted to FRL-certified students, and Head Start exposure is assigned at the time when a child was four years old.[19] To simplify the interpretation of the coefficient of interest, Head Start spending per child is scaled by $500, the equivalence of mean spending over the study's period. Thus, the coefficients should be interpreted as the effect of exposure to a $500 increase in Head Start spending per child. All monetary values are converted to 2014 dollars, and test scores are standardized using the entire student population.

A. Federal Head Start Funding and Test Scores

Panel A of Table 3 shows the effect of Head Start spending per child on combined test scores

[18] Starting in 2003, the state law requires the new pre-K establishments to coordinate and cooperate with Head Start (TEC §29.1533; TEC §29.158).

[19] Additional results are presented with Head Start exposure at ages three through eight in Section VII.

in math and reading. Column 1 reports the results for the main sample. The coefficients in Panel A indicate that a \$500 increase in Head Start spending per child leads to a statistically significant 0.031 standard deviation (σ) increase in third-grade test scores.[20]

Previous studies show that the returns to public education investments are higher for groups in the lower end of the skill distribution (for example, Bitler, Hoynes and Domina, 2014). Given that male students tend to be lower-achieving relative to females, Head Start may yield larger returns for males compared to females. In the next two columns of Table 3, I analyze whether exposure to more generous Head Start funding improves third-grade test scores differentially by sex. In Columns 2 and 3, additional Head Start funding exposure is associated with improvements in test scores for males and females, with slightly larger point estimates for males. However, the estimated coefficients are not statistically significant and statistically different from one another.

In the rest of the table, I present the estimates for non-Hispanic whites, blacks, and Hispanics, respectively. The race and ethnicity breakdown reveals that improvements for Hispanics are the main driver of the results. In particular, I find that a \$500 increase in Head Start spending per child leads to a 0.051σ increase in test scores in math and reading ($p<0.01$). These results suggest that around a \$500 increase in federal Head Start funding exposure closes more than 13 percent of the gap relative to the raw mean difference in test scores (= 0.051/0.38).[21] This estimate is similar to Currie and Thomas (1999), who find that Head Start participation closes at least 25 percent of the test score gap between Hispanics and whites.

Although the results for Hispanics are positive, large, and statistically significant, the analogous estimates for whites suggest that these cohorts did not experience improvements in test scores with exposure to additional Head Start funding. Also, the effects for blacks are positive and economically significant but statistically imprecise.[22]

[20] I present the results on combined test scores in math and reading for simplicity. To provide additional insight, in Appendix Table A.5, I examine the effect of Head Start in math and reading test scores separately and show that the estimated coefficients are larger and more statistically precise for math scores (a 0.045σ increase, $p<0.05$) relative to reading scores (a 0.023σ increase, $p<0.1$).

[21] Using Table 3, the average standardized test score for Hispanics is -0.449 and -0.069 for whites. The raw difference is 0.38.

[22] It is worth noting that the magnitude of the estimated effect on blacks is not negligible. Considering the enrollment increases for blacks, the estimated TOT is larger than the TOT for Hispanics.

A.1 Why Does Head Start Improve Test Scores for Hispanic Students?

As noted in Section B, Hispanic students lag behind both black and white students academically. In contrast to black children, who are historically underprivileged compared to whites, Hispanics often live in immigrant, Spanish-speaking families and communities (Currie and Thomas, 1999). During my study period, Hispanics make up more than 50 percent of FRL-certified third-grade students in Texas, and 37 percent have limited language proficiency. Table 1 reports that Hispanics were living in areas with more generous Head Start funding compared to whites and blacks. For example, average Head Start funding exposure for Hispanics was \$687, with a standard deviation of 1178; by contrast, funding exposure for blacks was approximately \$300, with a standard deviation of 211. Additionally, counties with a high fraction of Hispanics experienced more funding expansions relative to communities with a high fraction of blacks, as shown in Appendix Table A.3.

There are at least four channels by which Head Start could be beneficial for Hispanics. First, considering that Hispanics have higher participation in Head Start in Texas (11.4 percent Hispanic enrollment as a share of all poor children) compared to whites (5 percent) and blacks (4.7 percent), additional funding may induce more Hispanic children into the program. Table 2 presents enrollment effects of Head Start funding expansions among whites, blacks, and Hispanics separately. Indeed, the enrollment effects of additional Head Start funding are larger for Hispanics (6.5pp increase, $p<0.01$)[23] relative to whites (2.8pp increase, $p<0.01$) and blacks (1.5pp increase, $p<0.01$).

Second, Head Start could increase their exposure to English and develop their language skills early, which could influence their educational performance in both reading and math. Panel B of Table 3 reports that additional Head Start funding increases the likelihood of their language proficiency (0.01pp increase, $p<0.05$). One might expect that improved language proficiency would improve reading skills more than math skills. Given that the results on math scores (0.070σ) are larger than reading (0.037σ), this channel alone does not explain the result's overall

[23] Note that "pp" stands for percentage point throughout the paper.

pattern (see Appendix Table A.5).

Third, Head Start could help special needs children during early childhood by providing services to identify and potentially treat some types of special education placements such as behavior disorders or learning disabilities. This channel is plausible because the program has been required by federal law to reserve 10 percent of its enrollment for children with disabilities. Panel C of Table 3 shows that for Hispanic students, additional funding reduces the likelihood of having special education status in the third grade (0.01pp reduction, $p<0.1$).[24]

Finally, although not statistically testable, Head Start could promote cultural assimilation for Hispanics that may help children adapt to school more easily (Currie and Thomas, 1999; Bitler, Hoynes and Domina, 2014).

B. Do Test Score Effects Persist over Time?

Test score fade-out has been at the heart of the Head Start policy debate. As evidence against the program's effectiveness, critics point out that test score gains at the time of school entry dissipate quickly. In this section, I provide suggestive evidence on the effect of the Head Start program on test scores from the third through eighth grade. The main limitation here is that this analysis relies on cross-sectional data because I observe students in each grade who took standardized tests in Texas between 1994 and 1999.[25] This limits the scope of the findings, mainly because I am not able to distinguish between the nonlinearity of the effects in test scores over the years (the impact in the third grade is expected to decrease as students age) and the reduction in exposure to funding generosity across the years (older cohorts experience less exposure to Head Start relative to younger cohorts).

With this limitation in mind, Figure 4 presents estimated coefficients and their 95 percent

[24] Due to the lack of detailed data, I am not able to explore the effect of Head Start separately across categories of disability. With the detailed data from North Carolina (NC), Muschkin, Ladd and Dodge (2015) examine the effects of two early childhood education (ECE) initiatives, which occurred in the 1990s and early 2000s in NC, on the probability that children are placed into special education by the end of the third grade (the outcome I explore in my paper) and are placed in specific disability categories. Taking advantage of the rollout of two ECE programs' introductions across counties and over time, they show that the programs' introduction and expansion reduced special education replacement in the third grade (similar to my findings). They also find that one of the two programs reduced placements for educable handicaps, specific learning disabilities, and other health impairments, providing some evidence that supports the main argument.

[25] Unfortunately, I do not have access to panel data that would allow me to follow the same students across years.

confidence intervals from separate regressions with the standardized test scores combined, in math and in reading for each grade as the outcome variables and exposure to Head Start funding as the right-hand-side variable.[26] Panel A of Figure 4 shows the evolution of the results when the sample includes all FRL-certified students, and Panel B shows the estimated coefficients when the sample is restricted to Hispanic students who are FRL certified.

Overall, the results suggest that the effects on test scores appear to be positive throughout, but the effect sizes decline after the fifth grade. Although not conclusive, this pattern in Panel A is consistent with the results of previous literature (Currie and Thomas, 1995; Krueger and Whitmore, 2001; Deming, 2009).

In Panel B, for Hispanic students, the group with the highest returns, there is suggestive evidence of more persistence of test score gains in comparison to Panel A, especially in reading. The estimated effect declines from 0.051σ in the third grade to 0.036σ in the fifth grade in combined test scores. Although the point estimates in combined test scores are smaller, they are still sizable and statistically significant in the fifth grade. In reading test scores, the effect size remains around 0.04σ through sixth grade. There could be two explanations for this pattern. First, as presented in Section A.1, there are several reasons why Hispanics benefit from early childhood investments, which may explain more persistent effects on test scores among this group. Prior literature also shows supporting evidence that early childhood investments are more effective for Hispanics (Gibbs, 2014). Second, the results for Hispanics may take a longer time to fade out because the initial impact size for this group is larger.

In sum, this analysis suggests that for Hispanics, Head Start investments improved test scores from the third through fifth grade, reading test scores persist through the sixth grade, and the fade-out pattern takes a longer time to occur.

C. Interpretation and Magnitude of Estimates

The estimates reported in Table 3 represent ITT effects, which can be interpreted as the average effect of Head Start funding exposure on economically disadvantaged children. My main

[26] Appendix Table A.6 presents these results in a table form.

sample consists of FRL-certified students. These students would have been eligible for Head Start as children if their families' socioeconomic status had stayed approximately the same from early childhood to age nine.[27]

To make accurate comparisons with the literature, I attempt to convert the ITT estimate of Head Start funding on test scores to the TOT. This conversion requires having a "first-stage" that provides an estimate of exposure to Head Start funding on the likelihood of participating in Head Start for the sample. Unfortunately, Head Start participation is not observed in the main dataset. As a proxy, I use the estimated effect from Column (1) of Table 2, which reports a \$500 increase in funding is associated with an 11.2pp increase in Head Start enrollment at the county level for low-income children in Texas.[28]

Scaling up the test score impact of a 0.031σ by 0.112, the implied effect of Head Start enrollment at the county level corresponds to a 0.28σ increase in test scores. Although imperfect, this estimate provides a comparison point with the literature. However, this should not be interpreted as the "true" effect of participating in Head Start, given the issues mentioned above.

Comparing the test score impact of Head Start from my paper with Deming (2009) and the HSIS (Puma et al., 2012), my study reports a larger and statistically significant impact on third-grade test scores. My analysis and the other two papers differ in terms of the study period, the sample, and identification strategy, potentially explaining the larger findings here. While Deming (2009) studies the Head Start program in the 1980s (pre-1990) and Puma et al. (2012) examines the random assignment in the 2000s, my paper focuses on the 1990s, during which Head Start evolved significantly in terms of capacity and quality. Moreover, the other two papers analyze nationally representative samples, but my sample consists of low-income children in Texas who are predominantly Hispanic. Last, my identification strategy differs from the other two papers by taking advantage of the funding expansions across counties and over time.

[27] However, eligibility does not imply participation. Head Start only serves a proportion of the eligible children in a given year due to capacity constraints. There is also an issue with incomplete take-up, meaning that not all eligible children actually enroll in the program.

[28] This exercise requires a strong assumption that funding increases translated into "only" enrollment expansion. However, in the 1990s, additional funding was appropriated toward quality improvements, as described in detail in Section A. As a result, TOT here could be interpreted as an upper bound of the effect of Head Start participation.

VI. Discussion of Mechanisms

One channel by which Head Start funding increases could improve test scores is by serving more children. As expressed above, Table 2 reports the results that show that additional funding led to significant increases in Head Start enrollment for economically disadvantaged children in Texas.

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

To explore this, I examine the relationship between federal funding increases and program inputs such as number of teachers per child, enrollment per child, full-time enrollment per child, child-teacher ratio, and director's salary[29] as well as per child spending for education, health, nutrition, social services, and parent involvement in the programs. For this analysis, I employ data on program characteristics from the PIRs (available for 1988-1994), director's salary from the PIRs (available for 1992-1994), and program budgets for various types of spending available in an administrative data set on Head Start budgets called PCCOST (available for 1992-1994 for some programs).

Table 4, Panel A reports the findings on program inputs. The first three columns show that a $500 increase in Head Start funding per child is associated with significant increases in the number of teachers per child, enrollment per child, and full-time enrollment per child. This finding is

[29] Appendix Figure A.6 plots the average of some of these inputs from 1988 to 1995.

consistent with Walters (2015). In Column 4, the coefficient on the child-teacher ratio is negative as expected; however, it is not statistically significant. Finally, the results in Column 5 suggest that part of the increase in funding translated into increases in directors' salary.

Panel B of Table 4 shows that a $500 increase in Head Start funding per child is associated with increases in spending on education, health, nutrition, social services, and parent involvement. Given that budget data are only available for a subset of programs, these results are suggestive at best. Overall, there is suggestive evidence that on the margin, federal funding partially went to spending for services that might improve education and health development for children.

Since the 1990s, Head Start has aimed to increase teacher qualifications by setting aside funds to improve teachers' education. Although not testable, quality improvements may have been partly driven by the advances in teacher qualifications.[30]

VII. Robustness

In this section, I conduct various robustness exercises to address potential concerns with the estimation strategy. Table 5 presents multiple sensitivity checks to the main specification. As a point of comparison, I include the baseline estimates in Column 1 that show the effect of Head Start exposure on third-grade test scores. In Column 2, I omit county-specific linear trends, which leads to a small increase in the size of the effect. Column 3 presents results with no pre-K controls, which slightly decreases the magnitude of the effect.

In Column 4, I exclude the controls for average county-level income per capita at the time of birth, at the time of Head Start, and test year. With this restriction, the estimated coefficient remains significant, but the magnitude decreases slightly. In the next column, I omit controls for the county-level measures of per capita transfer payments for cash income support, Food Stamps, medical care, retirement, and disability programs and find that the estimate's magnitude decreases without these controls. This is expected because Head Start funding expansions are positively correlated with generosity of transfer payments. Excluding these critical controls biases the main effect downward.

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

In the last column, I include school-specific linear trends to control for possible improvements in students' neighborhoods. Adding these trends leads to smaller coefficients with similar standard errors, which means that school trends may be correlated with Head Start funding changes. Because the overall effect is relatively sensitive to adding school trends, I replicate the main results adding school trends. Appendix Table A.7 presents these estimates, which show that the effect for Hispanics remains statistically significant, though the effect size is slightly smaller.

Next, I analyze whether the results are sensitive to two additional Head Start exposure measures constructed using alternative population denominators and find that the results are consistent across all three measures. Column 1 of Appendix Table A.8 reports the estimates using my preferred measure, Head Start per child, and in the next two columns, I present results obtained with Head Start per capita (Column 2) and Head Start per poor child (Column 3).[31] Regardless of the differences in the measures, this table shows that an increase in Head Start funding, on average, leads to similar test score gains.[32]

As a falsification exercise, I analyze the effect of exposure to Head Start on test scores for the sample of children who are *not* FRL certified or living under poverty; hence they are not likely to benefit from Head Start. I present these results in Appendix Table A.10 and show that Head Start exposure does not affect test scores for this group.

To check for potential selection effects, I test whether increases in federal Head Start spending affect the composition of a particular grade within a school (similar to Carrell and Hoekstra, 2010). If the variation in Head Start spending is not correlated with selection into the sample, I would expect to find no correlation. Appendix Table A.1 presents regression results obtained by regressing exogenous student characteristics on Head Start exposure conditional on birth year, test year and school fixed effects, and other controls included in the main specification. The

[31] To construct age-specific poor population counts at the county-year level, I use the Small Area Income and Poverty Estimates (SAIPE) of the U.S. Census Bureau, for years 1989, 1993, 1995, 1997, 1998 and 1999. These data report estimated counts for the number of children aged 0-17 and children aged 5-17. Taking the difference between these two variables, I construct the number of children younger than age five. The number of poor children age three or four is two-fifth of the number of children younger than age five (Frisvold, 2006). For more details about the data construction, see Appendix II.

[32] In Appendix Table A.9, I show the results using test scores by sex and by ethnicity using Head Start per capita in Panel A and Head Start per poor child in Panel B. These results closely follow the patterns of the estimates in Table 3.

results suggest no evidence that a particular cohort's composition (probability of being a specific race or sex, likelihood of being FRL certified, and county's income composition) is correlated with Head Start funding. In the last column, I test whether Head Start funding increases are associated with the "predicted test scores" using students' observable characteristics and find no association between them. Overall, these tests suggest that the main results are not driven by selection bias.

Because Head Start serves children between the ages of three and five, one would not expect exposure to Head Start that occurred during other ages to be associated with improvements in outcomes (Thompson, 2018). As a falsification test, I estimate models where I analyze Head Start exposure at different ages.[33] Table 6 reports results from specifications that use the sample of third graders with the assignment of exposure to Head Start changing from age three through age eight as the right-hand-side variable.[34] These findings suggest that Head Start exposure at ages three and four matters the most, with the largest impact on test scores coming from exposure at age four. This is expected, considering that four-year-old children make up around 50 percent of the total children served in Head Start (PIR).

In sum, this section presents evidence that the results are robust to excluding county-specific linear trends, do not appear to be the result of the expansion of the alternative early childhood programs during this time period, and are not driven by selection bias. The results also show that Head Start funding exposure at age four leads to the largest improvements in third-grade test scores, and exposure at ages older than age five does not affect test scores.

VIII. Conclusion

This paper provides new evidence on the old debate of whether early childhood investments for economically disadvantaged children narrow the academic achievement gap in elementary school. Many early childhood interventions are effective in improving children's skill development

[33] Here exposure to Head Start is assigned based on ages from three through eight separately. For example, for a child who is born in 1985, their age of exposure as a three-year-old would be in 1988, a four-year-old in 1989, and a five-year-old in 1990.

[34] Appendix Figure A.7 visually represents the coefficients for each age of exposure with 95 percent confident intervals.

initially, yet the effects disappear when students reach the third grade (Bailey et al., 2017). Using new variation in the federal funding expansions for the Head Start program in the 1990s and student-level administrative data from Texas, I find that exposure to more generous Head Start funding during childhood led to substantial improvements in test scores.

The gains of low-income Hispanic students drive the overall improvements in test scores. In particular, a $500 increase in Head Start funding per child led to a 13 percent reduction in the test score gap in math and reading combined between Hispanics and whites. Hispanics benefited from funding expansions through increased access to Head Start and improvements in program inputs. These advances enhanced their language proficiency and reduced their likelihood of special education needs during elementary school.

Head Start has served low-income children for more than 50 years to reduce education and health disparities across socioeconomic groups. During the 1990s, the federal government expanded Head Start funding to improve both its capacity and quality. The significant national funding increase within a relatively short time created a natural experiment that resulted in a large variation of the adoption of funding expansions across counties and over time. Combining several sources on program characteristics and budgets, my findings provide new evidence on the ways public funds are spent and suggest that both program capacity and quality improvements are essential pathways for the ultimate effects on test scores.

Early childhood investments have been receiving significant political attention. Therefore, a comprehensive understanding of the potential benefits and costs of such investments is essential for informed policymaking. My findings indicate that Head Start passes the cost-benefit test by a wide margin and provides useful insight for policymakers considering future public investments in early childhood education.

References

Andrews, Rodney J, Paul Jargowsky, and Kristin Kuhne. 2012. "The Effects of Texas's Targeted Pre-Kindergarten Program on Academic Performance." National Bureau of Economic Research NBER Working Paper No. 18598.

Bailey, Drew, Greg J Duncan, Candice L Odgers, and Winnie Yu. 2017. "Persistence and Fade-out in the Impacts of Child and Adolescent Interventions." *Journal of Research on Educational Effectiveness*, 10(1): 7–39.

Bailey, Martha J, Shuqiao Sun, and Brenden Timpe. 2020. "Prep School for Poor Kids: The Long-Run Impacts of Head Start on Human Capital and Economic Self-Sufficiency." National Bureau of Economic Research NBER Working Paper No. 28268.

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

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

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

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

Barr, Andrew C., and Chloe Gibbs. 2019. "Breaking the Cycle? Intergenerational Effects of an Anti-Poverty Program in Early Childhood." Annenberg Institute at Brown University 141.

Bauer, Lauren, and Diane Schanzenbach. 2016. "The Long-Term Impact of the Head Start Program." Hamilton Project.

Bitler, Marianne P, Hilary W Hoynes, and Thurston Domina. 2014. "Experimental Evidence on Distributional Effects of Head Start." National Bureau of Economic Research.

Blau, David, and Janet Currie. 2006. "Pre-school, Day Care, and After-school Care: Who's Minding the Kids?" *Handbook of the Economics of Education*, 2: 1163–1278.

Carneiro, Pedro, and Rita Ginja. 2014. "Long-Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start." *American Economic Journal: Economic Policy*.

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

Consolidated Federal Funds Report, Fiscal Years 1982-2007: County Areas [CFFR-dataset]. Washington D. C.: Bureau of the Census and U. S. Department of Commerce. Retrieved from National Archives Catalog https://catalog.archives.gov/id/626196

Consolidated Federal Funds Report, Fiscal Years 1983-2010: County Areas [CFFR]. Washington D. C.: Bureau of the Census and U. S. Department of Commerce. Reports https://www.census.gov/library/publications/time-series/cffr.All.html

Community Services Act, Pub. L. No. 93-644, H.R. 14449, 93rd Cong. 1975 [Community Services Act of 1974] . Retrieved from https://www.congress.gov/bill/93rd-congress/house-bill/14449/all-info

Currie, Janet. 2001. "Early Childhood Education Programs." *Journal of Economic Perspectives*, 213–238.

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

Currie, Janet, and Duncan Thomas. 1999. "Does Head Start Help Hispanic Children?" *Journal of Public Economics*, 74(2): 235–262.

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

Deming, David. 2009. "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start." *American Economic Journal: Applied Economics*, 1(3): 111–34.

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

Duncan, Greg J, and Richard J Murnane. 2011. *Whither Opportunity?: Rising Inequality, Schools, and Children's Life Chances.* Russell Sage Foundation.

Feller, Avi, Todd Grindal, Luke W Miratrix, and Lindsay Page. 2016. "Compared to What? Variation in the Impact of Early Childhood Education by Alternative Care-type Settings." *Annals of Applied Statistics*.

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

Frisvold, David E. 2006. "Head Start Participation and Childhood Obesity." *Vanderbilt University Economics Working Paper No. 06-WG01*.

Garces, Eliana, Duncan Thomas, and Janet Currie. 2002. "Longer-Term Effects of Head Start." *American Economic Review*, 92(4): pp. 999–1012.

Gibbs, Chloe, Jens Ludwig, and Douglas L Miller. 2013. "Head Start Origins and Impacts." *Legacies of the War on Poverty*, 39–65.

Gibbs, Chloe R. 2014. "Experimental Evidence on Early Intervention: The Impact of Full-day Kindergarten." *Frank Batten School of Leadership and Public Policy Working Paper*, 4.

Head Start Act §641a, 42 USC §9801, Pub L No. 110-134. 2007. HHS Administration for Children and Families. Retrieved from https://eclkc.ohs.acf.hhs.gov/policy/head-start-act/sec-644-administrative-requirements-standards.

Heckman, James J. 2006. "Skill Formation and the Economics of Investing in Disadvantaged Children." *Science*, 312(5782): 1900–1902.

Hoynes, Hilary Williamson, and Diane Whitmore Schanzenbach. 2012. "Work Incentives and the Food Stamp Program." *Journal of Public Economics*, 96(1): 151–162.

Jackson, C. Kirabo. 2018. "Does School Spending Matter? The New Literature on an Old Question." National Bureau of Economic Research Working Paper.

Johnson, Rucker C, and C Kirabo Jackson. 2019. "Reducing Inequality Through Dynamic Complementarity: Evidence from Head Start and Public School Spending." *American Economic Journal: Economic Policy*, 11(4): 310–49.

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

Krueger, Alan B, and Diane M Whitmore. 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.

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

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

Muschkin, Clara G, Helen F Ladd, and Kenneth A Dodge. 2015. "Impact of North Carolina's Early Childhood Initiatives on Special Education Placements in Third Grade." *Educational Evaluation and Policy Analysis*, 37(4): 478–500.

National Cancer Institute. 2017. *Download U. S. Population Data- 1969-2016* [SEER - dataset]. Washington D. C.: National Institutes of Health. Retrieved from http://www.nber.org/data/seer_u.s._county_population_data.html

Office of Head Start Program Information Report. Office of Head Start [PIR – dataset]. Retrieved from http://eclkc.ohs.acf.hhs.gov/PIR.

Puma, Michael, Stephen Bell, Ronna Cook, Camilla Heid, and Michael Lopez. 2005. "Head Start Impact Study: First Year Findings." *Administration for Children & Families*.

Puma, Michael, Stephen Bell, Ronna Cook, Camilla Heid, Gary Shapiro, Pam Broene, Frank Jenkins, Philip Fletcher, Liz Quinn, Janet Friedman, et al. 2010. "Head Start Impact Study: Final Report." *Administration for Children & Families*.

Puma, Mike, Stephen Bell, Ronna Cook, Camilla Heid, Pam Broene, Frank Jenkins, Andrew Mashburn, and Jason Downer. 2012. "Third Grade Follow-Up to the Head Start Impact Study: Final Report. OPRE Report 2012-45." *Administration for Children & Families*.

Reardon, Sean F, Joseph P. Robinson, and Ericka S Weathers. 2008. "Patterns and trends in racial/ethnic and socioeconomic academic achievement gaps." *Handbook of Research in Education Finance and Policy*.

Sanders, Nicholas 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, Hilary M, Holly S Schindler, Katherine A Magnuson, Greg J Duncan, Hirokazu Yoshikawa, and Cassandra MD Hart. 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.

Texas Education Agency Prekindergarten Programs. Texas Education Agency 2014. Retrieved from: https://www.lbb.state.tx.us/Documents/Publications/Issue_Briefs/1383_TEA_Prek_Programs.pdf

Texas Education Code 29.158. Education Code Title 2. Public Education. Chapter 29. Educational Programs. Retrieved from https://statutes.capitol.texas.gov/Docs/ED/htm/ED.29.htm#29.158

Texas Education Code 29.1531. Education Code Title 2. Public Education. Chapter 29. Educational Programs. Retrieved from https://statutes.capitol.texas.gov/Docs/ED/htm/ED.29.htm#29.1531

Texas Education Code 29.1532. Education Code Title 2. Public Education. Chapter 29. Educational Programs. Retrieved from https://statutes.capitol.texas.gov/Docs/ED/htm/ED.29.htm#29.1532

Texas Education Code 29.1533. Education Code Title 2. Public Education. Chapter 29. Educational Programs. Retrieved from https://statutes.capitol.texas.gov/Docs/ED/htm/ED.29.htm#29.1533

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

U.S. Administration for Children Youth and Families.1990. Head Start Fact Sheet. Washington, DC: Head Start Bureau.

U.S. Census Bureau, Small Area Income and Poverty Estimate [SAIPE – dataset]. Retrieved from https://www.census.gov/programs-surveys/saipe/data/datasets.html

U.S. Department of Commerce. Bureau of Economic Analysis. Regional Economic Information System [REIS – dataset]. Retrieved from https://www.bea.gov/data/economic-accounts/regional

U.S. Department of Education, National Center for Education Statistics, Common Core of Data [CCD – dataset]. Retrieved from https://nces.ed.gov/ccd/pubschuniv.asp

U.S. Department of Health and Human Services (HHS), Administration for Children and Families, Head Start Bureau. Head Start. 1999 [PCCOST – dataset]. PCCost: The computerized version of the grant application package SF424A version 4.01 Y2K.

United States. Bureau of the Census. County and City Data Book [United States], 1983. Inter-university Consortium for Political and Social Research [distributor], 2008-06-18. https://doi.org/10.3886/ICPSR08256.v1

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

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

IX. Tables

Table 1: Sample Characteristics of Third Grade Students in Texas

	Full Sample	Not Disav. Sample	Free/Reduced Lunch Eligible Sample					
			All	Male	Female	White	Black	Hispanic
Head Start								
Head Start Funding per Child	402.66	312.52	518.47	518.98	517.96	335.79	298.95	687.47
	(690.86)	(432.03)	(909.30)	(905.40)	(913.14)	(361.80)	(210.60)	(1177.93)
Outcomes								
Reading and Math Composite Score	0.00	0.31	-0.36	-0.46	-0.25	-0.07	-0.38	-0.45
	(1.00)	(0.77)	(1.11)	(1.16)	(1.05)	(0.99)	(1.04)	(1.15)
Standardized Reading Score	0.00	0.30	-0.34	-0.47	-0.21	-0.08	-0.33	-0.43
	(1.00)	(0.78)	(1.11)	(1.16)	(1.04)	(1.02)	(1.04)	(1.14)
Standardized Math Score	0.00	0.29	-0.33	-0.40	-0.27	-0.05	-0.39	-0.41
	(1.00)	(0.77)	(1.11)	(1.16)	(1.06)	(0.97)	(1.03)	(1.17)
Limited English Proficiency	0.10	0.02	0.21	0.22	0.20	0.01	0.00	0.37
	(0.30)	(0.14)	(0.41)	(0.41)	(0.40)	(0.09)	(0.07)	(0.48)
Participates in a Special Education Program	0.09	0.07	0.11	0.15	0.07	0.15	0.12	0.10
	(0.29)	(0.26)	(0.32)	(0.36)	(0.26)	(0.35)	(0.32)	(0.30)
Controls								
Year of Birth	1987.19	1987.08	1987.32	1987.33	1987.30	1987.27	1987.34	1987.30
	(2.06)	(2.05)	(2.06)	(2.05)	(2.07)	(2.08)	(2.04)	(2.06)
Female	0.50	0.49	0.50	0.00	1.00	0.50	0.51	0.50
	(0.50)	(0.50)	(0.50)	(0.00)	(0.00)	(0.50)	(0.50)	(0.50)
White, not of Hispanic Origin	0.53	0.78	0.22	0.22	0.22	1.00	0.00	0.00
	(0.50)	(0.41)	(0.41)	(0.41)	(0.41)	(0.00)	(0.00)	(0.00)
African American	0.14	0.07	0.23	0.22	0.23	0.00	1.00	0.00
	(0.35)	(0.26)	(0.42)	(0.42)	(0.42)	(0.00)	(0.00)	(0.00)
Hispanic	0.31	0.12	0.54	0.55	0.54	0.00	0.00	1.00
	(0.46)	(0.33)	(0.50)	(0.50)	(0.50)	(0.00)	(0.00)	(0.00)
Free/Reduced Lunch Eligible	0.44	0.00	1.00	1.00	1.00	1.00	1.00	1.00
	(0.50)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Bilingual	0.06	0.01	0.13	0.14	0.13	0.00	0.00	0.24
	(0.24)	(0.07)	(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
	(0.17)	(0.10)	(0.23)	(0.24)	(0.23)	(0.07)	(0.06)	(0.29)
N	742660	407349	335311	166684	168627	72377	76436	182021

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

Table 2: The Effect of Head Start Funding per Child on Head Start and Pre-K Enrollment

	Head Start Enrollment				State Pre-K Enrollment
	Total	Whites	Blacks	Hispanics	Total
Head Start Funding	0.112***	0.028***	0.015***	0.065***	-0.013
	(0.023)	(0.009)	(0.004)	(0.018)	(0.014)
Mean Y	0.210	0.050	0.047	0.114	0.524
Obs	1771	1771	1771	1771	1771

Notes: This table reports the estimated coefficients for Head Start and pre-K enrollment per poor population of children 3 and 4 in Texas when the independent variable is real federal Head Start (HS) spending per child (2014$). HS spending per child is scaled by $500 (average spending during the period of study), thus the coefficients should be interpreted as the effect of a $500 increase in funding. HS 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. Head Start enrollment data are from the Program Information Reports (PIR). Total HS enrollment data are available in all years. The enrollment data by race/ethnicity are available in 1988 and between 1992-1994, the missing data in between are linearly interpolated. Pre-K enrollment data are from the Common Core Data (CCD). Poor children 3-4 counts are estimated using the SAIPE data. For more details about the data description, see Appendix II. The data are at the county-level and include years between 1988 and 1994. All regressions include county, year fixed effects and 1980 county characteristics interacted with linear trends. Standard errors are clustered at the community level. * p<0.10, ** p<0.05, *** p<0.01.

Table 3: Baseline Estimates of the Effect of Head Start Funding Exposure on Student Outcomes in Third Grade

	All	Males	Females	Whites	Blacks	Hispanics
	(1)	(2)	(3)	(4)	(5)	(6)
A: Reading and Math Composite Score						
Head Start Funding	0.031**	0.036	0.028	-0.013	0.053	0.051***
	(0.013)	(0.023)	(0.018)	(0.044)	(0.043)	(0.014)
Mean Y	-0.355	-0.459	-0.253	-0.069	-0.382	-0.449
B: Language Profeciency						
Head Start Funding	0.008**	0.009**	0.007***	-0.000	-0.004	0.010**
	(0.003)	(0.005)	(0.002)	(0.001)	(0.003)	(0.004)
Mean Y	0.789	0.783	0.795	0.992	0.995	0.631
C: Special Education						
Head Start Funding	-0.006	-0.008	-0.007*	-0.001	-0.032**	-0.010*
	(0.004)	(0.006)	(0.004)	(0.012)	(0.013)	(0.005)
Mean Y	0.112	0.150	0.074	0.145	0.117	0.097
Obs	335311	166684	168627	72377	76436	182021

Notes: This table contains results obtained when the independent variable is real federal Head Start (HS) spending per child (2014$) when the child was four years old. HS spending per child is scaled by $500 (average spending during the period of study), thus the coefficients should be interpreted as the effect of a $500 increase in funding. All regressions include controls for demographics and county-level characteristics, school, test year and birth year fixed effects, along with county-specific linear trends. Sample consists of third-grade students who are FRL certified or who are identified as economically disadvantaged, based on the description by the Texas Education Agency (TEA). Student data are from the TEA which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 and 1999. HS spending data are from the Consolidated Federal Files Reports (CFFR) and include years between 1988 and 1994. Standard errors are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.

Table 4: Mechanisms: The Relationship between Head Start Funding and Head Start Program Characteristics and Budget

	# Teachers per child	# Enrolled per child	# Full-Time Enrolled per child	Child Teacher Ratio	Director's Salary per child
	(1)	(2)	(3)	(4)	(5)
A: Program Characteristics:					
Head Start Funding per Child	0.002***	0.036***	0.018**	-1.486	2.964*
	(0.000)	(0.010)	(0.007)	(1.097)	(1.615)
Mean Y	0.003	0.077	0.029	29.060	15.465
Obs	1771	1771	1771	1339	759
	Education	Health	Nutrition	Social	Parent Involvement
	(1)	(2)	(3)	(4)	(5)
B: Program Budgets for Per Child Spending on:					
Head Start Funding per Child	69.851***	3.946	20.191**	17.132*	6.285**
	(22.267)	(4.124)	(8.426)	(9.744)	(2.339)
Mean Y	296.25	37.72	28.85	19.56	19.76
Obs	322	322	322	322	322

Notes: The independent variable is real federal Head Start (HS) spending per child (2014$), scaled by $500 (average spending during the period of study). Thus the coefficients should be interpreted as the effect of a $500 increase in funding. HS spending data are from the Consolidated Federal Funds Reports (CFFR). Program characteristics and director's salary are from the PIRs (Panel A), and budget spending breakdowns are from the PCCOST data (Panel B). Director's salary is available for the years 1992-1994. Data on budgets are only available for the years 1992-1994 for some programs. The rest of the variables are available for years 1988-1994. For more details about the data description, see Appendix II. All the monetary values in outcome variables are converted into 2014 dollars. All outcome variables (except for pupil teacher ratio) are in "per child" terms. All regressions include county and year fixed effects. In Panel B, the sample is restricted to programs that experience positive HS funding expansions. Standard errors are clustered at the community level. * p<0.10, ** p<0.05, *** p<0.01.

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

Sensitivity of Results to Alternative Specifications

	Main Results	Omit County Trends	Omit Pre-K Controls	Omit Income Controls	Omit Safety Net Controls	Add School Trends
	(1)	(2)	(3)	(4)	(5)	(6)
Head Start Funding	0.031**	0.034**	0.029**	0.027**	0.015	0.021
	(0.013)	(0.014)	(0.013)	(0.014)	(0.013)	(0.013)
Mean Y	-0.355	-0.355	-0.355	-0.355	-0.355	-0.355
Obs	335311	335311	335311	335311	335311	335311

Notes: This table reports results obtained when the dependent variable is third grade combined standardized test scores in math and reading. The independent variable is real federal Head Start (HS) spending per child (in 2014$) when the child was four years old. HS spending per child is scaled by $500 (average spending during the period of study), thus the coefficients should be interpreted as the effect of a $500 increase in funding. All regressions include controls for demographics, school, test year and birth year fixed effects. Sample consists of third-grade students who are FRL certified or who are identified as economically disadvantaged, based on the description by the Texas Education Agency (TEA). Student data are from the TEA which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 and 1999. HS pending data are from the Consolidated Federal Files Reports (CFFR) and include years between 1988 to 1994. Standard errors are clustered at the county level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

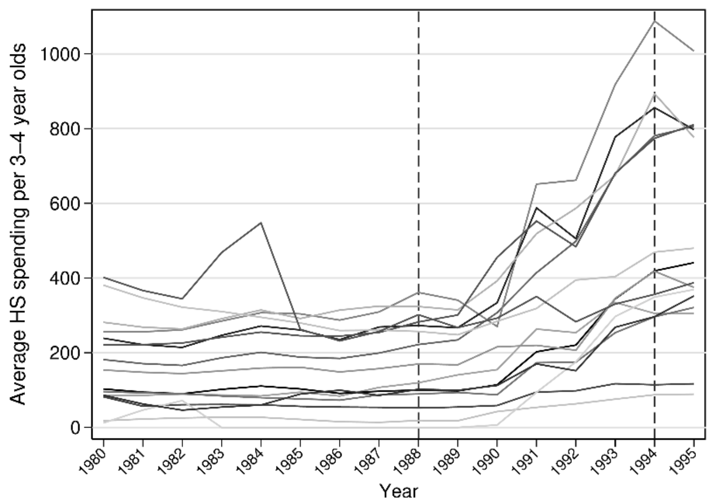
Table 6: Falsification Test: The Effect of Head Start Funding Exposure on Third Grade Combined Standardized Test Scores - Differential Effects by Age of Exposure, 3-8

	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8
Head Start Funding	0.022*	0.031**	0.020	0.007	0.003	0.009
	(0.013)	(0.013)	(0.018)	(0.016)	(0.012)	(0.012)
Mean Y	-0.355	-0.355	-0.355	-0.355	-0.355	-0.355
Obs	335311	335311	335311	335311	335311	335311

Notes: This table reports results obtained when the dependent variable is third grade combined standardized test scores in math and reading, and the independent variable is real federal Head Start (HS) spending per child (in 2014$), assigned when the child was at ages 3 to 8 years old. HS spending per child is scaled by $500 (average spending during the period of study), thus the coefficients should be interpreted as the effect of a $500 increase in funding. All regressions include controls for demographics and county-level characteristics, school, test year and birth year fixed effects, along with county-specific linear trends. Sample consists of third-grade students who are FRL certified or who are identified as economically disadvantaged, based on the description by the Texas Education Agency (TEA). Student data are from the TEA which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 and 1999. HS spending data are from the Consolidated Federal Files Reports (CFFR) and include years between 1988 and 1994. Standard errors are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.

X. Figures

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

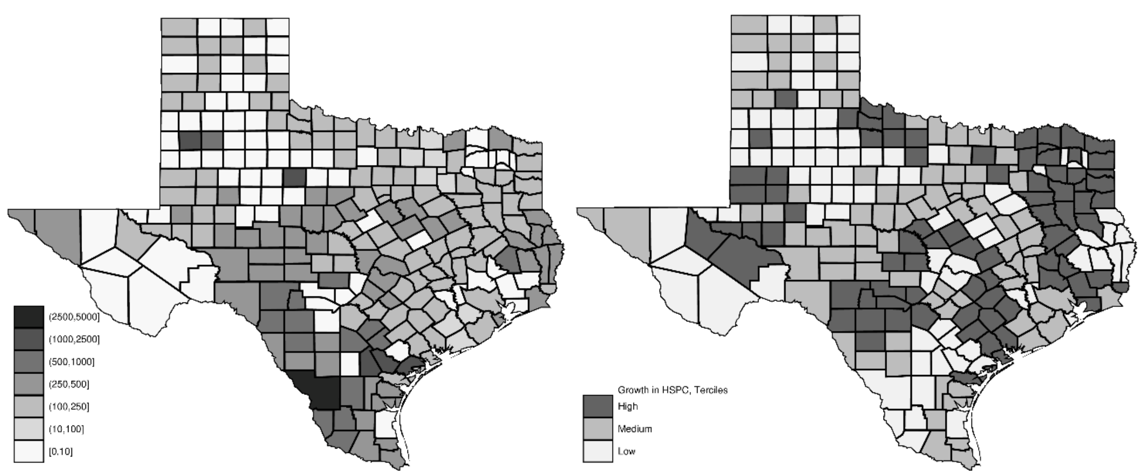


Notes: Head Start spending (in 2014$) data are from the Consolidated Federal Funds Reports (CFFR), coupled with the population counts for three- and four-year-olds at the county-level from the SEER. For more details about data construction, see Section III. 15 most populous counties include Bend, Bexar,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

Figure 2: Geographic Variation in Head Start Funding per Child

(a) Spending per child in 1988

(b) Growth in spending per child 1988-1994

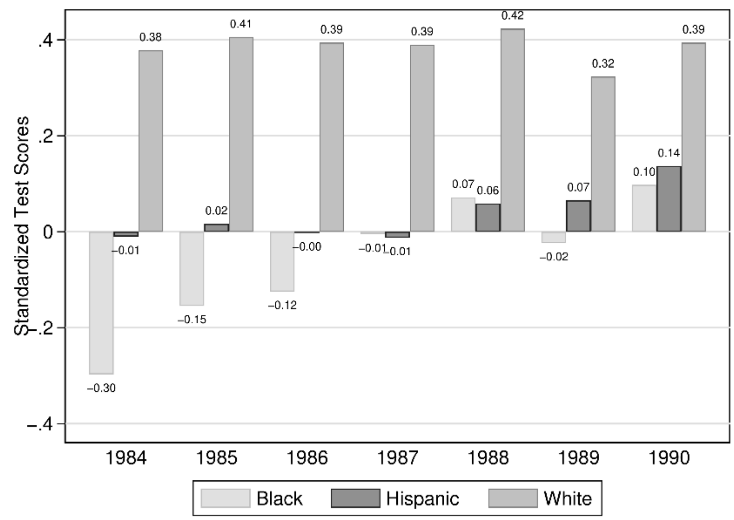


Notes: Head Start spending (in 2014$) data are from the Consolidated Federal Funds Reports (CFFR), coupled with the population counts for three- and four-year-olds at the county-level from the SEER. Growth measure is calculated using 1988 as the base period, and the shades in the map on the right are determined based on the terciles of the growth distribution. For more details about data construction, see Section III.

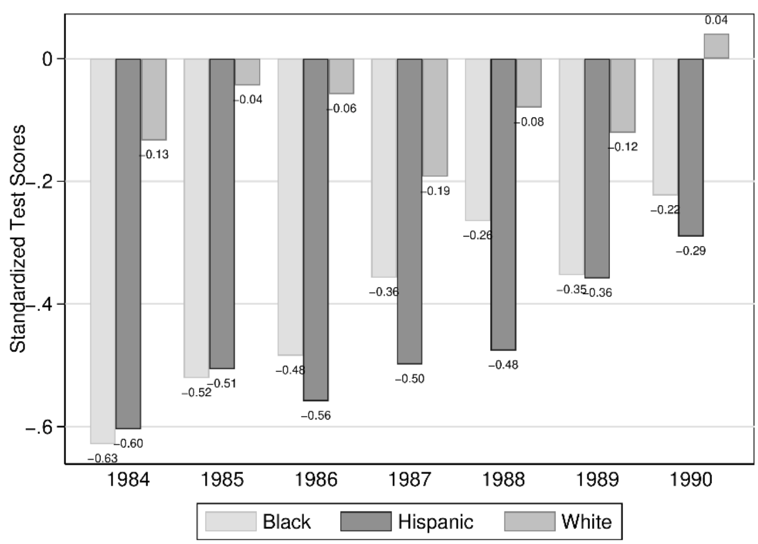
Figure 3: Histogram of Standardized Test Scores across Birth Cohorts,

by Free-or-reduced Lunch Status

(a) *Not* Free/Reduced Lunch Certified



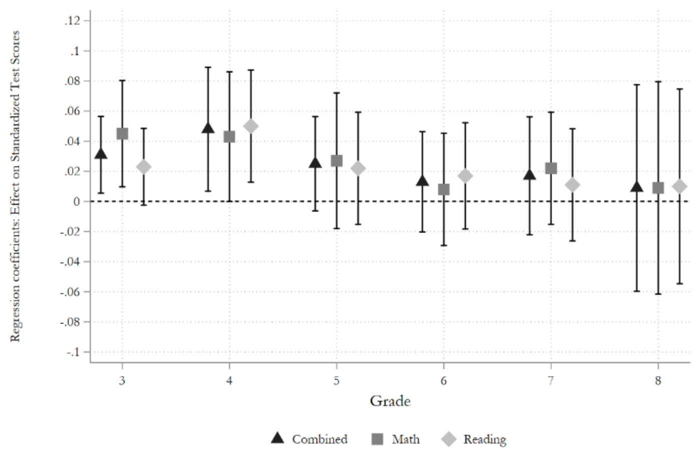
(b) Free/Reduced Lunch Certified



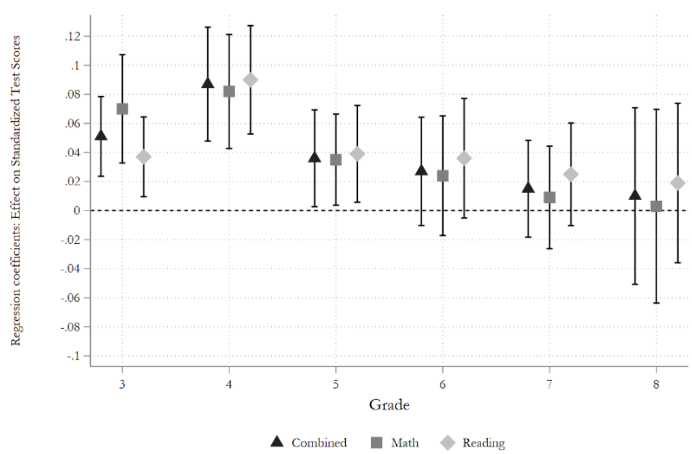
Notes: The bars indicate average test scores for black, Hispanic, and white students separately. Test scores are math and reading composite scores, standardized using all third-grade students who took the standardized test in Texas between 1994 and 1999, from the Texas Education Agency (TEA). The sample is divided into two groups: (a) students who are *not* identified as economically disadvantaged and (b) students who are FRL certified or who are identified as economically disadvantaged based on their families' welfare eligibility.

Figure 4: Effect of Head Start Exposure on Test Scores for Each Grade

(a) All Students



(b) Hispanic Students



Notes: This figure plots coefficients and their 95% confidence intervals obtained when the dependent variable is standardized test scores combined, in math, and in reading separately for each grade and the independent variable is real federal Head Start spending per child (in 2014$) when the child was four years old. Head Start spending per child is scaled by $500 (average spending during the period of study), thus the coefficients should be interpreted as the effect of a $500 increase in funding. All regressions include controls for demographics and county-level characteristics, school, test year and birth year fixed effects, along with county-specific linear trends. Sample consists of students from third through eighth grade who are FRL certified or who are identified as economically disadvantaged, 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 1994. Standard errors are clustered at the county level.

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

Esra Kose

May 1st, 2021

I. Empirical Appendix

A. Tables

Table A.1: Falsification Tests: The Effect of Head Start Exposure on Exogenous Student Characteristics, County-level Income, and Predicted Test Scores

	Black	Hispanic	Female	Free/Reduced Meal	Income PC	Predicted Test Scores
Head Start Funding	-0.001	0.001	0.002	-0.008	0.000	-0.000
	(0.001)	(0.001)	(0.006)	(0.008)	(0.001)	(0.000)
Mean Y	0.145	0.312	0.498	0.394	35164.821	0.010
Obs	742660	742660	742660	742660	742660	742660

Notes: This table reports results obtained when the dependent variables are race/ethnicity, sex, FRL status, income per capita at the county level and predicted test scores, separately. The independent variable is real federal Head Start (HS) spending per child (in 2014$) when the child was four years old. HS spending per child is scaled by $500 (average spending during the period of study), thus the coefficients should be interpreted as the effect of a $500 increase in funding. All regressions include controls for demographics and county-level characteristics, school, test year and birth year fixed effects, along with county-specific linear trends. Sample consists of all the third-grade students. Student data are from the TEA which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 and 1999. HS spending data are from the Consolidated Federal Files Reports (CFFR) and include years between 1988 and 1994. Standard errors are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.

Table A.2: Correlations between County-Level Variables and Average Head Start Funding per Child, 1988-1994

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Percent Urban	0.289											0.891
	(2.054)											(1.338)
Percent Black		-20.678*										-12.964*
		(10.522)										(7.759)
Percent Hispanic			12.048*									1.527
			(6.442)									(2.457)
Percent Farmland				6.321**								0.650
				(2.971)								(1.173)
Education Expenditures per Capita					2.190**							0.486
					(1.061)							(0.445)
Welfare Expenditures per Capita						79.074						11.080
						(84.746)						(10.697)
Income per Capita							-0.076**					-0.019
							(0.037)					(0.023)
Percent Single Mother								82.841				58.801
								(61.262)				(51.039)
Percent of Children 0-18 under Poverty									27.861**			5.819
									(13.610)			(6.052)
Fraction of Pop Under 5										177.278		24.701
										(109.938)		(35.460)
Fraction of Pop Older than 65											8.712	16.562
											(6.862)	(11.556)
Log population												56.120
												(54.299)
Obs	254	254	254	254	254	254	254	254	254	254	254	254
R-Squared	0.000	0.072	0.218	0.056	0.058	0.057	0.110	0.090	0.192	0.092	0.003	0.194
F-test	0.020	3.862	3.498	4.527	4.262	0.871	4.298	1.829	4.191	2.600	1.612	2.183
p-value	0.888	0.052	0.064	0.035	0.041	0.353	0.040	0.179	0.043	0.110	0.207	0.017

Notes: The dependent variable is the average federal Head Start (HS) spending per child (2014$) from the Consolidated Federal Funds Reports (CFFR), coupled with the population counts for three- and four-year-olds from the SEER. The 1980 county controls are from the City and County Data Book. For more details about the data, see Appendix II. Estimates are weighted by the population of three- and four-year-olds. Standard errors are clustered at the community level. * p<0.10, ** p<0.05, *** p<0.01.

Table A.3: Correlations between County-Level Variables and Change in Head Start Funding per Child, 1988-1994

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Percent Urban	-0.502											0.735
	(0.565)											(0.596)
Percent Black		-6.093***										-4.672*
		(1.982)										(2.587)
Percent Hispanic			3.084***									-1.596
			(0.958)									(1.330)
Percent Farmland				1.662**								0.219
				(0.730)								(0.577)
Education Expenditures per Capita					0.908**							0.214
					(0.405)							(0.241)
Welfare Expenditures per Capita						10.127						1.461
						(13.355)						(2.551)
Income per Capita							-0.022***					-0.000
							(0.006)					(0.012)
Percent Single Mother								17.957*				10.337
								(10.301)				(16.791)
Percent of Children 0-18 under Poverty									8.079***			6.653*
									(2.359)			(3.374)
Fraction of Pop Under 5										43.327**		9.821
										(19.223)		(17.683)
Fraction of Pop Older than 65											6.124**	4.550
											(2.776)	(5.893)
Log population												20.987
												(27.821)
Obs	254	254	254	254	254	254	254	254	254	254	254	254
R-Squared	0.005	0.087	0.199	0.054	0.138	0.013	0.128	0.059	0.225	0.077	0.021	0.102
F-test	0.789	9.452	10.370	5.193	5.015	0.575	13.869	3.039	11.725	5.080	4.867	2.249
p-value	0.376	0.003	0.002	0.024	0.027	0.450	0.000	0.084	0.001	0.026	0.029	0.013

Notes: The dependent variable is long-change in real federal Head Start (HS) spending per child (2014$) from 1988 to 1994 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 the City and County Data Book. For more details about the data, see Appendix II. Estimates are weighted by the population of three- and four-year-olds. Standard errors are clustered at the community level. * p<0.10, ** p<0.05, *** p<0.01.

Table A.4: Head Start Funding and Director Quality, 1992-1994

	(1)	(2)	(3)
Director has at least college degree	-0.003		
	(0.119)		
Director's yrs of eduation		0.042	
		(0.036)	
Director's yrs of experience			-0.003
			(0.005)
Mean X	0.26	5.36	7.92
Obs	470	194	207

Notes: The dependent variable is real federal Head Start (HS) spending per child (2014$) from Consolidated Federal Funds Reports (CFFR). Head Start program directors' data are from the Program Information Reports (PIR), and include years between 1992 to 1994. Director's years of education is a categorical variable which takes the value "1" for less than high school graduate, "2" for high school graduate,"3" for completed less than two years college, "4" for completed two to four years college but did not graduate, "5" for graduated college, "6" for some post-graduate credits, and "7" for graduate degree. For more details about the data description, see Appendix II. All regressions include county and year fixed effects. Estimates are weighted by the population of three- and four-year-olds. Standard errors are clustered at the community level. * p<0.10, ** p<0.05, *** p<0.01.

Table A.5: Baseline Estimates of the Effect of Head Start Funding on Standardized Math and Reading Test Scores in Third Grade

	All	Males	Females	Whites	Blacks	Hispanics
	(1)	(2)	(3)	(4)	(5)	(6)
A: Standardized Math Scores						
Head Start Funding	0.045**	0.052*	0.041***	-0.011	0.052	0.070***
	(0.018)	(0.028)	(0.013)	(0.045)	(0.043)	(0.019)
Mean Y	-0.332	-0.397	-0.268	-0.048	-0.388	-0.412
B: Standardized Reading Scores						
Head Start Funding	0.023*	0.027	0.023	-0.011	0.033	0.037***
	(0.013)	(0.018)	(0.025)	(0.042)	(0.038)	(0.014)
Mean Y	-0.339	-0.470	-0.209	-0.083	-0.333	-0.434
Obs	335311	166684	168627	72377	76436	182021

Notes: This table reports results obtained when the dependent variable is third-grade standardized test scores in math and reading separately, and the independent variable is real federal Head Start (HS) spending per child (in 2014$) when the child was four years old. HS spending per child is scaled by $500 (average spending during the period of study), thus the coefficients should be interpreted as the effect of a $500 increase in funding. All regressions include controls for demographics and county-level characteristics, school, test year and birth year fixed effects, along with county-specific linear trends. Sample consists of third-grade students who are FRL certified or who are identified as economically disadvantaged, based on the description by the Texas Education Agency (TEA). Student data are from the TEA which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 and 1999. HS spending data are from the Consolidated Federal Files Reports (CFFR) and include years between 1988 and 1994. Standard errors are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.

Table A.6: Effect of Head Start Funding on Standardized Test Scores for Grades Three - Eighth

	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
	(1)	(2)	(3)	(4)	(5)	(6)
A: All - Combined Test Score						
Head Start Funding	0.031**	0.048**	0.025	0.013	0.017	0.009
	(0.013)	(0.021)	(0.016)	(0.017)	(0.020)	(0.035)
Mean Y	-0.355	-0.356	-0.359	-0.287	-0.300	-0.217
B: All - Math Score						
Head Start Funding	0.045**	0.043**	0.027	0.008	0.022	0.009
	(0.018)	(0.022)	(0.023)	(0.019)	(0.019)	(0.036)
Mean Y	-0.332	-0.335	-0.337	-0.251	-0.276	-0.195
C: All - Reading Score						
Head Start Funding	0.023*	0.050***	0.022	0.017	0.011	0.010
	(0.013)	(0.019)	(0.019)	(0.018)	(0.019)	(0.033)
Mean Y	-0.339	-0.361	-0.366	-0.305	-0.307	-0.227
Mean X($)	518	397	355	357	324	316
Obs	335311	209898	263722	264377	254212	143135
D: Hispanic - Combined Test Score						
Head Start Funding	0.051***	0.087***	0.036**	0.027	0.015	0.010
	(0.014)	(0.020)	(0.017)	(0.019)	(0.017)	(0.031)
Mean Y	-0.449	-0.398	-0.370	-0.293	-0.313	-0.216
E: Hispanic - Math Score						
Head Start Funding	0.070***	0.082***	0.035**	0.024	0.009	0.003
	(0.019)	(0.020)	(0.016)	(0.021)	(0.018)	(0.034)
Mean Y	-0.412	-0.368	-0.332	-0.241	-0.275	-0.183
F: Hispanic - Reading Score						
Head Start Funding	0.037***	0.090***	0.039**	0.036*	0.025	0.019
	(0.014)	(0.019)	(0.017)	(0.021)	(0.018)	(0.028)
Mean Y	-0.434	-0.411	-0.389	-0.327	-0.334	-0.238
Mean X($)	687	525	466	456	413	403
Obs	182021	121150	153815	157156	155284	87516

Notes: This table reports results obtained when the dependent variable is standardized test scores combined, in math and reading separately. The independent variable is real federal Head Start (HS) spending per child (in 2014$) when the child was four years old. HS spending per child is scaled by $500 (average spending during the period of study), thus the coefficients should be interpreted as the effect of a $500 increase in funding. All regressions include controls for demographics and county-level characteristics, school, test year and birth year fixed effects, along with county-specific linear trends. Sample consists of students from third through eighth grade who are FRL certified, or who are identified as economically disadvantaged, 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 1994. Standard errors are clustered at the county level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.7: Baseline Estimates of the Effect of Head Start Funding on Standardized Test Scores - Adding School Trends

	All	Males	Females	Whites	Blacks	Hispanics
	(1)	(2)	(3)	(4)	(5)	(6)
Head Start Funding	0.021	0.018	0.020	-0.050	-0.083	0.047***
	(0.013)	(0.025)	(0.021)	(0.064)	(0.093)	(0.015)
Mean Y	-0.355	-0.459	-0.253	-0.069	-0.382	-0.449
Obs	335311	166684	168627	72377	76436	182021

Notes: This table reports results obtained when the dependent variable is combined standardized test scores in math and reading. The independent variable is real federal Head Start (HS) spending per child (in 2014$) when the child was four years old. HS spending per child is scaled by $500 (average spending during the period of study), thus the coefficients should be interpreted as the effect of a $500 increase in funding. 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 third-grade students who are FRL certified, or who are identified as economically disadvantaged, based on the description by the Texas Education Agency (TEA). Student data are from the TEA which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 and 1999. HS spending data are from the Consolidated Federal Files Reports (CFFR) and include years between 1988 and 1994. Standard errors are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.

Table A.8: The Effect of Head Start Exposure on Third Grade Standardized Test Scores Sensitivity of Results to Different Measures of Head Start Exposure

	Per Child (Age 3-4)	Per Capita	Per Poor Child (Age 3-4)
Head Start Funding Per Child (Main)	0.031**		
	(0.013)		
Head Start Funding Per Capita		0.026***	
		(0.009)	
Head Start Funding Per Poor Child			0.027**
			(0.013)
Mean Y	-0.355	-0.355	-0.355
Mean X($)	518	18	480
Obs	335311	335311	335311

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

Table A.9: The Effect of Alternative Head Start Funding Measures on Third Grade Standardized Test Scores

	All	Males	Females	Whites	Blacks	Hispanics
	(1)	(2)	(3)	(4)	(5)	(6)
Head Start Funding Per Capita	0.026***	0.037**	0.018	-0.023	0.057	0.041***
	(0.009)	(0.015)	(0.013)	(0.052)	(0.051)	(0.010)
Mean Y	-0.355	-0.459	-0.253	-0.069	-0.382	-0.449
Mean X($)	18447	18450	18444	10752	9725	25335
Panel B:						
Head Start Funding Per Poor Child	0.027**	0.033	0.023	0.000	0.037	0.050***
	(0.013)	(0.020)	(0.017)	(0.029)	(0.036)	(0.014)
Mean Y	-0.355	-0.459	-0.253	-0.069	-0.382	-0.449
Mean X($)	480346	480912	479786	416108	363267	556897
Obs	335311	166684	168627	72377	76436	182021

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

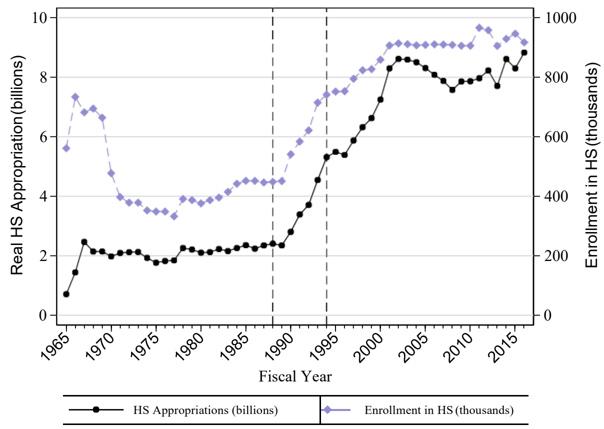
Table A.10: Falsification Tests: The Effect of Head Start Exposure on Third Grade Standardized Test Scores for Students Not Certified for Free/Reduced Lunch

	All	Males	Females	Whites	Blacks	Hispanics
	(1)	(2)	(3)	(4)	(5)	(6)
Head Start Funding	0.002	-0.013	0.006	0.006	0.017	-0.008
	(0.012)	(0.023)	(0.014)	(0.017)	(0.154)	(0.026)
Mean Y	0.312	0.255	0.369	0.392	-0.037	0.050
Obs	407349	206309	201040	316015	31524	49827

Notes: This table reports results obtained when the dependent variable is combined standardized test scores in math and reading, and the independent variable is real federal HS spending per child (in 2014$) when the child was four years old. HS spending per child is scaled by $500 (average spending during the period of study), thus the coefficients should be interpreted as the effect of a $500 increase in funding. All regressions include controls for demographics and county-level characteristics, school, test year and birth year fixed effects, along with county-specific linear trends. Sample consists of third-grade students who are *not* FRL certified based on the description by the Texas Education Agency (TEA). Student data are from the TEA which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 and 1999. HS spending data are from the Consolidated Federal Files Reports (CFFR) and include years between 1988 and 1994. Standard errors are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.

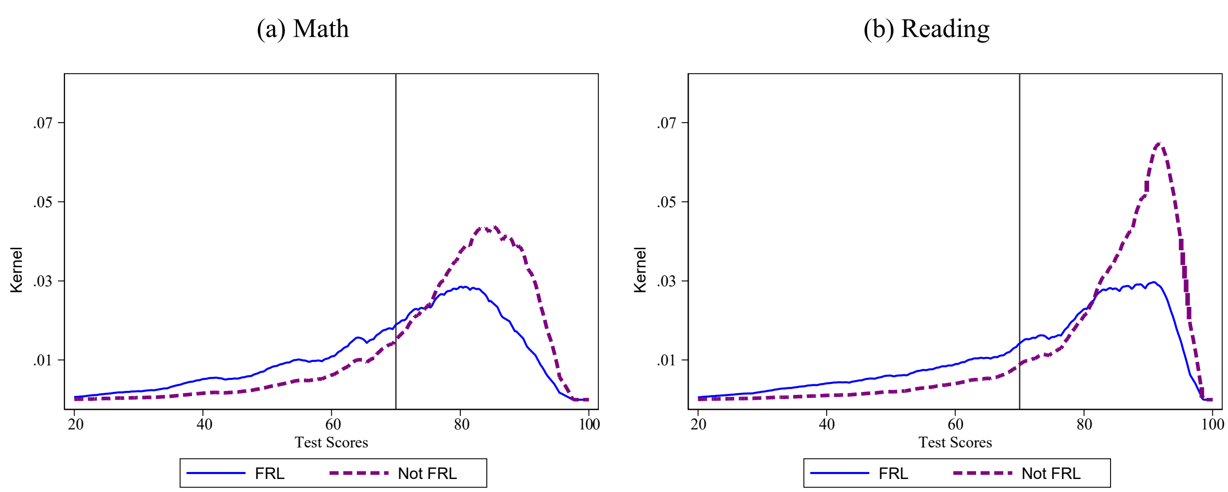
B. Figures

Figure A.1: Head Start Program Facts



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

Figure A.2: Kernel Density of Test Scores, by Free-or-reduced Lunch Status

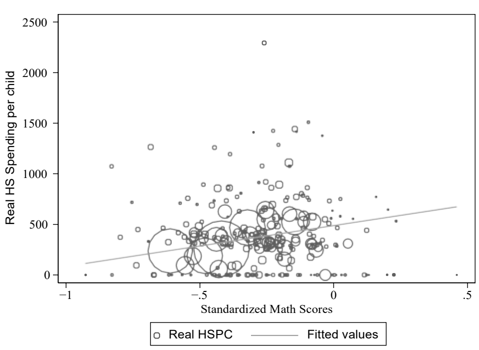


Notes: Test score data include all third-grade students who took the standardized test in Texas between 1994 and 1999, from the Texas Education Agency (TEA). The sample is divided into two groups: (i) students who are *not* identified as economically disadvantaged (Not FRL) and (ii) students who are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

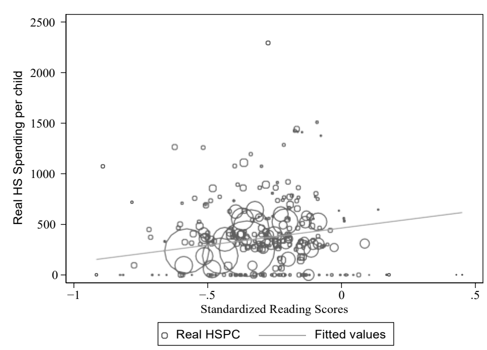
Figure A.3: Raw Correlations between Head Start Funding per Child

and Standardized Test Scores - Free/Reduced Lunch Certified Sample

(a) Math

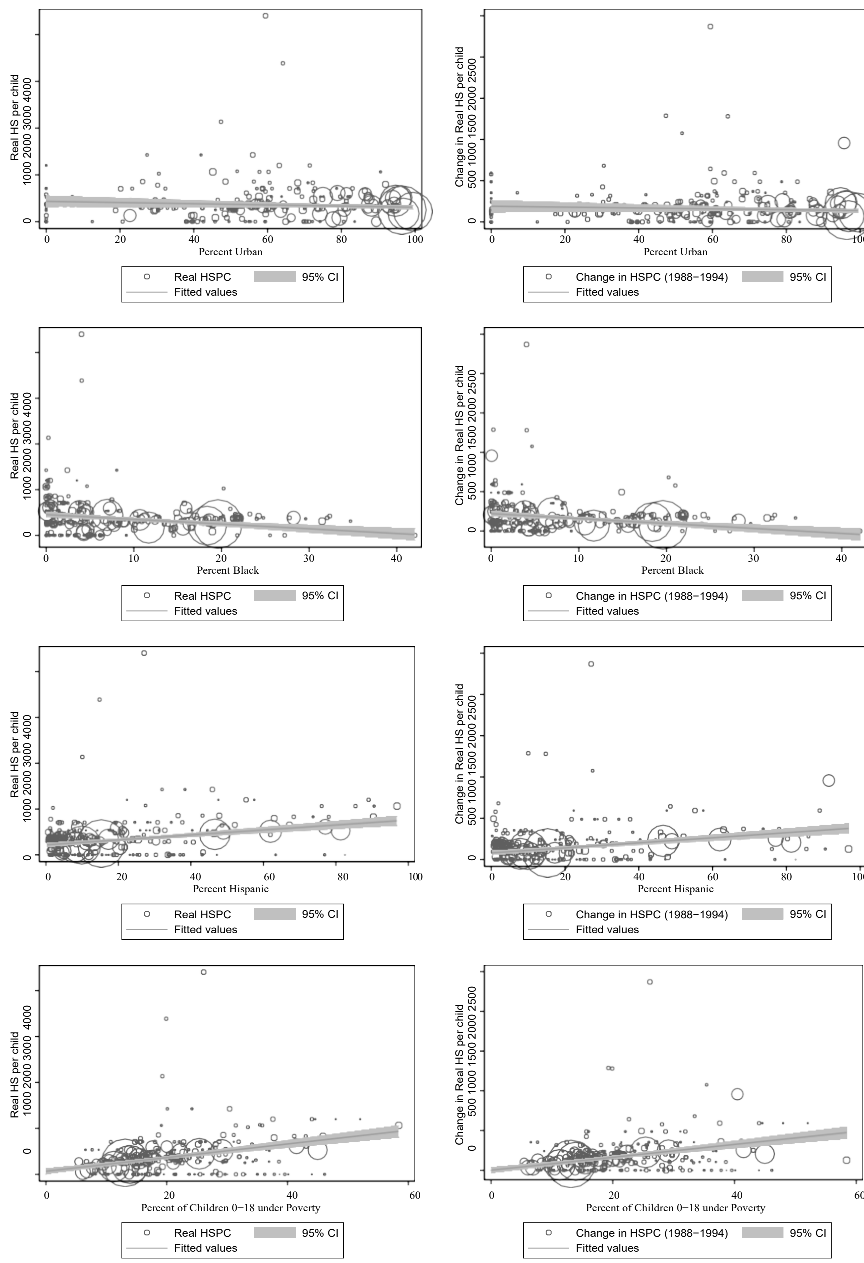


(b) Reading



Notes: Head Start spending (in 2014$) data are obtained from the Consolidated Federal Funds Reports), coupled with the population counts for three- and four-year-olds at the county-level from the SEER. 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 represent counties, and weighted by the population of three- and four-year-olds.

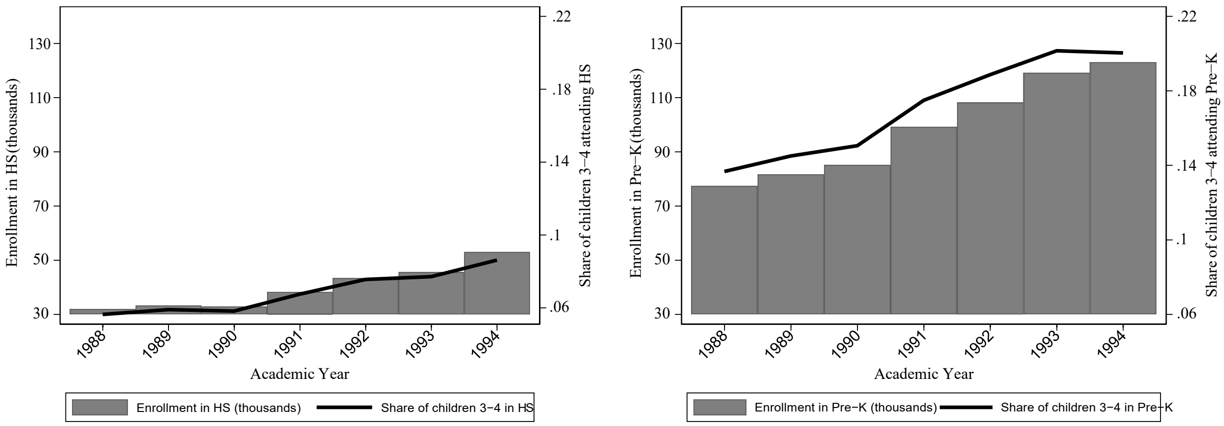
Figure A.4: Correlations between 1980 County Characteristics and Head Start Funding per Child



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

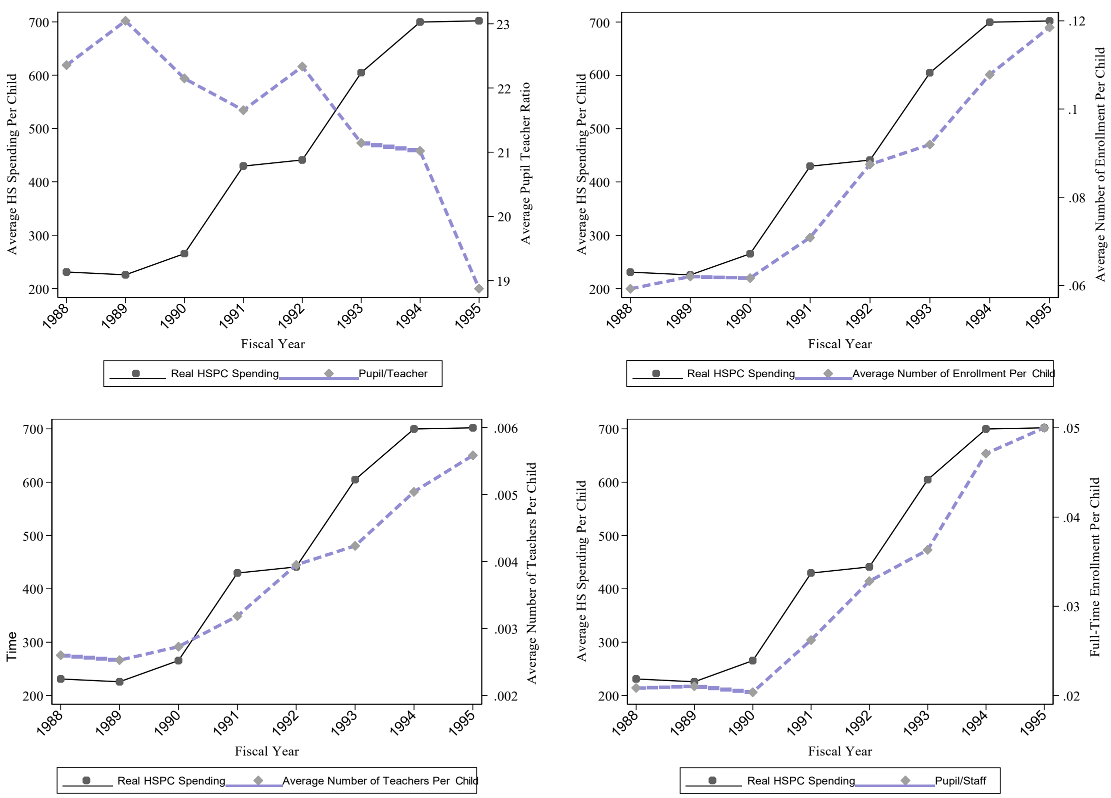
Figure A.5: Early Childhood Education Expansions in Texas

(a) Head Start expansions (b) Pre-K expansions



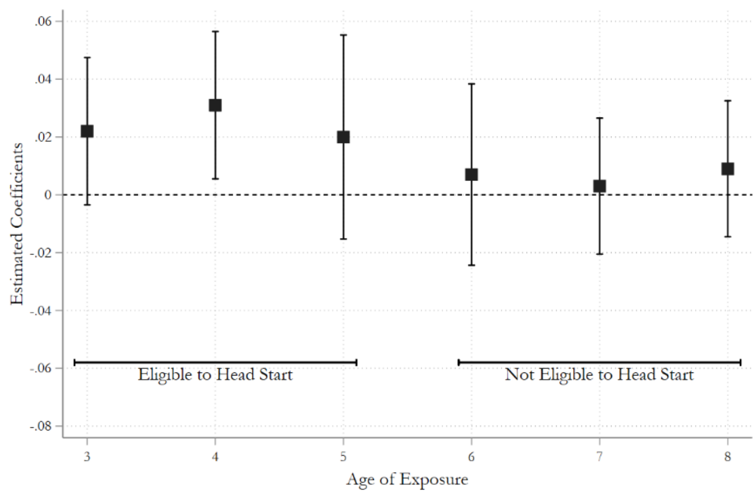
Notes: Head Start enrollment data are from the Program Information Reports (PIR). Pre-K enrollment data are from the Common Core Data (CCD). The share measure is calculated using the population counts for three- and four-year-olds from the SEER.

Figure A.6: Head Start Funding and Program Quality Trends



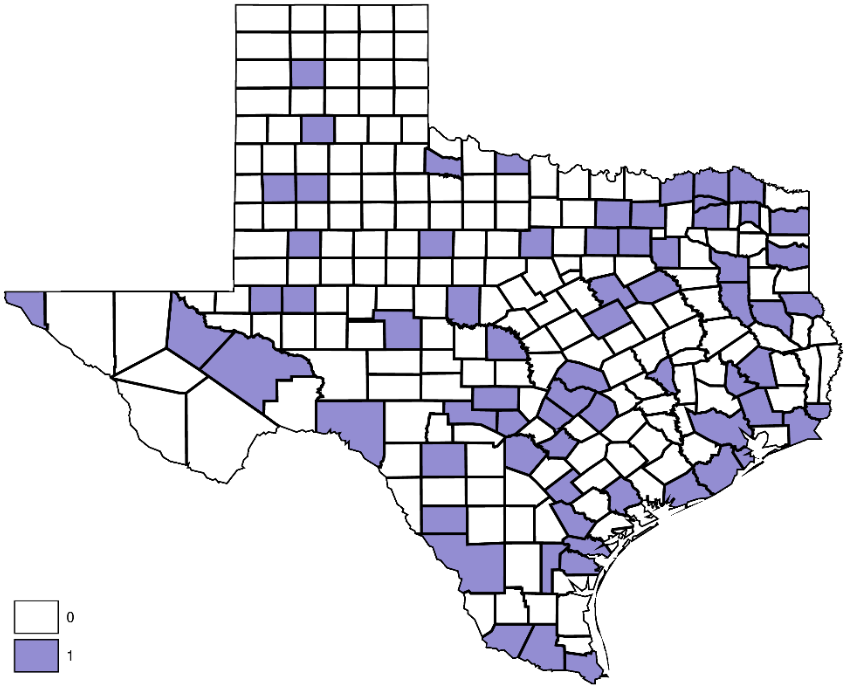
Notes: Head Start spending (in 2014$) data are from Consolidated Federal Funds Reports (CFFR), coupled with the population counts for three- and four-year-olds at the county-level from the SEER. Head Start program data are from the Program Information Reports (PIR) for the years 1988 to 1995.

Figure A.7: Falsification Test: The Effect of Head Start Funding Exposure on Third Grade Test Scores - Differential Effects by Age



Notes: This figure reports coefficients and their 95% confidence intervals obtained from separate regressions when the dependent variable is combined standardized test scores in math and reading. The independent variable is real federal HS spending per child (in 2014$), assigned when the child was at ages 3 to 8 years old. HS spending per child is scaled by $500 (average spending during the period of study), thus the coefficients should be interpreted as the effect of a $500 increase in funding. All regressions include controls for demographics and county-level characteristics, school, test year and birth year fixed effects, along with county-specific linear trends. Sample consists of third-grade students who are FRL certified or who are identified as economically disadvantaged, based on the description by the Texas Education Agency (TEA). Student data are from the TEA which include information on year of birth, ethnicity, economic disadvantage indicators and test scores conducted between 1994 and 1999. HS spending data are from the Consolidated Federal Files Reports (CFFR) and include years between 1988 and 1994. Standard errors are clustered at the county level.

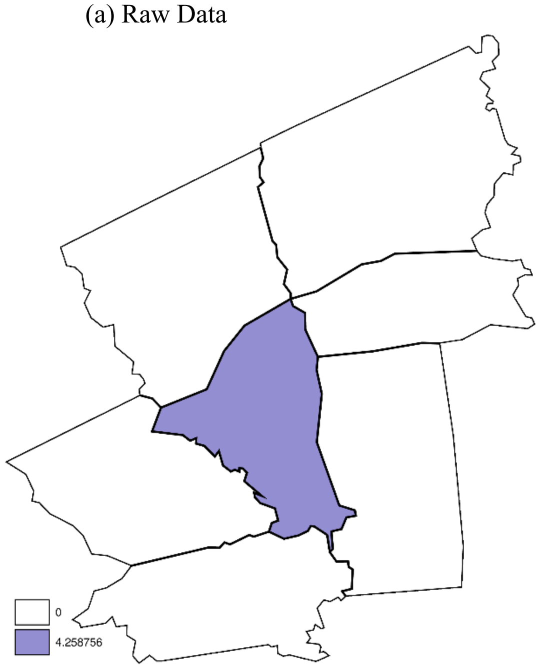
Figure A.8: Raw Data: Indication of Positive Head Start Spending in 1994



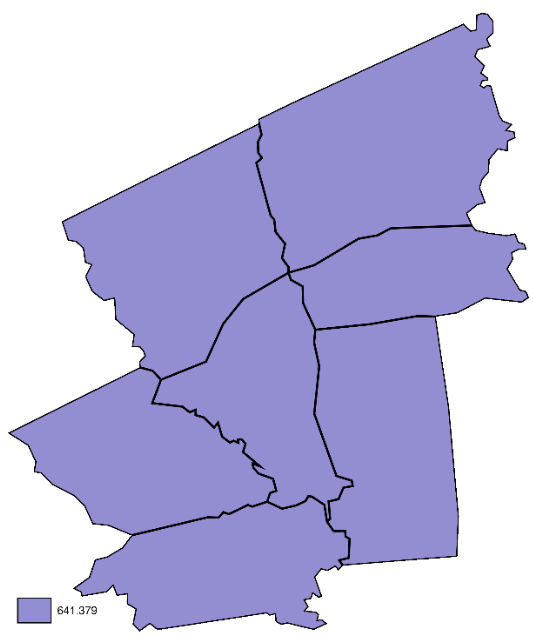
Notes: Raw federal Head Start spending data at the grantee-level, from the Consolidated Federal Funds Reports. There are 69 grantees that served 196 counties in 1994.

Figure A.9: Brazos Valley Community Action Agency, Before and After Reallocation

(a) Raw Data



(b) After Allocation



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

II. Data Appendix

A. Public Use Data

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

  Federal government expenditures data are reported annually in states and counties of the U.S. by program type. CFFR reports detailed information on the geographic distribution of federal program expenditures, including Head Start grants, using data submitted by federal departments and agencies. Head Start is administered by the Department of Health and Human Services (HHS), Administration for Children and Families (ACF), and Office of Head Start (OHS). HHS describes "grantees" as the agencies that receive grant awards directly. "Delegates" are other agencies grantees may contract services. CFFR data span from 1983-2007, available through the National Archives. The program identification code for Head Start expenditures is 93.600.

  When the federal government announces grant availability for a specific local community, it also announces which counties each grantee is expected to serve. However, the retrospective grant announcements are not available to researchers. To determine the serving counties for each grantee, I use an administrative data set on Head Start budgets (called PCCOST) provided by Currie and Neidell (2007). PCCOST includes detailed information on the allocation of total expenditures for education, health, and other services for each grantee and its network. These data cover the years 1990-2000 with comprehensive coverage of networks for the year 1994. I assigned the networks using the 1994 data. Based on my web search for a randomly chosen subset of grantees in Texas, I did not find evidence that the serving counties changed from 1988 to 1994.[1] I then confirmed these networks of counties using grantees' websites and an additional data set provided by Frisvold (2006).

[1] 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.

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. Appendix Figure A.8 maps out the raw data in Texas in 1994, with 69 grantees. These grantees serve 196 counties in Texas.

To give a concrete example of the construction of these networks, consider the Brazos Valley Community Action Agency's headquarters in Texas, established in 1967,[2] located in Bryan, Texas (Brazos County). This grantee serves Head Start programs in six other counties.[3] Thus, the raw data as shown in Appendix Figure A.9a records around $4.3 million in expenditures for Brazos County in 1994 and zero dollars for all the serving counties. Using the network of counties that I constructed, I reallocated dollars for the serving counties in proportion to the number of age-eligible children in each network (see Appendix Figure A.9b for reallocation map).

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

  (a) 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 includes county-level population counts for each age group starting in 1969.

  (b) County-level population to construct the per capita measure also comes from SEER.

  (c) 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 younger than 17 and children 5-17 are available. Using these two variables, I construct the number of children younger than age five by taking

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

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

the difference. To create age-eligible poor child counts, I follow Frisvold (2006) that states that children who are age 3-4 years old are two-fifths of children younger age 5. These data are only available for years 1989, 1993, 1995, 1997-1999, the years in between are determined through linear interpolation.

- **Program Information Reports (PIR)**: Starting in 1988, the Office of Head Start has collected comprehensive data from all grantees and delegates on the services, staff, children, and families served by the program. These data are essential for my analysis as they provide information on the number of funded enrollees, number of staff, demographic composition of children, qualifications of directors and teachers, and so on. I use this information to show how much funding expansions translate into enrollment versus the quality of the Head Start programs. PIR data are not commonly used because the format of these data and variables collected changed over time. The main variables I use from the PIR dataset, enrollment (all and full-time), and number of teachers, are available all the years starting in 1988. Enrollment by race/ethnicity is available in 1988, and between 1992-1994, the missing data in the years between (1989-1991) are linearly interpolated.

  PIR data have become more comprehensive with additional information (i.e., educational background of directors and teachers) included in the surveys during and after the 1990s. For example, relevant but not usable in my analysis, the number of teachers with AA or BA degrees is collected starting in 1999. For my purposes, data on directors' salary and education levels are available starting in 1992. Part of these data from 1988 and 1998 were generously provided to me by Currie and Neidell (2007).

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

CCD's pre-K attendance data report public pre-K enrollment that takes place in public elementary schools, excluding public pre-K programs that take place in private preschools or daycare. Therefore, it has measurement error.

- **County-level Demographics**: I use the Regional Economic Information Systems (REIS) to construct county-year data on per capita income, per capita transfer payments for cash income support (Aid to Families with Dependent Children and SSI), medical benefits (Medicare, Medicaid, and Children's Health Insurance Program), food assistance (food stamps), retirement, and disability programs. I 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 the survey time.

  Using the 1983 City and County Data Book (before the 1990s Head Start spending expansion), I construct other county demographics. These include the 1980 population living in an urban area, black, Hispanic, single parent, less than age 5, ages 65 or older, and percentage of 0- to 18-year-olds living in poverty as well as income, education, and welfare spending per capita (in 2014 dollars). To control for exposure to business cycles at birth, I use the county-year unemployment rate from the Bureau of Labor Statistics.

  Finally, I use SEER data to control for the composition of the population demographics at the county level by racial and age groups.

B. Data from Texas Education Agency (TEA)

Student-level data from the TEA[4] include information on test scores monitored through the Texas Academic Assessment System (TAAS) for grades 3 to 8 between 1994 and 1999. These data are de-identified and are not linked across the years. They contain information on birth year, sex, race/ethnicity, FRL status for all students, language proficiency, and

[4] These data are currently administered and available to approved researchers by the University of Texas Education Research Centers in Austin and Dallas. The data available at these centers are more detailed and available to be linked across years.

special education status for students in the third grade. Descriptions for two critical variables:

i. Main outcome variable: Texas Learning Index (TLI) reading and math scores.

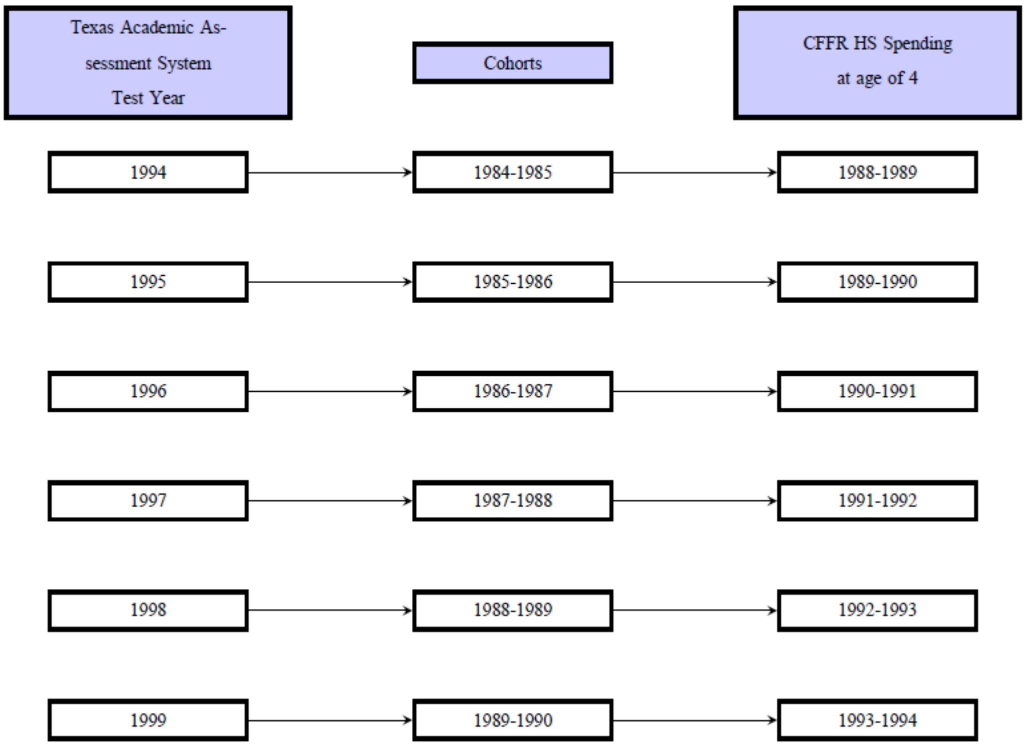
ii. Economic disadvantage variable:

1. *0*: Not identified as economically disadvantaged
2. *1*: certified for free meals under the National School Lunch and Child Nutrition Program
3. *2*: Eligible for reduced-price meals under the National School Lunch and Child Nutrition Program
4. *9*: Other economic disadvantage, including:
    a. from a family with an annual income at or below the official federal poverty line
    b. eligible for Temporary Assistance to Needy Families (TANF) or other public assistance
    c. received a Pell Grant or comparable state program of need-based financial assistance
    d. eligible for programs assisted under Title II of the Job Training Partnership Act (JTPA)
    e. eligible for benefits under the Food Stamp Act of 1977

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

C. Data Structure

The chart below shows the correspondence of test years with third-grade cohorts and their Head Start exposure at age four.



III. Cost-Benefit Analysis

This section provides a back-of-the-envelope calculation of a cost-benefit analysis of federal Head Start spending expansions. Several studies have attempted to calculate the social benefits of the Head Start 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 requires strong assumptions. Here, I attempt to obtain the costs and benefits associated with a $500 increase in federal Head Start funding, assuming that the Head Start program accrues test score gains in the third grade and that these estimates will translate into later earnings.[5] This analysis adopts the cost-benefit formulation constructed by Kline and Walters (2016) for one year of Head Start attendance using the HSIS. All monetary values are converted to 2014 dollars. The marginal cost is $500 per child because the analysis is based on the test score impact of a $500 increase in federal funding per child. To calculate the marginal benefit, I need two parameters: (1) the potential link between third-grade test scores and earnings and (2) a prediction of average earnings for my sample (FRL-certified students in Texas). Although I cannot directly measure the impact on earnings due to data limitations, other studies that examine the link between test scores and earnings provide estimates. Following Kline and Walters (2016), I use a conservative estimate that earnings rise by 10 percent for each standard deviation increase in test scores.

Chetty et al. (2011) calculate that the present value of earnings at age 12 for the average individual in the U.S. is approximately $566,720 (in 2014 dollars). The average present discounted value of the predicted earnings at age four corresponds to around $434,000, with a discount rate of 3 percent. Adjusting that the median earnings in Texas are about 94 percent of the median earnings in the U.S., the average present discounted value corresponds to $407,960.[6]

[5] 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 who attended Head Start in similar years. There is also an implicit assumption that third-grade test score effects do not fade out over time.

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

Children who participate in FRL are economically disadvantaged and are 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.[7] Using the estimate for intergenerational income elasticity reported by Lee and Solon (2009) of 0.4, the average child in FRL is expected to earn 75.2 percent of the average $(1 - (1 - 0.38) * 0.4)$. These predictions yield a present value of earnings of approximately $307,000.

Putting the pieces together, 10 percent of $307,000 is $30,700 and multiplying it with the test score impact of 0.031 yields a projected earnings impact of roughly $950. These calculations show that a conservative estimate of the benefit-cost ratio is around 1.9. This estimate of 1.9 is much larger than the estimated rates of return associated with the Earned Income Tax Credit (0.88) and the Food Stamps Program (0.66) reported in Hendren (2016).

Finally, I compute the marginal value of public funds (MVPF) using the framework developed by Kline and Walters (2016) and Hendren and Sprung-Keyser (2020) and assuming the tax and transfer rate of 20 percent on the earnings. The MVPF is the ratio of the willingness to pay (the net after-tax benefits to individuals) divided by the net cost of the policy to the government (including the upfront cost and the other fiscal externalities) (Hendren and Sprung-Keyser, 2020). In my setting, the government's tax collection would be close to $190 (=$950*0.2) from a $950 increase in earnings with a 20 percent tax and transfer rate, which would imply that the net after-tax benefit to the participants of $760 (=$950-$190) and the net cost of a $500 increase in funding for Head Start to the government of $310 (=$500-$190). These calculations show a conservative estimate of the MVPF of 2.45 (=$760/$310), which is similar to the estimated MVPF of one year of Head Start attendance reported in Kline and Walters (2016) for the HSIS cohorts (around 2.41) but larger than the estimated MVPF for the EITC and the nutrition assistance program expansions (Hendren and Sprung-Keyser, 2020).[8]

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

References for Data Sources

Consolidated Federal Funds Report, Fiscal Years 1982-2007: County Areas [CFFR-dataset]. Washington D. C.: Bureau of the Census and U. S. Department of Commerce. Retrieved from National Archives Catalog https://catalog.archives.gov/id/626196

National Cancer Institute. 2017. *Download U. S. Population Data- 1969-2016* [SEER - dataset]. Washington D. C.: National Institutes of Health. Data set retrieved from http://www.nber.org/data/seer_u.s._county_population_data.html

Office of Head Start Program Facts. Office of Head Start [HS Program Facts – dataset]. Retrieved from https://eclkc.ohs.acf.hhs.gov/hslc/data/factsheets/2015-hs-program-factsheet.html

Office of Head Start Program Information Report. Office of Head Start [PIR – dataset]. Retrieved from http://eclkc.ohs.acf.hhs.gov/PIR.

U.S. Administration for Children Youth and Families. 2015. Head Start Fact Sheet. Washington, DC: Head Start Bureau. Retrieved from https://eclkc.ohs.acf.hhs.gov/hslc/data/factsheets/2015-hs-program-factsheet.html

U.S. Census Bureau, Small Area Income and Poverty Estimate [SAIPE – dataset]. Retrieved from https://www.census.gov/programs-surveys/saipe/data/datasets.html

U.S. Department of Commerce. Bureau of Economic Analysis. Regional Economic Information System [REIS – dataset]. Retrieved from https://www.bea.gov/data/economic-accounts/regional

U.S. Department of Education, National Center for Education Statistics, Common Core of Data [CCD – dataset]. Retrieved from https://nces.ed.gov/ccd/pubschuniv.asp

U.S. Department of Health and Human Services (HHS), Administration for Children and Families, Head Start Bureau. Head Start. 1999 [PCCOST – dataset]. PCCost: The computerized version of the grant application package SF424A version 4.01 Y2K.

United States. Bureau of the Census. County and City Data Book [United States], 1983. Inter-university Consortium for Political and Social Research [distributor], 2008-06-18. https://doi.org/10.3886/ICPSR08256.v1

References

Carneiro, Pedro, and Rita Ginja. 2014. "Long-Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start." *American Economic Journal: Economic Policy*.

Chetty, Raj, John N Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2011. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star*." *Quarterly Journal of Economics*, 126(4).

Consolidated Federal Funds Report, Fiscal Years 1983-2010: County Areas [CFFR]. Washington D. C.: Bureau of the Census and U. S. Department of Commerce. Reports https://www.census.gov/library/publications/time-series/cffr.All.html

Currie, Janet, and Matthew Neidell. 2007. "Getting Inside the Black Box of Head Start Quality: What

Matters and What Doesn't." *Economics of Education Review*, 26(1): 83–99.

Elango, Sneha, Jorge Luis García, James J Heckman, and Andrés Hojman. 2015. "Early Childhood Education." National Bureau of Economic Research.

Frisvold, David E. 2006. "Head Start Participation and Childhood Obesity." *Vanderbilt University Economics Working Paper No. 06-WG01*.

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

Hendren, Nathaniel, and Ben Sprung-Keyser. 2020. "A Unified Welfare Analysis of Government Policies." *The Quarterly Journal of Economics,* 135 (3): 1209-1318.

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

Lee, Chul-In, and Gary Solon. 2009. "Trends in Intergenerational Income Mobility." *The Review of Economics and Statistics*, 91(4): 766–772.