

**PREP SCHOOL FOR POOR KIDS:
THE LONG-RUN IMPACTS OF HEAD START ON HUMAN CAPITAL AND
ECONOMIC SELF-SUFFICIENCY**

Martha J. Bailey, Shuqiao Sun, and Brenden Timpe

November 20, 2018

[\[click here for most updated draft\]](#)

Abstract

This paper evaluates the long-run effects of Head Start using large-scale, restricted 2000-2013 Census-ACS data linked to date and place of birth in the SSA's Numident file. Using the county-level rollout of Head Start between 1965 and 1980 and state age-eligibility cutoffs for school entry, we find that participation in Head Start is associated with increases in adult human capital and economic self-sufficiency, including a 0.29-year increase in schooling, a 2.1-percent increase in high-school completion, an 8.7-percent increase in college enrollment, and a 19-percent increase in college completion. These estimates imply sizable, long-term returns to investing in large-scale preschool programs.

JEL Codes: I2, J24, J6

Contact Information

Bailey: Department of Economics, University of Michigan, 611 Tappan Street, Ann Arbor, Michigan 48109; Email: baileymj@umich.edu; Website: www-personal.umich.edu/~baileymj. Sun: Department of Economics, University of Michigan, 611 Tappan Street, Ann Arbor, Michigan 48109; sqsun@umich.edu; Timpe: Department of Economics, University of Michigan, 611 Tappan Street, Ann Arbor, Michigan 48109; btimpe@umich.edu.

Acknowledgements

Data collection for the War on Poverty project was generously supported by the National Institutes of Health (R03-HD066145). Data linkage and analyses for this project were generously supported by the Laura and John Arnold Foundation. The opinions and conclusions expressed herein are solely those of the authors and should not be construed as representing the opinions or policy of any agency of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. We gratefully acknowledge the use of the services and facilities of the Population Studies Center at the University of Michigan (funded by NICHD Center Grant R24 HD041028). We are also grateful for support for the Michigan RDC from the NSF (ITR-0427889). During work on this project, Timpe was partially supported by the NIA (T32AG000221) as a UM Population Studies Center Trainee. We are grateful to Doug Almond, Hilary Hoynes, and Diane Schanzenbach for sharing the Regional Economic Information System (REIS) data for the period of 1959 to 1978 and the data and dofiles to replicate their analysis with the PSID; and Clint Carter for the many hours spent helping us disclose these results. Evan Taylor and Bryan Stuart provided exceptional assistance in translating string names in the SSA's NUMIDENT file into GNIS codes. Ariel Binder, Dorian Carloni, and Bryan Stuart also contributed substantially to the cleaning of the restricted census data. We are grateful for comments from Liz Cascio, Janet Currie, Doug Miller, Chloe Gibbs, Greg Duncan, and Gary Solon.

Convincing evidence on the longer-term impacts of scaled-up pre-k programs on academic outcomes and school progress is sparse, precluding broad conclusions. The evidence that does exist often shows that pre-k-induced improvements in learning are detectable during elementary school, but studies also reveal null or negative longer-term impacts for some programs.

~ Brookings Pre-Kindergarten Task Force of Interdisciplinary Scientists, Phillips et al. (2017)

In 1965, the U.S. began a new experiment in the provision of public preschool for disadvantaged children. The motivation was simple: “the creation of and assistance to preschool, day care, or nursery centers for 3- to 5-year-olds...will provide an opportunity for a *head start* by canceling out deficiencies associated with poverty that are instrumental in school failure” (United States Senate Committee on Labor and Public Welfare 1964). The program that ensued is the now-famous “Head Start”, a “prep school for poor kids” which aimed to help millions of children escape poverty (Levitan 1969).

More than fifty years later, Head Start is one of the most popular of the War on Poverty’s programs, serving around 900,000 children annually at a cost of \$9.6 billion in 2017 to the federal government.¹ Unlike expensive, small-scale “model” programs such as Perry Preschool and Abecedarian, Head Start’s architects prioritized widespread access, calculating that a massive preschool expansion would maximize its poverty-fighting (and political) benefits. Skepticism about the quality of this large-scale preschool program coupled with difficulties in evaluation have generated controversy about its short-term benefits for decades (Duncan and Magnuson 2013, Currie 2001, Westinghouse Learning Corporation 1969). Convincing evidence regarding Head Start’s long-term effects has remained more elusive, thanks to the lack of program randomization in its early years, small sample sizes of longitudinal surveys, and the difficulty of measuring adults’ access to Head Start decades ago. Consequently, the best estimates of Head Start’s long-term effects are limited by lingering concerns about endogeneity (sibling comparison designs, Currie and Thomas 1995, Garces et al. 2002, Deming 2009) and imprecision (measurement error in funding and access, Ludwig and Miller 2007, or small sample sizes, Carneiro and Ginja 2014). Whether Head Start achieved its goal of increasing life opportunities for children remains an open question.

¹ <https://eclkc.ohs.acf.hhs.gov/sites/default/files/pdf/hs-federal-funding-enrollment-history.pdf>

This paper uses large-scale data to estimate Head Start’s long-term effects on human capital and economic self-sufficiency. By linking the restricted long-form 2000 Census and 2001-2013 American Community Surveys (ACS) to the exact date and place of birth from the Social Security Administration’s (SSA) Numident file, we observe outcomes for one-quarter of U.S. adults as well as a high quality measure of their access to and eligibility for Head Start as children. The resulting sample is four orders of magnitude larger than longitudinal surveys, and information on place of birth and exact date of birth ameliorates (potentially non-classical) measurement error in childhood access to Head Start.

Our research design exploits the county-level roll-out of Head Start programs from 1965 to 1980 at the Office of Economic Opportunity (OEO) (Levine 1970, Bailey and Goodman-Bacon 2015, Bailey and Duquette 2014, Bailey and Danziger 2013, Bailey 2012). This approach exploits the well-documented “great administrative confusion” at the OEO (Levine 1970), mitigating problems of measurement error in archival funding data (Barr and Gibbs 2017) and concerns about the endogeneity of Head Start funding *levels*. An additional strength of our design is that it leverages Head Start’s age-eligibility guidelines, comparing cohorts who were age-eligible when it launched (ages 5 and younger) to cohorts born in the same county that were age-*ineligible* (children 6 and older). The substantial variation in adult outcomes means that even our large dataset does not permit the estimation of a regression discontinuity or regression kink design, but our approach is based on similar assumptions. Much like a regression kink design (Card et al. 2015), our key identifying assumption is that Head Start’s causal effect is the *only* reason for a change in the *relationship* between a child’s age at the program’s launch and her outcomes as an adult. Because even our large-scale data are not large enough to estimate a regression kink formally, we present both event-study estimates as well as trend-break tests based on spline parameterizations.

The results suggest that Head Start increased the human capital and economic self-sufficiency of disadvantaged children. An index of adult human capital rose by 10 percent of a standard deviation among Head Start participants relative to children born in the same county who were age 6 when the program began. Participating children achieved 0.29 more years of education, were 2.1 percent more likely to complete high school, 8.7 percent more likely to enroll in college, and 19 percent more likely to complete

college. In addition, Head Start increased economic self-sufficiency in adulthood by almost 4 percent of a standard deviation—gains driven largely by a 12-percent reduction in adult poverty and a 29-percent reduction in public assistance receipt. We find no evidence of reductions in incarceration. Heterogeneity in Head Start’s effects suggests that they reflect, in part, practices outside of pre-school curriculum. In particular, health screenings and referrals as well as more nutritious meals appear to be important mechanisms for the program’s effects on disadvantaged children. In addition, the effects of Head Start appear to be complement greater family and public resources arising from a stronger economy. Overall, Head Start appears to have achieved the goals of its early architects, both increasing children’s economic opportunities and reducing poverty.

A final analysis quantifies the private, internal rates of return to dollars spent on Head Start in the 1960s and 1970s. Rather than using changes in wage income directly, we use the *National Longitudinal Survey of Youth 1979* (NLSY79) to predict changes in earnings for the relevant cohorts net of any ability differences (Neal and Johnson 1996, Deming 2009). Using potential earnings accounts for Head Start-induced negative selection in men’s employment (driven by reductions in disability) and positive selection in women’s employment (due to the income effect dominating for the least skilled women). This exercise suggests a private internal rate of return to Head Start of 7.7 percent, which ranges from around 4 percent for women to 11 percent for men. Using *only* savings on public assistance expenditures as a conservative method to calculate the program’s social returns, we find that the internal rate of return of putting one child through Head Start is 2.4 percent. In short, this paper’s estimates suggest substantial long-run returns to America’s first scaled-up, public preschool program. While the results do not imply that all of today’s large-scale preschool programs work, they suggest that some less-than-model preschool programs may have lasting effects—a key finding for current policy deliberations (Phillips et al. 2017).

I. THE LAUNCH OF HEAD START IN THE 1960S AND EXPECTED EFFECTS

In the 1960s, the idea that preschool could improve children’s cognitive development was revolutionary. Challenging the conventional notion that IQ was immutable and fixed at birth, Joseph McVicker Hunt’s (1961) book, *Intelligence and Experiences*, persuasively argued that children’s intelligence could be significantly improved by altering their experiences. Benjamin Bloom further

emphasized that the first four years of children’s lives was a “critical period,” noting that “intelligence appears to develop as much from conception to age 4 as it does from age 4 to 18” (1964). This idea suggested an innovative strategy for poverty prevention. Because poor children started school with significantly less educational background, comprehensive preschool could give them a “Head Start,” improving their success in school and addressing a root cause of poverty.

A. *A Brief History of Head Start’s Launch*

Funded by the OEO, Head Start began as an 8-week summer program in 1965. After a successful first summer, President Lyndon Johnson announced that Head Start would become a full-year program for children ages 3 to 5.² The director of the OEO wrote 35,000 letters to public health directors, school superintendents, mayors and social services commissioners to encourage applications. The OEO also made a special effort to generate applications in America’s 300 poorest counties (Ludwig and Miller 2007).

Head Start’s political popularity led to an even faster launch than other War on Poverty programs. Figure 1 shows the program’s quick expansion. By 1966, Head Start had begun in more than 500 counties where over half of the nation’s children under age 6 resided. By 1970, federal expenditures on the program reached \$326 million, or \$2.1 billion in 2018 dollars (OEO 1970). This early expansion ensured that by 1970, Head Start existed in roughly half of U.S. counties, putting preschool programs within a short drive for 83 percent of children under age six (Appendix Table A1).

The exact timing of Head Start’s launch depended on many idiosyncratic factors. The OEO’s “wild sort of grant-making operation” has been well documented in oral histories (Gillette 1996: 193) as well as in more recent, quantitative analyses (Bailey and Duquette 2014). In the case of Head Start, other factors were key as well: how excited were local institutions or politicians about the program? Was their adequate and available space to launch? Could the program be integrated within the public school system or would it remain separate? The final result of the grant making process and local constraints was that Head Start began in areas that were significantly more populous and urban (Appendix Table A2)—areas where more children could be served. In addition, urban areas were funded earlier. After accounting for population and

² “Preschool” also included five-year-olds, because public kindergarten was not yet universal (Cascio 2009).

urban differences, the roll-out of Head Start was not strongly affected by other county characteristics (Appendix Table A3). Consistent with the historical evidence that this national program was rushed into existence, exactly when Head Start began after 1965 does not appear to be systematically related to pre-existing local characteristics.

B. Head Start's Mission

Head Start's architects adopted a holistic approach that aimed to develop children's mental and physical abilities by improving health; self-confidence; verbal, conceptual, and relational skills; and raising parent involvement. Levitan (1969) notes that Head Start's 1966-7 budget included early childhood education (daily activities and transport, 70 percent), health services (including immunizations, screenings and medical referrals), and nutrition (17-20 percent). Parent involvement, social services (e.g., helping families cope with crises), and mental health services accounted for the remaining budget.

The expected effects of the program on adult outcomes could flow directly from the early learning facilitated by the program. But the role of health services and nutrition may be important as well. Head Start's vaccinations and screening (e.g., tuberculosis, diabetes, vision, hearing) and referrals to local physicians may have prevented complications from childhood diseases (North 1979, Ludwig and Miller 2007) and helped parents obtain simple, cost-effective technologies to improve learning (e.g., eye glasses and hearing aids or antibiotics to reduce hearing damage from ear infections). Healthy meals and snacks may have also raised children's ability to learn. Early estimates suggest that more than 40 percent of children entering Head Start were receiving less than two-thirds of the recommended allotment of iron, and 10 percent were extremely deprived in terms of their daily calories (Fosburg et al. 1984). Among children who received blood tests in the 1968 full-year program, 15 percent were found to be anemic (DHEW 1970). Reducing these nutritional deficiencies could also translate into significant gains in educational achievement in both the short and longer term (Frisvold 2015).

The challenges of *quickly* starting a new *national* program meant that implementation often deviated from ideals. Not only did Head Start lack curricular standardization, but programs struggled to find high-quality teachers to achieve the suggested pupil-to-teacher ratio of 15:1. As a practical solution many centers relied on para-professionals, most of whom lacked post-secondary education; thirty percent

had not finished high school (Hechinger 1966, Braun and Edwards 1972).³ In addition, many components of Head Start phased in slowly. For instance, the OEO wrote that in 1965, “the proportion of children receiving treatment for conditions discovered in Head Start medical and dental examinations...was probably under 20 percent. It rose to over 65 percent in 1966, and in 1967 we fully expect it to have reached over 90 percent” (OEO 1967).

Consequently, Head Start in its earliest years was far from a model preschool program. Nevertheless, even the less-than-ideal implementation of Head Start was likely higher quality than the alternatives available to low income children in the 1960s (Currie 2001). Importantly, similar concerns hold today: Head Start’s quality score from the National Institute for Early Education Research places Head Start program quality around the median of the score distribution (Espinosa 2002) but the program may still be much better than informal child care (Loeb 2016).

II. LITERATURE REGARDING THE LONG-TERM EFFECTS OF HEAD START

Previous evaluations of Head Start provide suggestive evidence of the program’s long-term effects on human capital and economic self-sufficiency.⁴ One pioneering approach was the use of family fixed effects with longitudinal data. Building on work by Currie and Thomas (1995), Garces, Thomas, and Currie (2002) used the Panel Study of Income Dynamics (PSID) to compare children who participated in Head Start to their siblings who did not. They show that Head Start increased high school graduation rates and college enrollment among whites and reduced arrest rates among blacks. Using a similar research design for more recent cohorts in the National Longitudinal Survey of Youth (NLSY), Deming (2009) finds that Head Start participation had large and positive effects on a summary index of adult outcomes (including high school graduation, college attendance, “idleness,” crime, teen parenthood, and health status). Well-known critiques caution that sibling comparisons may suffer from well-known sources of endogeneity bias (Griliches 1979, Bound and Solon 1999). In addition, small sample sizes in longitudinal surveys may

³ Sizable variation in preschool quality persists today. For an overview see Currie (2001), Cascio and Schanzenbach (2013), and Duncan and Magnuson (2013).

⁴ See reviews of studies of Head Start’s short-term effects (Currie 2001, Cascio and Schanzenbach 2013, Duncan and Magnuson 2013, Gibbs, Ludwig, and Miller 2014).

provide unreliable estimates of Head Start’s effects (Grosz, Miller, and Shenhav 2017).

More recent work exploits shifts in access to Head Start using three distinct research designs. The path-breaking application of RD in Ludwig and Miller (2007) exploited the OEO’s special effort to generate grant proposals from the 300 poorest counties. Comparing the outcomes of children on either side of this threshold, they find evidence that Head Start reduced childhood mortality and increased the receipt of high-school degrees and college enrollment. However, because the 1990 and 2000 Censuses required them to use county of residence in adulthood to proxy for childhood Head Start access, measurement error causes their education results to be sensitive to specification and often statistically insignificant.⁵ Carneiro and Ginja (2014) use an RD in state-, year-, and household-based income eligibility cutoffs for more recent Head Start programs. They find that Head Start decreased behavioral problems, the prevalence of some health conditions (including obesity) between the ages of 12 and 17, and crime rates around age 20. They find a positive though statistically insignificant effect on receiving a high-school diploma as well as suggestive evidence that Head Start *reduced* college enrollment.

In work closely related to this paper, three studies make use of county-year variation in Head Start funding in the 1960s and 1970s to quantify the program’s long-term effects. Using a sample of likely eligible children from the NLSY, Thompson (2017) finds that greater funding for Head Start at ages 3 to 6 raised college graduation rates, reduced the incidence of health limitations, and tended to raise adult household income. Focusing on a “high impact” sample, Johnson and Jackson (2017) find that an average level of Head Start and education spending increases the likelihood that children graduated from high school by 8 percentage points and gained 0.39 years of schooling. These children also experienced a 7.8 log-point increase in adult wages, a 14.4 log-point increase in adult family income at ages 20 to 50, a 3.6 percentage-point reduction in poverty at ages 20 to 50, and a 3 percentage-point reduction in adult incarceration. Finally, Barr and Gibbs (2017) examine the intergenerational effects of Head Start using the NLSY and two research designs: family fixed effects and variation in program *availability* across birth counties (also

⁵ Also, limited evidence shows the poorest 300 counties were more likely to get funding for Head Start (see Ludwig and Miller 2007: Table II and Pihl 2017).

referred to as “roll-out”). To alleviate concerns about the endogeneity of funding levels and measurement error in the National Archives data, their roll-out design uses a binary measure of Head Start access that is equal to one if funding exceeds the 10th percentile of observed funding per four-year-old. They find evidence of large first-generation effects on women (including a gain of a half a year of schooling) and large second-generation effects on their children’s high school graduation and completed education.

III. DATA AND RESEARCH DESIGN

This study combines the long-form 2000 Census and 2001-2013 ACS with the SSA’s Numident file to shed new light on Head Start’s long-term effects. The Census/ACS data represent almost one quarter of the U.S. population and are four orders of magnitude larger than previously used longitudinal samples. Another advantage of these combined data is that the Numident contains county of birth (rather than adulthood residence) and exact date of birth, which allows a high-quality proxy for Head Start access and age eligibility in childhood. The data’s main disadvantage is that they contain no information on family background.⁶ This lack of covariates means that we cannot model many determinants of adult outcomes—which limits precision even in this large dataset—or model treatment effect heterogeneity by childhood characteristics.

Our sample is comprised of children born from 1950 to 1980 in U.S. states where the school-entry age cutoff is known. We additionally limit our sample to individuals who are in their prime earning years (ages 25 to 54).⁷ We collapse these data to means by birth year, survey year, county of birth, and school age.⁸ We also weight our regressions using the number of observations in each cell (Solon, Haider, and Wooldridge 2015). To minimize disclosure concerns at the Census Bureau, we use only observations with non-allocated and non-missing values for all outcomes.

Our outcomes of interest are summary measures of human capital and economic self-sufficiency,

⁶ For instance, we cannot focus on a high impact sample: adults who were very poor as children and would have been much more likely to participate in Head Start. In 1970, 62 percent of Head Start’s participants were from families with annual incomes less than the poverty line for a family of four (~\$4,000) (OEO 1970).

⁷ We find no evidence that Head Start affected survival to 2000 (see Appendix Table A11).

⁸ School age is defined using exact date of birth and school-entry age cutoffs. More details are provided in the subsection below.

which permit tests of co-movements of related adult outcomes and limit the number of statistical tests (Kling, Liebman, and Katz 2007).⁹ A shortcoming of this approach is that, because indices weight each component equally, large changes in one dimension are averaged with potentially opposite-signed or zero effects in other dimensions. We, therefore, also examine the individual index components. The human capital index includes four binary variables indicating achievement of a given level of education or greater: high school or GED, some college, a 4-year college degree, and a professional or doctoral degree; years of schooling, and an indicator for working in a professional occupation. Our index of self-sufficiency includes binary indicators of employment, poverty status, income from public sources, family income, and income from other non-governmental sources; continuous measures of weeks worked, usual hours worked, the log of labor income, log of other income from non-governmental sources, and log ratio of family income to the federal poverty threshold.

A. *Measuring Exposure to Head Start*

Combining data on the launch of Head Start programs from Bailey and Goodman-Bacon (2015) with the Census/ACS-Numident permits two refinements to previously used research designs (Barr and Gibbs 2017, Johnson and Jackson 2017, Thompson 2017). First, we use only variation in the launch of the Head Start program rather than a continuous measure of Head Start spending. This refinement (1) addresses the potential endogeneity of Head Start funding levels to the program's *performance* and (2) sidesteps issues of measurement error in the National Archives grant data (Barr and Gibbs 2017).¹⁰ Second, we examine changes in outcomes for children who were *age-eligible* for Head Start (ages 3-5 or younger) relative to those who were age-ineligible (ages 6+) when it launched, allowing for the effects to vary by the number of years each cohort was potentially eligible. Age eligibility is based on exact date of birth in the Numident and school-entry age cutoffs, which alleviates measurement error in defining the potential treatment and

⁹ Following Kling, Liebman, and Katz (2007), we standardize outcome measures for each individual using the mean and standard deviation in the control group (ages 6-7 at launch), and we recode outcomes so that increases indicate improvements in human capital.

¹⁰ Thompson (2017) also tries this strategy but notes that his estimates in the NLSY are statistically insignificant.

control groups. Finally, our large dataset allows us to use state-by-birth-year fixed effects to adjust estimates for state economic and policy changes that could have affected children’s outcomes independently of Head Start.¹¹ Our identifying assumption in the analysis that follows is that the causal effect of Head Start is the only reason for a *change* in the relationship between a child’s age at the program’s launch and her outcomes as an adult. (See Appendix for more description of our sample.)

B. Event-Study Regression Model

Our research design uses a flexible event-study framework and roll-out of Head Start to estimate the effect of exposure to Head Start on long-term human capital and economic outcomes,

$$(1) \quad Y_{bct} = \theta_c + \alpha_t + \delta_{s(c)b} + \mathbf{Z}'_c b \boldsymbol{\beta} + \text{HeadStart}_c \text{Age}'_{bs(c)} \boldsymbol{\varphi} + \varepsilon_{bct}.$$

Children’s birth years are indexed by $b=1950–1980$, county of birth by c , and Census/ACS year by $t=2000–2013$. Specifications include fixed effects for county of birth, θ_c , year, α_t , and state-by-birth-year, $\delta_{s(c)b}$, which, respectively, capture time-invariant differences across counties, national changes affecting all individuals, and changes in state policies that differentially affect birth cohorts. Although covariates matter little, we follow the literature and include county characteristics, \mathbf{Z}_c interacted with a linear trend in year of birth, b (Hoynes, Page, and Stevens 2011, Bailey 2012, Bailey and Goodman-Bacon 2015).¹²

HeadStart is a binary variable equal to 1 if a child was born in a county that received a Head Start grant before 1980. *Age* is a set of dummy variables for a child’s “school age” at the time of Head Start’s launch, $1(T_c^* - b = a)$ where $a=-15$ to 30 (or $T_c^* = 1965$ and $b = 1980$ to $T_c^* = 1980$ and $b = 1950$) and T_c^* is the year Head Start began in county c . We omit school age 6 (age 6 before the school entry cut-off date), because these children would have been unlikely to have attended Head Start rather than public school.¹³ Our point estimates of interest, $\boldsymbol{\varphi}$, describe the evolution of the intent-to-treat (ITT) effects of

¹¹ Appendix Tables A12-A15 document how the estimated effects of Head Start change in our Census/ACS-Numident dataset using (1) alternative measures of access to Head Start including funding per capita and years of access to Head Start as well as (2) including state-by-birth-cohort fixed effects.

¹² County characteristics include the 1960 poverty rate, log county population, population share over age 65, under age 5, living in an urban setting, and non-white.

¹³ As an example, consider a child born October 1, 1960, in a county where Head Start started in fall of 1966. If the state’s age cutoff for turning age 6 for 1st grade entry was December 1, we would code the child as “school age” 6 in fall of 1966. However,

Head Start on long-term human capital and economic self-sufficiency. Standard errors are corrected for heteroskedasticity and adjusted for an arbitrary within-birth-county covariance structure (Arellano 1987, Bertrand, Duflo, and Mullainathan 2004).¹⁴ In our tables, we also report p-values corrected for multiple hypothesis testing using the Bonferroni-Holm method in our tables (Holm 1979, Duflo, Glennerster, and Kremer 2007).

C. *Expected Effects of Exposure to Head Start by Age at Launch*

The event-study model provides a flexible approach that imposes few restrictions on the relationship between Head Start and adult outcomes. Although economic theory does not make predictions as to the magnitudes of the event-study coefficients, the program’s phased implementation and the greater potential for some children to enroll (due to multiple years of exposure) predict a specific *pattern*. Figure 2 plots this pattern under the following set of assumptions.

First, if we assume Head Start had no effect on children who were over age 5 when the program launched, then the relationship between adult outcomes and Head Start for these children should be zero. This is the equivalent to a test for a pre-trend in our analysis and is illustrated as a flat line for children ages 6 to 12 in Figure 2.

Second, if Head Start has a positive causal effect on adult outcomes, we expect to see a *change* in those outcomes for children under age 5 when it launched, because these cohorts would have been the first to have been age-eligible *and* have access. This would not result in an immediate shift in the *level* of outcomes (akin to a regression discontinuity, RD) but rather a shift in the *slope* (akin to a regression kink, RK). The reason is that Head Start’s capacity grew over time, both because new programs were added but also as individual programs matured.¹⁵ Program quality also increased over time with better hiring and

if the state’s age cutoff for 1st grade entry was September 1, we would code the child as “school age” 5 in fall of 1966 and, therefore, age-eligible to participate in Head Start.

¹⁴ We also implement alternative standard error corrections for clustering by birth state and, separately, two-way clustering by birth-county and year (Cameron, Gelbach, and Miller 2011). Because the Census Bureau has requested that we reduce disclosures for this project and because these alternative corrections have little effect on our conclusions, we have not disclosed these additional estimates.

¹⁵ The full-year program served only 20,000 children in 1965 but 160,000 in 1966, 215,000 in 1967 and 1968, and 257,700 in 1970 (OEO 1965, 1966, 1967, 1968, 1970). Enrollment in summer Head Start was much higher, but we expect the summer program to

training of teachers, curriculum development, and the implementation of auxiliary services (e.g., health).¹⁶ Studies of other War on Poverty programs such as family planning or community health centers suggest that many of these programs reached maturity around 4 to 5 years after launch (Bailey 2012, Bailey and Goodman-Bacon 2015). Figure 2 illustrates this as the implementation curve (line with square markers), which rises from zero to 100 percent.

In addition to these gradual changes in program quality and capacity, we also expect larger effects for children who were younger when Head Start launched, simply because they would have been age-eligible for a larger share of their preschool years. For instance, a child 5 years old when Head Start launched could participate for at most one year, whereas a 3-year-old child would be age-eligible for three years. This does not mean that the 3-year-old enrolled for more than one year. However, it is more likely that a child enrolls if s/he had three years to do so. Figure 2 illustrates this cumulative potential access to Head Start as a linear relationship (dashed line with circle markers), but differences in the likelihood of enrollment by age could make this relationship more S-shaped as well (because enrollment in the early years was more likely at ages 4 than 5).

The combination of phased implementation and cumulative potential access to Head Start implies a *non-linear* change in the relationship between age at Head Start launch and adult outcomes (solid, bold line). In Figure 2's stylized example, children ages -1 or younger at Head Start's launch would have been age-eligible for a fully implemented program for each of their three years of eligibility. Assuming that Head Start *did not* continue to mature and that it *did not* have any complementarities with other War on Poverty programs, the relationship should level off for children ages -1 or younger at launch, because all cohorts born after this would have had the same potential exposure to Head Start as the -1 cohort.

Note, however, that Figure 2B shows how relaxing two assumptions implies a slightly different

have smaller effects. Even in the 1960s, few experts believed that an 8-week summer program could produce lasting benefits (Vinovskis 2008 citing Edward Zigler). Moreover, 30 to 40 percent of children in summer Head Start were aged six and older, whereas only 10 percent of those in full-year programs were older than five, which means that many of these children will be in our control group. For the age distribution of children in Head Start, see Thompson (2017)'s Table 1.

¹⁶ We suspect that the speed of implementation varied with the year of Head Start's implementation—programs starting later could adopt best practices faster. However, the rapid roll-out of Head Start programs limits our ability to test for this heterogeneity.

shape. First, allowing for effects on children ages 6 and older implies that the curve would begin to slope up *before age 6*. This is possible because 10 percent of children in full-year Head Start were 6 or older (Vinovskis 2008), and age-ineligible children could still benefit from their younger siblings' participation (Garces, Thomas, and Currie 2002). Because our subsequent analysis standardizes the effects at age 6 to zero, this relationship would appear as the flat part of the line falling below zero. Second, if the Head Start program continued to mature after 5 years or was complemented by other programs (e.g., Medicaid continued to expand into the 1970s, Goodman-Bacon (2018)), we would expect to see a *slope* for cohorts ages -1 and younger when the program launched.

D. Spline Summary Specification

Our event-study estimates impose none of the restrictions used to outline the expected effects of Head Start in Figure 2. However, we expect estimates from this flexible event-study specification to be noisy, in part because so many unobserved factors determine adult outcomes. To improve precision and test formally for trend breaks, we use Figure 2's predictions to guide the specification of a three-part spline with knots at ages 6 and -1. We implement this by replacing the $\mathbf{Age}_{bs(c)}$ in equation (1) with components of the spline in age, $a = T_c^* - b$:

$$(2) \quad Y_{bct} = \theta_c + \alpha_t + \delta_{s(c)b} + \mathbf{Z}'_c b \boldsymbol{\beta} + \text{HeadStart}_c (\mathbf{D}'_{cb} \boldsymbol{\rho}_1 + a \mathbf{D}'_{cb} \boldsymbol{\rho}_2) + \varepsilon_{bct}.$$

where \mathbf{D}'_{cb} is a vector of dummy variables, $1(-10 \leq a \leq -1)$, $1(-1 \leq a \leq 6)$, $1(6 \leq a \leq 15)$, $1(a < -10)$, and $1(15 > a)$ and the other variables remain as previously defined. We constrain the estimates of $\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2$ to ensure that the spline joins at $a=6$ and -1 . While the spline specification is more restrictive than our flexible event-study, it improves precision by imposing restrictions that should be true. The spline allows a parsimonious method to test for a pre-trend (captured in the slope of the segment for $1(6 \leq a \leq 15)$) as well as a formal trend-break test between components $1(6 \leq a \leq 15)$ and $1(-1 \leq a \leq 6)$. This final test captures whether the relationship between adult outcomes and age at Head Start's introductions *changes* at age 6—the age at which older children tend to stop participating in Head Start and begin first grade.

IV. TESTS OF IDENTIFYING ASSUMPTIONS

The research design outlined in the previous section relies on two crucial assumptions: (1) the launch of a Head Start program increased participation in Head Start and (2) the launch of a Head Start program did not coincide with other county-level changes that would affect the outcomes of preschool children. This section provides further evidence on both assumptions.

A. *How Much Did Head Start Increase Preschool Enrollment?*

While there is little doubt that introducing a Head Start program increased children's attendance in this program, the magnitude of this relationship net of crowd-out is crucial for interpreting the ITT effects recovered in equations (1) and (2). Administrative data suggest that the launch of a Head Start program significantly increased children's enrollment. The OEO reported that full-year Head Start served over 600,000 children before 1968, rising from 20,000 children in 1965, to 160,000 in 1966, to around 215,000 in 1967 and 1968 (OEO 1965, 1966, 1967, 1968, 1970).¹⁷ About 257,700 children attended full-year Head Start in 1970. Three-quarters of the children were aged 4 or 5, three-quarters were nonwhite, and 62 percent came from families with less than \$4,000 in annual income. Between 1971 and 1978, enrollment and directory information suggest that the average county with a Head Start program served roughly 309 children. These sources imply that the average Head Start program served from about 10 percent of resident age-eligible children in 1971 to 15.8 percent in 1978.

If Head Start substituted for private preschool for some children (Cascio and Schanzenbach 2013, Kline and Walters 2016, Bassok, Fitzpatrick, and Loeb forthcoming), administrative data may overstate the role of Head Start programs in increasing exposure to preschool. To examine this possibility, we use the 1970 Census, which was the first to ask children *younger than age 5* about school enrollment as of February 1—a date *during* the school year, which should capture enrollment in *full-year* Head Start.¹⁸ The Census

¹⁷ Enrollment in summer Head Start was much higher, but we expect the summer program to have smaller effects than full-year exposure to a Head Start program. Even at the beginning of the program, few experts on the planning committee believed that an 8-week summer program could produce lasting benefits (Vinovskis 2008 citing Edward Zigler). Moreover, in this period, 30 to 40 percent of children in summer Head Start were aged six and older, whereas no more than 10 percent of those in full-year programs were older than five. See also Table 1 in Thompson (2017).

¹⁸ Although the 1960 Census asked about school enrollment (including kindergarten), the question was only asked of children ages

data show that four-year-old children in counties without Head Start programs were 3.4 percentage points less likely to be enrolled in school (16.8 versus 20.2 percentage points, see Figure A1). Five-year-old children were 17 percentage points less likely to be enrolled in school (48.9 versus 65.9 percentage points). These gaps are 5.9 percentage points among 4 year olds and 21.3 percentage points among 5 year olds when looking only at children of mothers with less than a high school education.¹⁹

We use a linear probability model to adjust these gaps for state fixed effects (to account for age-invariant, state-level factors that determine the local supply of preschools) and 1960 county characteristics (share of county population in urban areas, in rural areas, under 5 years of age, 65 or older, nonwhite, with 12 or more years of education, with less than 4 years of education, in households with income less than \$3,000, in households with incomes greater than \$10,000, local government expenditures, income per capita, and whether the county was among the 300 poorest counties). Details are presented in Appendix section 6 and summarized here for parsimony. The regression results show that school enrollment was 14.9 percentage points higher for all five year olds, 15.1 percentage points higher for boys, and 14.5 percentage points higher for girls (Appendix Table A5). These results are highly robust to the inclusion (or exclusion) of different covariates.

Consistent with crowd-out being minimal, the 14.9 percent increase in enrollment in the Census is only slightly smaller than the 15.8 percent contained in administrative data. This estimate is also comparable to other studies. Garces, Thomas, and Currie (2002) estimates the national Head Start participation rate was between 10 percent and 17 percent for the 1964 to 1970 cohorts in the PSID (p. 1002). A slightly higher estimate comes from Ludwig and Miller (2007), who estimate that children's enrollment in Head Start was 17 percentage points higher at the 300-poorest county discontinuity in the 1988 National Educational Longitudinal Survey (NELS).

5 and older which precludes analysis of preschool aged children.

¹⁹ Note that Head Start was not exclusively for poor kids in the 1960s and 1970s. To encourage interaction between poor children and those from less disadvantaged backgrounds, OEO policy allowed 15 percent, and later 10 percent, of children to come from families that did not meet its poverty criteria. Roughly two-thirds of children in the full-year 1969 and 1970 programs came from families in which the mother had less than a high school education, although the mothers of about 7 percent of children had attended or graduated from college.

Based on this evidence, we use our best estimate of 0.149 to transform the ITT effects from our event-study and spline specifications into average treatment-effects-on-the-treated (ATET). We also construct confidence intervals using a parametric bootstrap procedure with 10,000 independent draws from normal distributions with means and standard deviations equal to the point estimates and standard errors from the reduced-form and first-stage estimates (Efron and Tibshirani 1993).

B. Did Head Start’s Launch Correspond to Other Policy Changes?

Another key assumption of our analysis is that the launch of a Head Start program is the only reason why a cohort’s adult outcomes changed for age-eligible children relative to 6+ year olds. The decentralization of U.S. governance and the process of applying for OEO grants makes it unlikely that communities across the nation coordinated in their grant applications. More worrisome is that OEO could have provided multiple grants to a community in the same year, making it difficult to disentangle the effects of Head Start from other OEO programs. To test this hypothesis, we use information on other War on Poverty programs compiled from the National Archives and estimate regressions similar to equation (1). In particular, we replace the dependent variable with an indicator for receiving funding in county c in fiscal year t , or $Y_{ct} = \theta_c + \delta_{s(c)t} + \mathbf{X}'_{ct}\boldsymbol{\beta} + \sum_k^K \pi_k 1(t - T_c^* = k) + \varepsilon_{cbt}$. We also include county and state-year fixed effects, θ_c and $\delta_{s(c)t}$, and county level covariates, \mathbf{X}_{ct} , as we do in our primary regression specifications. Our variable of interest is event-time, $t - T_c^* = k$, the year of observation relative to the date Head Start launched (the year before Head Start began, $k = -1$, is omitted).

Figure 3 shows the relationship between the launch of Head Start and the launch of other OEO programs. As expected, 100 percent of treated counties in our sample first received a Head Start grant in event-year 0. This is by design. In subsequent years, the share of counties receiving a Head Start grant tapers off to 70 percent after around five years—this reflects the fact that some counties received multi-year grants and also the fact that not all of the early programs continued.

For our estimates to be confounded by changes in other federal funding, these funding changes would need to happen around the time Head Start launched, or event-year 0. However, our analysis of Food Stamps, Community Health Centers, and other child health programs show no such pattern. The one

program that shows a small change in funding after Head Start began is the CAP health project. Importantly for our inferences, funding was very small for this program (around \$8 in 2013 dollars per person; by contrast, annual Head Start funding was nearly \$1,500 per 4-year-old during the same period). Furthermore, there is little reason to think this program would affect children who were 5 or younger but not those who were 6. Although we cannot rule out funding changes in programs we do not measure, these patterns provide little evidence that our research design is confounded by other OEO programs.

V. HEAD START'S EFFECTS ON HUMAN CAPITAL

Figure 4 plots the event-study estimates for all outcomes in the human capital index for the set of compositionally balanced county-birth-cohorts,²⁰ or individuals ages 15 to as young as -1 when Head Start launched. The solid line plots the event-study estimates. Consistent with the patterns anticipated in Figure 2, the human capital index and each of its components exhibit little relationship to adult outcomes for cohorts ages 6 to 15 when Head Start launched (i.e., there is little evidence of a pre-trend among ineligible cohorts). However, the index and many of its subcomponents exhibit a trend-break around age 6, suggesting that access to Head Start improved the human capital of adults. Notably, this relationship is not as sharp as in regression kink designs, because (1) school age-entry cutoffs were not strictly enforced, (2) older children ages 6 and 7 participated in Head Start (although at lower rates), and (3) older siblings of participants may have benefitted indirectly from their younger sibling's involvement. Any benefits of Head Start for these other ages would lead the event-study estimates for ages higher than 6 to fall slightly below zero, as the data show. They also weaken the visual evidence and formal tests for a trend-break at exactly age 6.

Table 1 summarizes both the event-study and spline estimates at -1. Column 1 presents the mean and standard deviation of the outcome for cohorts ages 6 and 7 at the time of Head Start's launch (our control group). Column 2, which shows the ITT-event-study estimate at -1 (our estimate for cohorts that were age eligible for up to three years for a fully implemented program), suggests that Head Start significantly improved adult human capital. The standardized index increases by 1.5 percent of a standard deviation for the fully exposed cohort and 10 percent of a standard deviation for treated children (column

²⁰ See Appendix Table A4 for a table describing program roll-out by cohort and age.

6). Across outcomes, column 3's ITT-spline estimates are identical to those in column 2 to the hundredth.

Supporting the impression left by the event-study plots, column 4 shows that a formal test for pre-trend (ages 6 to 15), the slope estimate for the spline component for ages 6 to 15). Figure 5A additionally plots the magnitude of the t-tests for the pre-trend. For the index and for each of its subcomponents, the data find no evidence of a pre-trend—a conclusion strengthened by Bonferroni-Holm (BH) p-value adjustments for multiple hypothesis testing. T-tests for pre-trends fail to reject the slope is zero in all cases. However, column 5 of Table 1 and Figure 5B shows that the data reject the null hypothesis of no trend-break at age 6 for the adult human capital index at the 1-percent level. With the exception of high school graduation (which is just below the threshold for statistical significance at the 10-percent level), the data show evidence of a trend break for each of the subcomponents of the human capital index. It appears that, even with Head Start's spill-overs to children older than 6, the relationship between adult human capital *changed* for children age-eligible for Head Start relative to children in the same county who were old enough to enroll in first grade.

The data show strong effects on some of the most commonly studied outcomes in the preschool literature, including both high school graduation and college enrollment. Figure 4B and Table 1 show that treated children were 1.9 percentage points more likely to complete high school/GED (column 6)—a 2.1-percent increase relative to the control mean (column 7). The magnitude of this estimate is precisely estimated, but smaller than other estimates of Head Start's effects in the literature. Figure 6A shows that the effect is roughly half the size of Garces, Thomas, and Currie (2002)'s sibling comparison in the PSID and Thompson (2017)'s spending design in the NLSY. In addition, it is one-fifth the size of Johnson and Jackson (2017)'s spending design estimates for the very disadvantaged sample in the PSID; and one-ninth the size of Ludwig and Miller (2007)'s RD estimates using the Census. (It is one quarter the size of Deming et al. (2009)'s sibling comparison for Head Start in the 1990s for more recent cohorts.) Although our estimate falls within the confidence intervals of previous studies, this reflects the imprecision of those estimates.

Figure 4C and Table 1 also shows a statistically significant effect of Head Start on college

enrollment. Head Start raised college enrollment by 5.4 percentage points, or 8.7 percent. This estimate is half the size of Garces, Thomas, and Currie (2002) and one quarter the size of Ludwig and Miller (2007) (Figure 6B). (The magnitude of the increase in college enrollment of 0.05 is only slightly smaller than Deming (2009)'s NLSY sibling comparison for Head Start in the 1990s.) Again, consistent with the visual impression of a trend break in Figure 4C, we find no evidence of a pre-trend for children older than 6 at launch and reject the null hypothesis of no trend break at age 6 at the 5-percent level.

In addition to generating more precise estimates for these commonly studied outcomes, our large-scale data permit a novel evaluation of the effects of Head Start on other dimensions of human capital, including college completion or higher degrees, which previous data have not been able to detect. Table 1 shows that participating children achieved 19 percent higher college graduation rates (the trend break is statistically significant at the 1-percent level).²¹ These estimates are one-quarter to one-fifth the size of those found for the Abecedarian Project (Currie 2001, Barnett and Masse 2007, Duncan and Magnuson 2013). Similarly, completion of professional or doctoral degrees increased by 50 percent among treated children, although evidence of a significant pre-trend caution against a causal interpretation (this is also consistent with Figure 4F). These gains across the education distribution are summarized in a 0.29-year increase in schooling. This estimate is smaller than Johnson and Jackson (2017)'s estimate of 0.52 years for very disadvantaged children, but it is highly statistically significant and is not driven by a pre-trend.²²

These large effects on college and higher degrees may be surprising, given that no other study of preschool has documented effects on post-secondary education. This lack of evidence may reflect, in part, the small longitudinal samples or the small scale of model preschool programs. Differences in the participating children may also matter. Abecedarian and Perry's participants were *very* disadvantaged children and mostly black, and Perry's participants had low IQs.²³ In contrast, Head Start was not

²¹ The similarity of column 2 for some college and college completion is correct. This does not imply that every person that Head Start induced to enter completed college. Head Start likely helped some individuals enter and others complete college.

²² Jackson and Johnson's ITT estimate is 0.0967 per \$1000 spent per poor 4-year-old. They translate this into an ATET by multiplying the coefficient by 4 (the average Head Start spending per poor 4-year-old measured in thousands) and dividing by 0.75 (their estimate of take-up among income-eligible 4-year-olds in counties with Head Start programs), so that $0.0967 * 4 / 0.75 = 0.52$

²³ The model Perry Preschool Program, which focused on lower IQ children, had no measured effects on postsecondary outcomes (Anderson 2008).

exclusively for poor, African-American, or low-IQ children. Consequently, Head Start's participants in the 1960s and 1970s likely faced fewer socio-economic and cognitive disadvantages and less racism relative to model programs. Differences in the background characteristics of Head Start's participants make it less surprising that they experienced gains in post-secondary education.

Because analyses of model preschool programs have found different educational effects for boys, Table 2 stratifies our sample by sex. Among participating men, the human capital index increased by a statistically significant 14 percent of a standard deviation. For this group, high school completion rose by a statistically insignificant 2.7 percent, college attendance rose by 13 percent, and college completion rose by 27 percent. The high school estimates are smaller than others in the literature, but the college attendance estimates tend to be larger. Head Start cumulatively raised years of education among treated men by 0.41 years and the likelihood of completing a professional/doctoral degree by 59 percent. The evidence suggests that men treated with Head Start were 19 percent more likely to hold professional jobs.

The human capital index increased by less among women, at only 7 percent of a standard deviation. Completion of high school (or a GED) rose by a statistically insignificant 1.5 percent, and college attendance rose by 5.7 percent (although the trend-break is not statistically significant). For women, changes in the human capital index appear driven by increases in higher degrees, including an 11-percent increase college completion and 36-percent increase in professional degrees. Treated women's schooling rose by 0.17 years and their likelihood of holding a professional job rose by 9.5 percent.

Our Appendix Tables A7-A8 reports estimates of Head Start's effects on human capital by race. Unfortunately, even our large sample size is too small to precisely quantify effects by race, because less than 1/6 of our sample is nonwhite and, unlike longitudinal data, we have no background covariates to model the many other determinants of adult outcomes. The broad patterns in these estimates suggest that Head Start's effects are largest among white men (13 percent of a standard deviation) and smaller among white and non-white women (5-6 percent of a standard deviation, respectively). Effects for non-white men were generally small and imprecise.

VI. HEAD START'S EFFECTS ON ECONOMIC SELF-SUFFICIENCY

The substantial effects of Head Start on human capital suggest a potential for effects on economic

self-sufficiency. Figure 7A plots the event study estimates for the self-sufficiency index, and Table 3 shows that an index of economic self-sufficiency aggregated over both sexes was by 4 percent of a standard deviation higher than for children ages 6-7 at the time Head Start began. Consistent with Head Start affecting less skilled individuals, the program decreased the likelihood of adult poverty by 12 percent and receipt of public assistance income by 29 percent, though these results are imprecise. Figure 7B and 7C show striking evidence of a trend break at age 6, which is also reflected in column 5 of Table 3. However, Table 3 shows little effect of Head Start on labor-force participation or wage income. (We omit the event study estimates for these outcomes for parsimony, because they are noisy and show no effect.) Null effects for wages and labor-force participation also affect the magnitude of the change and the trend-break in the economic self-sufficiency index, which does not exhibit the sharp trend-break at age 6 as in the human capital outcomes (see also Figure 6B). This result may reflect the fact that men's and women's work effort changed in offsetting ways, resulting in selection for both groups. Whereas Head Start's effect on men's human capital may have led them to increase employment (e.g., the substitution effect dominates), the reverse may be true for women (e.g., the income effect dominates as more education allows them to marry higher-earning men).

Table 4 stratifies our sample by sex and provides evidence consistent with this hypothesis. Because the self-sufficiency estimates are noisier and stratifying by sex reduces sample sizes, we focus our discussion on the spline estimates. For treated men, the self-sufficiency index increased by 3 percent of a standard deviation. We also find positive effects of Head Start exposure on both the extensive and intensive margins of men's labor-force participation. Treated men were 2.1 percent more likely to have worked for pay (column 7), worked an average of one more week and one more hour per week (column 6). Consistent with these estimates reflecting the causal effect of Head Start, we find no evidence of a pre-trend and a marginally significant trend-break at age 6, which does not survive the Bonferroni-Holmes standard error correction. At first glance, it is curious that the combined effects of increased human capital and labor-force participation do not appear to have affected annual wages. Upon further investigation, this appears consistent with Head Start inducing *negative* selection into the labor-force: the marginal participants tended

to be less skilled, and therefore lowered the cohort's wages on average.²⁴ Head Start had little effect on men's poverty, but the program is associated with a 27-percent decline in public assistance receipt among treated men. Reductions in public assistance are also consistent with *negative* selection, because male public assistance recipients receive high rates of disability income.²⁵

The pattern is different for women. The self-sufficiency index increased by 4 percent of a standard deviation among women treated with Head Start, largely driven by a 28-percent reduction in public assistance receipt and a 16-percent reduction in poverty. However, women's labor-force participation on the extensive and intensive margins *fell* slightly, albeit not significantly. These reductions in work appear to have *increased* annual wages of working women by around 4 percent, which is consistent with Head Start inducing *positive* selection (e.g., less-skilled women opting out). In our conclusion, we examine the effects of negative selection among men and positive selection among women.

As with the human capital outcomes, our Appendix Tables A9-A10 report estimates of Head Start's effects on self-sufficiency by race and sex. Similarly, the effects for nonwhites are generally statistically insignificant, owing to the fact that nonwhites comprise less than 15 percent of the sample and that we have few background characteristics to model the considerable variation in outcomes. As with human capital, the patterns of these estimates suggest that Head Start's effects on self-sufficiency are largest among white men and women (3 and 4 percent of a standard deviation, respectively).

VII. HETEROGENEITY IN HEAD START'S LONG-RUN EFFECT

This final section seeks to shed light on the potential mechanisms for Head Start's effects by examining how the estimates vary with access to other public programs and local economic conditions. We implement this analysis by interacting a binary indicator for whether cohorts lived in counties with "high" or "low" exposure to a program with the spline components in equation (2), where "high" is equal to one for counties above the median in the characteristics and 0 otherwise. We use only the main indices for

²⁴ Assuming all new labor market entrants came from the left-hand tail of the skill distribution, the NLSY shows a negative impact of such selection on wages is 7.4 percent. Because we find a 2.2 percent increase in log annual wages, this implies that Head Start led to an estimated 9.6-percent increase. This number is consistent with our NLSY estimate that men's potential wages increased by 9.4 percent (see Appendix Table A6).

²⁵ We find no evidence of decreases in incarceration among men (see Appendix Table A11).

human capital and economic self-sufficiency as the dependent variables. We caution that the lack of randomization of alternative programs and conditions means that, while these relationships are suggestive, they should not be interpreted as causal. Additionally, uncertainty about how program enrollment varied means these estimates are less precise than desired. We nevertheless provide additional guidance about the magnitudes of the ATET effects using our estimates of differential take-up across locations to scale our ITT estimates.

We first investigate the hypothesis that Head Start's long-run effects relate to their complementarities with other health programs for disadvantaged children. If health screening and referrals to health services (a sizable share of Head Start's budget) played a role in driving long-term effects, we would expect Head Start's effects to be larger for children with greater access to these health services through community health centers (CHCs) and/or Medicaid.²⁶ (We would also expect the selection effects on wages to mirror those we observe for men in Table 4.) Table 5 provides suggestive evidence of this mechanism, showing that the ATETs of Head Start for human capital were slightly larger in areas with CHCs and three times as large in states where more children were eligible for Medicaid (17 percent increase relative to 5 percent for less-exposed children). The 95-percent confidence interval in the difference between these two treatment effects on the treated suggests we can reject the null hypothesis of equal effects. Consistent with health services bringing more previously disabled workers into the labor market, Head Start's ATETs on economic self-sufficiency are more muted in locations with greater health services.

A related hypothesis is that Head Start affected adult outcomes by providing healthy meals and snacks, improving child nutrition which increased both health and learning. If nutrition is an important mechanism for Head Start's long-run effects, we would expect the program's effects to be *smaller* for children with greater access to Food Stamps, which also supported the provision of health meals.²⁷

²⁶ To construct high and low CHC exposure, we use data from Bailey and Goodman-Bacon (2015). We use the first year in which the program began in the county to construct the number of person-years an individual in each county and cohort would have been exposed to CHCs between ages 0 and 5. For Medicaid, we use Goodman-Bacon's (2018) data on the year states adopted Medicaid and the share of children covered by Aid for Families with Dependent Children (AFDC) in the year of Medicaid's launch. Because AFDC recipients were categorically eligible for Medicaid, coverage rates are strongly correlated with take-up of Medicaid in the early years. By combining the year of launch with the child AFDC rate for each county and cohort, we can obtain a measure of person-years of access to Medicaid.

²⁷ We use data from Almond et al. (2011). Access to Food Stamps is constructed in the same manner as access to CHCs.

Consistent with this, Table 5 shows that children participating in Head Start with more access to Food Stamps experienced smaller—although still statistically significant—gains in human capital and economic self-sufficiency. This suggests that the Food Stamps program provided a partial *substitute* for Head Start’s nutritional component.

The OEO’s larger effort to set up Head Start programs in the poorest 300 counties could also lead the Head Start program to be more intensive in these areas, potentially having larger effects. For this test, we report the poorest 300 counties in the column for “above median” and report the effects for counties outside this group in the column “below median.” While Table 5 shows little evidence of differential effects on human capital across these two groups of counties, there is suggestive evidence that children in poorer counties benefitted more in terms of their economic self-sufficiency—although this difference is imprecise.

A final hypothesis is that Head Start’s effects should be larger in areas with greater subsequent economic growth. Strong economic conditions should both increase the resources of children’s parents, the provision of public goods (such as schools), and create stronger incentives for children to invest in themselves, as children could expect higher and more certain returns. Rather than using actual economic growth (that may be endogenous), we use predicted economic growth between 1965 and 1985.²⁸ Table 5 suggests that the benefits of a strong economy complement Head Start’s effects. The ATETs of Head Start for human capital were twice as large in areas with strong predicted economic growth than in areas with weaker predicted economic growth. The ATETs of Head Start for economic self-sufficiency were fifty percent larger large in areas with strong predicted economic growth.

All in all, these results suggest that Head Start’s long-run effects may be driven by many factors beyond a preschool curriculum, including health screenings and referrals and more nutritious meals for an a population with otherwise little access to health care and under-nourished. Unsurprisingly, the effects of Head Start appear to be complementary to the family and public resources arising from a stronger economy.

²⁸ We use data from the Bureau of Economic Analysis (BEA) and County Business Patterns (CBP). We create a county-level panel of log real earnings from 1965 through 1985. We then regress growth in real earnings between 1965 and 1985 on a number of county characteristics from the 1960 Census: log of total population, share of a county in farmland, share of the population living in an urban setting, share black, share under age 5, share over age 65, and share living in poverty. We use predicted growth from this regression to select counties that were likely to have high economic growth from 1965-1985 and those that were not.

VIII. NEW EVIDENCE ON THE LONG-TERM RETURNS TO HEAD START

Over the past 20 years, substantial evidence has accumulated that *model* preschool programs have sizable economic returns (Almond and Currie 2011, Duncan and Magnuson 2013, Heckman et al. 2010, Cunha and Heckman 2007). However, convincing evidence on the long-run returns to larger-scale, public preschool has remained sparse (Phillips et al. 2017).

Using large-scale restricted Census/ACS data, this paper provides new evidence of the long-term effects of Head Start, the nation's longest-running, large-scale public preschool program. We find that Head Start had large effects on participants' human capital. Head Start children achieved 0.29 more years of schooling, reflecting the fact that they were 2.1 percent more likely to complete high school, 8.7 percent more likely to enroll in college, and 19 percent more likely to complete college. A second finding is that Head Start increased adult self-sufficiency, reducing the likelihood of adult poverty by 12 percent and public assistance receipt by 29 percent. Heterogeneity tests suggest that these long-run effects may reflect many aspects of the Head Start program beyond its curriculum: health screenings and referrals and more nutritious meals appear to be important mechanisms for the program's effects on disadvantaged children. In addition, the effects of Head Start appear to be complement greater family and public resources arising from a stronger economy.

A full accounting of the costs and benefits of Head Start is beyond the scope of this paper, but we summarize the implications of our estimates using *potential* earnings to account for selection (Neal and Johnson 1996, Deming 2009). Following Neal and Johnson (1996) and Deming (2009), the advantage of this approach is that it allows us to account for the effects of Head Start on employment (which differed for women and men). Like Deming, we use the NLSY79 to predict wages for individuals born from 1957 to 1965 (ages 14 to 22 in 1979)—a time frame that overlaps our Census/ACS analysis. The NLSY data allow us to estimate the relationship between wages and components of the human capital and economic self-sufficiency indices *after* flexibly controlling for ability using the AFQT. Although AFQT is not available in the Census/ACS, using this as a covariate helps mitigate omitted variables bias in ability in the education and earnings relationship. These regressions are reported in our Appendix Table A6. After accounting for ability, the NLSY79 suggest a private internal rate of return to Head Start of 7.7 percent, which ranges from

4 percent for women to 11 percent for men. As an alternative, the internal rate of return of putting one child through Head Start is 2.4 percent using only savings on public assistance expenditures (estimated at \$9,967 in the *Survey of Income and Program Participation*).

Several reasons suggest that these estimates are conservative. First, our research design differences out sibling spill-over effects, which tends to reduce the estimated effect sizes. Second, reports of income and public assistance receipt may be severely underreported in major national surveys (Meyer, Mok, and Sullivan 2015, Bound, Brown, and Mathiowetz 2001), suggesting estimates of Head Start's effect on public assistance may be understated. Third, adding increases in tax revenues, reductions in deadweight loss from public assistance transfers, or underreporting in public assistance income would serve to increase our estimates of the returns to Head Start. Finally, estimates of the returns to Head Start ignore benefits through improvements in outcomes not measured here. For instance, they ignore the extent to which more education engenders better health, longevity, or well-being. These potential limitations, however, tend to strengthen the conclusion that Head Start achieved its goal of reducing adult poverty, delivering sizable returns to investments made in the 1960s and 1970s. The results suggest potentially larger social returns.

The long-run returns to today's public preschool programs may be different for a number of reasons. Today, the curriculum is different, the target population is different, and the alternative programs and resources available to poor children are radically different than in the 1960s. Of course, researchers will need to wait another 50 years to evaluate the long-run effects of today's preschool programs. In the meantime, the sizable returns to the "less-than-model" Head Start preschool program of the 1960s suggest productive avenues for improving the lot of poor children today.

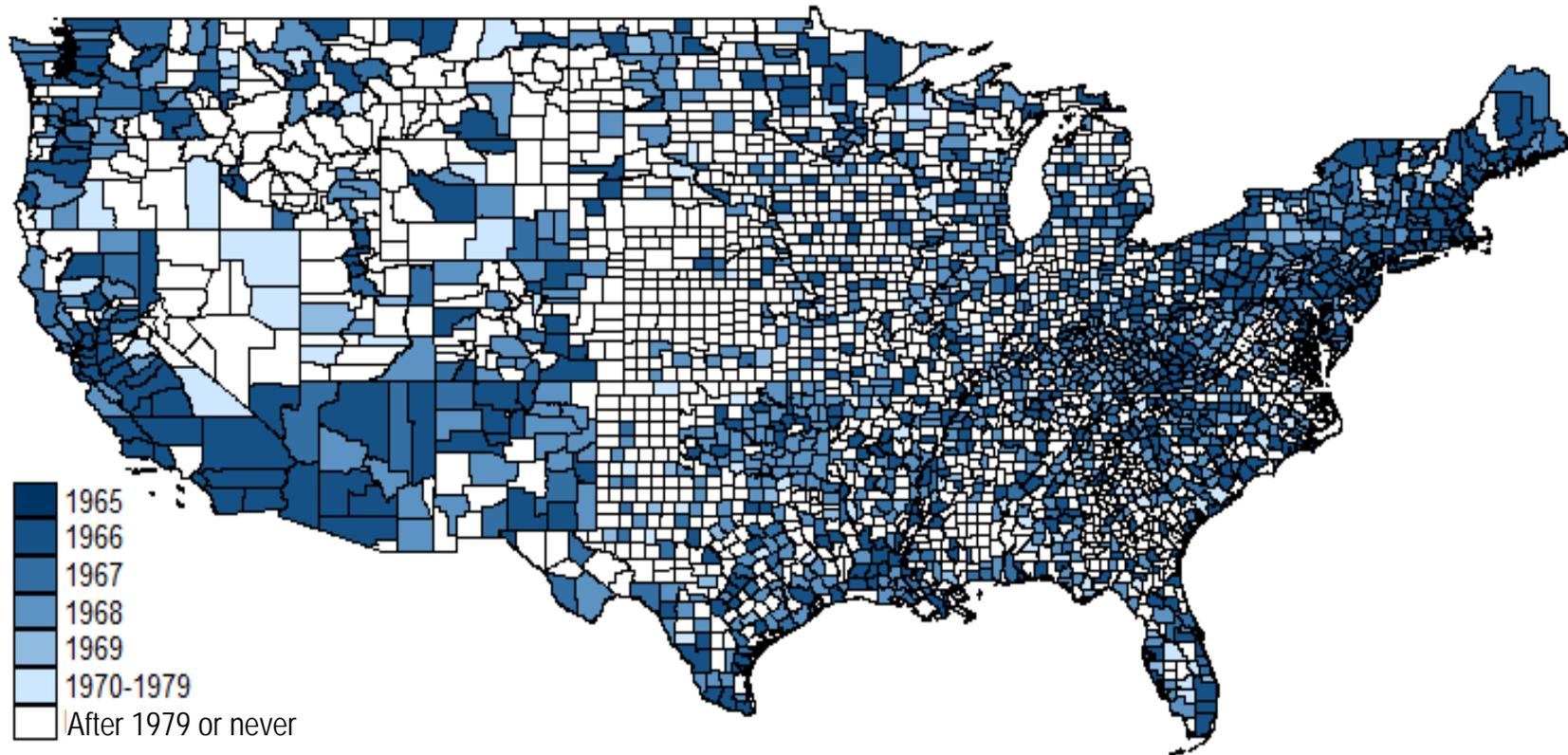
IX. REFERENCES

- Almond, Douglas, and Janet Currie. 2011. "Human capital development before age five." *Handbook of labor economics* 4:1315-1486.
- Almond, Douglas, Hilary W. Hoynes, and Diane Whitmore Schanzenbach. 2011. "Inside the War on Poverty: The Impact of Food Stamps on Birth Outcomes." *Review of Economics and Statistics* 93 (2):387-403.
- Anderson, Michael Anderson. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484):1481-1491.
- Arellano, Manuel. 1987. "Computing Robust Standard Errors for Within-Groups Estimators." *Oxford Bulletin of Economics and Statistics* 49 (4):431-434.
- Bailey, Martha J. 2012. "Reexamining the Impact of U.S. Family Planning Programs on Fertility: Evidence from the War on Poverty and the Early Years of Title X." *American Economic Journal: Applied Economics* 4 (2):62-97.

- Bailey, Martha J., and Sheldon Danziger. 2013. "The Legacies of the War on Poverty." In *Legacies of the War on Poverty*, edited by Martha J. Bailey and Sheldon Danziger. New York: Russell Sage Foundation.
- Bailey, Martha J., and Nicolas J. Duquette. 2014. "How Johnson Fought the War on Poverty: The Economics and Politics of Funding at the Office of Economic Opportunity." *Journal of Economic History* 74 (2):351-388.
- Bailey, Martha J., and Andrew J. Goodman-Bacon. 2015. "The War on Poverty's Experiment in Public Medicine: The Impact of Community Health Centers on the Mortality of Older Americans." *American Economic Review* 105 (3):1067-1104.
- Barnett, W., and Leonard N. Masse. 2007. "Comparative benefit-cost analysis of the Abecedarian program and its policy implications." *Economics of Education Review* 26 (1):113-125.
- Barr, Andrew, and Chloe R. Gibbs. 2017. "Breaking the Cycle? Intergenerational Effects of an Anti-Poverty Program in Early Childhood." https://curry.virginia.edu/sites/default/files/files/EdPolicyWorks_files/61_Anti_Proverty_Effects_ECE.pdf.
- Bassok, Daphna, Maria Fitzpatrick, and Susanna Loeb. forthcoming. "Does State Preschool Crowd-Out Private Provision? The Impact of Universal Pre-Kindergarten on the Childcare Sector in Oklahoma and Georgia." *Journal of Urban Economics*.
- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119 (1):249-275.
- Bloom, Benjamin S. 1964. *Stability and Change in Human Characteristics*. New York: Wiley.
- Bound, John, Charles Brown, and Nancy Mathiowetz, eds. 2001. *Measurement Error in Survey Data*. Edited by James J. Heckman and Edward Learner. Vol. 5, *Handbook of Econometrics*. Amsterdam: Elsevier.
- Bound, John, and Gary Solon. 1999. "Double Trouble: On the Value of Twins-Based Estimation of the Return to Schooling " *Economics of Education Review* 18 (2):169-182.
- Braun, Samuel J., and Esther P. Edwards. 1972. *History and Theory of Early Childhood Education*. Belmont, CA: Wadsworth.
- Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2011. "Robust Inference With Multiway Clustering." *Journal of Business & Economic Statistics* 29 (2):238-249. doi: 10.1198/jbes.2010.07136.
- Card, David, David S. Lee, Zhuang Pei, and Andrea Weber. 2015. "Inference on Causal Effects in a Generalized Regression Kink Design." *Econometrica* 83 (6):2453-2483.
- 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* 6 (4):135-173. doi: 10.1257/pol.6.4.135.
- Cascio, Elizabeth U. 2009. "Maternal Labor Supply and the Introduction of Kindergartens into American Public Schools." *Journal of Human Resources* 44 (1):140-70.
- Cascio, Elizabeth U., and Diane Whitmore Schanzenbach. 2013. "The Impacts of Expanding Access to High-Quality Preschool Education." *Brookings Papers on Economic Activity*:127-178.
- Cunha, Flavio, and James Heckman. 2007. "The Technology of Skill Formation." *American Economic Review* 97:31-47.
- Currie, Janet. 2001. "Early Childhood Education Programs." *Journal of Economic Perspectives* 15 (2):213-238.
- Currie, Janet, and Duncan Thomas. 1995. "Does Head Start Make a Difference? ." *American Economic Review* 85 (3):341-364.
- Deming, David. 2009. "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start " *American Economic Journal: Applied Economics* 1 (3):111-134.
- Duflo, Esther, Rachel Glennerster, and Michael Kremer. 2007. "Chapter 61 Using Randomization in Development Economics Research: A Toolkit." In *Handbook of Development Economics*, edited by T. Paul Schultz and John A. Strauss, 3895-3962. Elsevier.
- Duncan, Greg J., and Katherine Magnuson. 2013. "Investing in Preschool Programs." *Journal of Economic Perspectives* 27 (2):109-32.
- Efron, Bradley, and Robert J. Tibshirani. 1993. *An Introduction to the Bootstrap*. New York: Chapman & Hall.
- Espinosa, Linda M. 2002. "High-Quality Preschool: Why We Need It and What it Looks Like." *NIEER Preschool Policy Matters* 1.
- Fosburg, Linda B., Nancy N.; Goodrich, Mary K.; Fox, Patricia; Granahan, Janet; Smith, John H.; Himes, and Michael; Weitzman. 1984. "The Effects of Head Start Health Services: Report of the Head Start Health Evaluation." In *Report Prepared for the Administration for Children, Youth and Families, U. S. Department of Health and Human Services*. Cambridge, MA: Abt Associates.
- Frisvold, David E. 2015. "Nutrition and cognitive achievement: An evaluation of the School Breakfast Program." *Journal of Public Economics* 124:91-104. doi: <https://doi.org/10.1016/j.jpubeco.2014.12.003>.
- Garces, Eliana, Duncan Thomas, and Janet Currie. 2002. "Longer Term Effects of Head Start." *American Economic Review* 92 (4):999-1012.
- Gibbs, Chloe, Jens Ludwig, and Douglas L. Miller. 2014. "Does Head Start Do Any Lasting Good?" In *Legacies of the War on Poverty*, edited by Martha J. Bailey and Sheldon Danziger, 39-65. New York, NY: Russell Sage Foundation.
- Goodman-Bacon, Andrew J. 2018. "Public Insurance and Mortality: Evidence From Medicaid Implementation." *Journal of Political Economy* 126 (1):216-62.
- Griliches, Zvi. 1979. "Sibling Models and Data in Economics: Beginnings of a Survey " *Journal of Political Economy* 87 (5, Part 2):S37-S64.
- Grosz, Michel Z., Douglas L. Miller, and Na'ama Shenhav. 2017. "All In the Family: Assessing the External Validity of Family Fixed Effects Estimates and the Long-Term Impact of Head Start." Accessed July 31, 2017. http://conference.nber.org/confer/2017/SI2017/CH/Miller_Shenhav_Grosz.pdf.
- Hechinger, Fred M., ed. 1966. *Pre-School Education Today: New Approaches to Teaching Three-, Four-, and Five-Year-Olds*. New York: Doubleday.
- Heckman, James J., Seong H. Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz. 2010. "The Rate of the Return to the

- High Scope Perry Preschool Program " *Journal of Public Economics* 94 (1):114-128.
- Holm, Sture. 1979. "A Simple Sequentially Rejective Multiple Test Procedure." *Scandinavian Journal of Statistics* 6 (2):65-70.
- Hoynes, Hilary W., Marianne E. Page, and Ann Huff Stevens. 2011. "Can Targeted Transfers Improve Birth Outcomes? Evidence from the Introduction of the WIC Program." *Journal of Public Economics* 95 (7):813-827.
- Hunt, Joseph McVicker. 1961. *Intelligence and Experience*. New York: Ronald Press.
- Johnson, Rucker, and C. Kirabo Jackson. 2017. "Reducing Inequality through Dynamic Complementarity: Evidence from Head Start and Public School Spending." *NBER Working Paper No. 23489* June.
- Kline, Patrick, and Christopher Walters. 2016. "Evaluating Public Programs with Close Substitutes: The Case of Head Start." *Quarterly Journal of Economics* 131 (4):1795-1848.
- Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental analysis of neighborhood effects." *Econometrica* 75 (1):83-119.
- Levine, Robert A. 1970. *The Poor Ye Need Not Have with You: Lessons from the War on Poverty*. Cambridge: MIT Press.
- Levitan, Sar A. 1969. *The Great Society's Poor Law: A New Approach to Poverty*. Baltimore: Johns Hopkins Press.
- Loeb, Susanna. 2016. "Missing the Target: We Need to Focus on Informal Care Rather than Pre-School." *Economic Studies at Brookings* 1 (19).
- Ludwig, Jens, and Douglas L. Miller. 2007. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design." *Quarterly Journal of Economics* 122 (1):159-208.
- Meyer, Bruce D., Wallace K. C. Mok, and James X. Sullivan. 2015. "Household Surveys in Crisis." *Journal of Economic Perspectives* 29 (4).
- Neal, Derek, and William Johnson. 1996. "The Role of Pre-market Factors in Black-White Wage Differences." *Journal of Political Economy* 104 (5):869-95.
- North, A. Fredrick. 1979. "Health Services in Head Start." In *Project Head Start: A Legacy of the War on Poverty*, edited by Edward Zigler and Jeanette Valentine. New York: Free Press.
- OEO. 1965. "1st Annual Report: A Nation Aroused."
- OEO. 1966. "2nd Annual Report: The Quiet Revolution."
- OEO. 1967. "3rd Annual Report: The Tide of Progress."
- OEO. 1968. "4th Annual Report: As the Seed is Sown."
- OEO. 1970. "Annual Report Fiscal Years 1969-70."
- Phillips, Deborah A., Mark W. Lipsey, Kenneth A. Dodge, Ron Haskins, Daphna Bassok, Margaret R. Burchinal, Greg J. Duncan, Mark Dynarski, Katherine A. Magnuson, and Weiland Christina. 2017. "The Current State of Scientific Knowledge on Pre-Kindergarten Effects." Washington, DC. Accessed April 18, 2017. https://www.brookings.edu/wp-content/uploads/2017/04/duke_prekstudy_final_4-4-17_hires.pdf.
- Solon, Gary, Steven J. Haider, and Jeffrey M. Wooldridge. 2015. "What are We Weighting For?" *Journal of Human Resources* 50 (2):301-316.
- Thompson, Owen. 2017. "Head Start's Long Run Impact: Evidence form the Program's Introduction." *Journal of Human Resources* 53 (4).
- United States Senate Committee on Labor and Public Welfare. 1964. "Economic Opportunity Act of 1964." Washington, DC: U.S. Government Printing Office.
- Westinghouse Learning Corporation. 1969. "The Impact of Head Start: An Evaluation of the Effects of Head Start on Children's Cognitive and Affective Development (Executive Summary)." Washington, D.C.: Office of Economic Opportunity,. Retrieved from: <https://eric.ed.gov/?id=ED036321>.

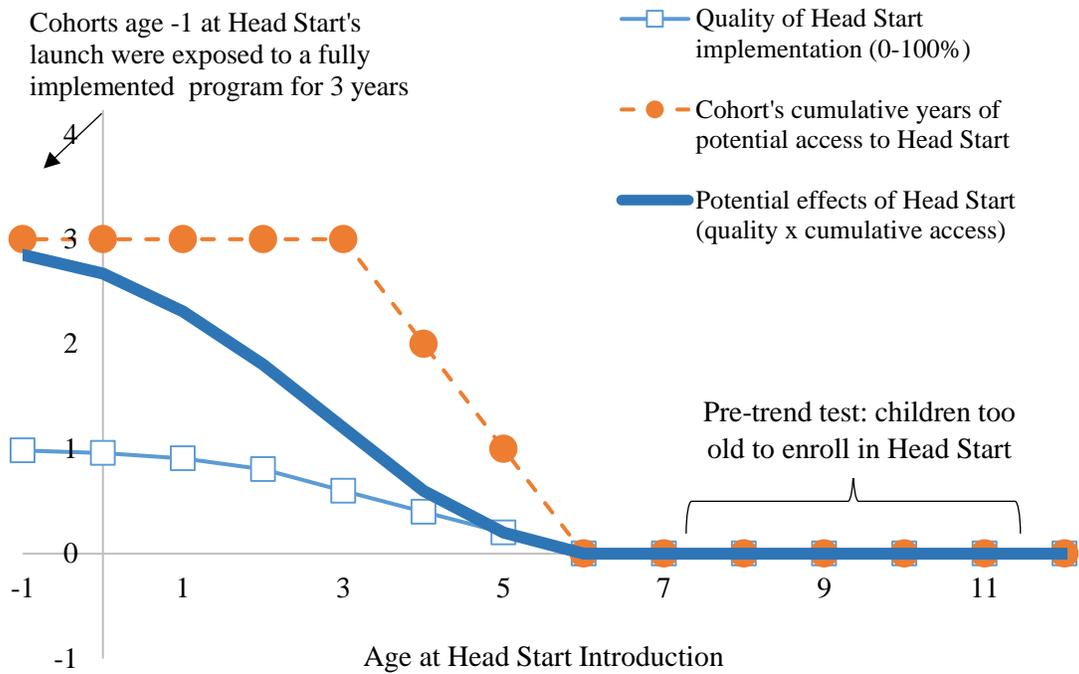
Figure 1. The Launch of Head Start between 1965 and 1980



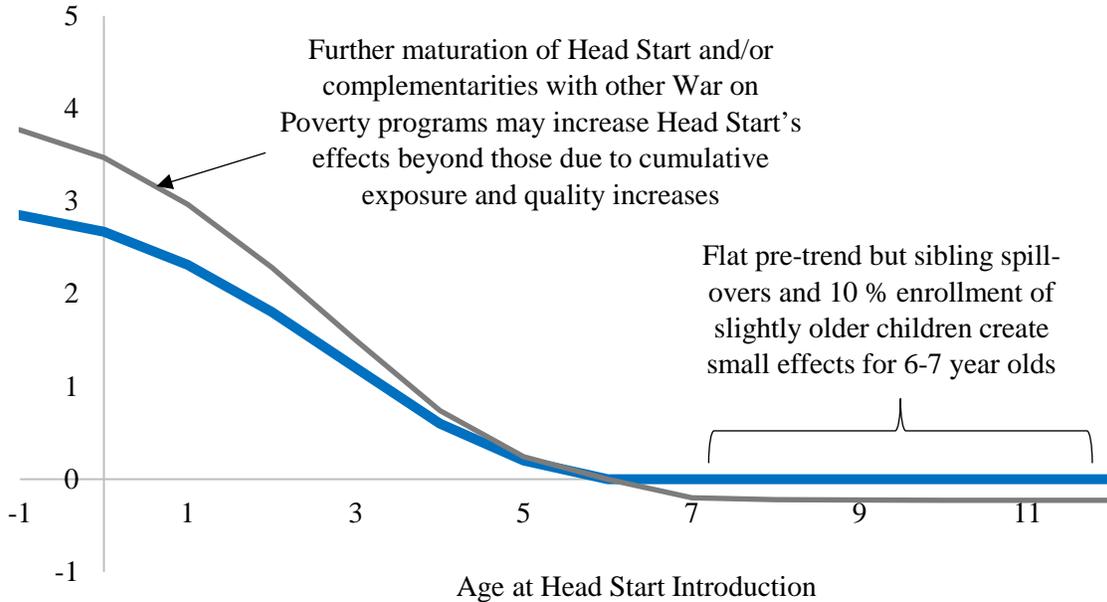
Notes: Counties are grouped by the fiscal year that Head Start launched between 1965 and 1980. Data on federal grants are drawn from the National Archives and Records Administration (NARA). See Bailey and Duquette (2014) and Bailey and Goodman-Bacon (2015) for details on data and variable construction.

Figure 2. The Expected Pattern of Effects on Adult Outcomes by Age of Child at Head Start's Launch

A. No Sibling Spill-overs or Complementarities with Other Programs

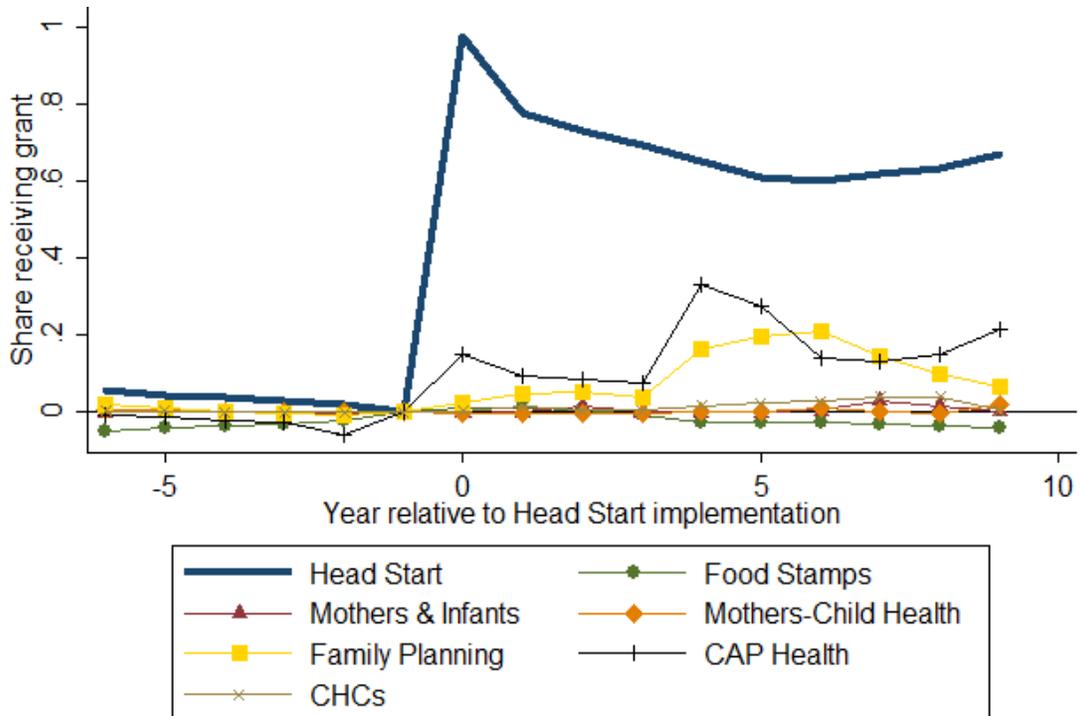


B. With Spill-Over to Siblings over 6 and Complementarities with Other Programs



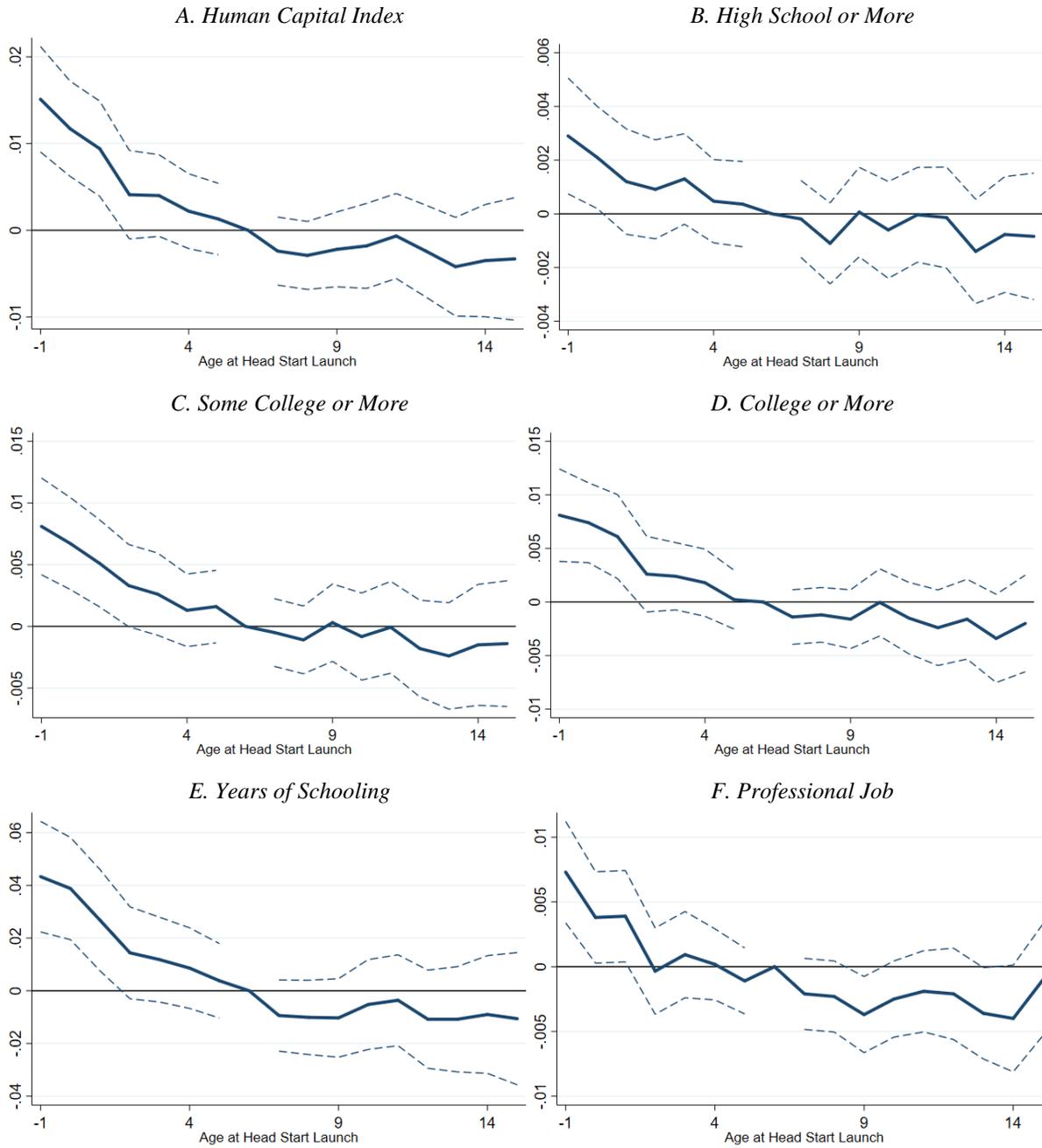
Note: This figure illustrates the potential effects of Head Start assuming there are no effects on children 6 and over, no spillovers to older siblings, and no complementarities with other programs.

Figure 3. Funding for Other OEO Programs Relative to the Year Head Start Began



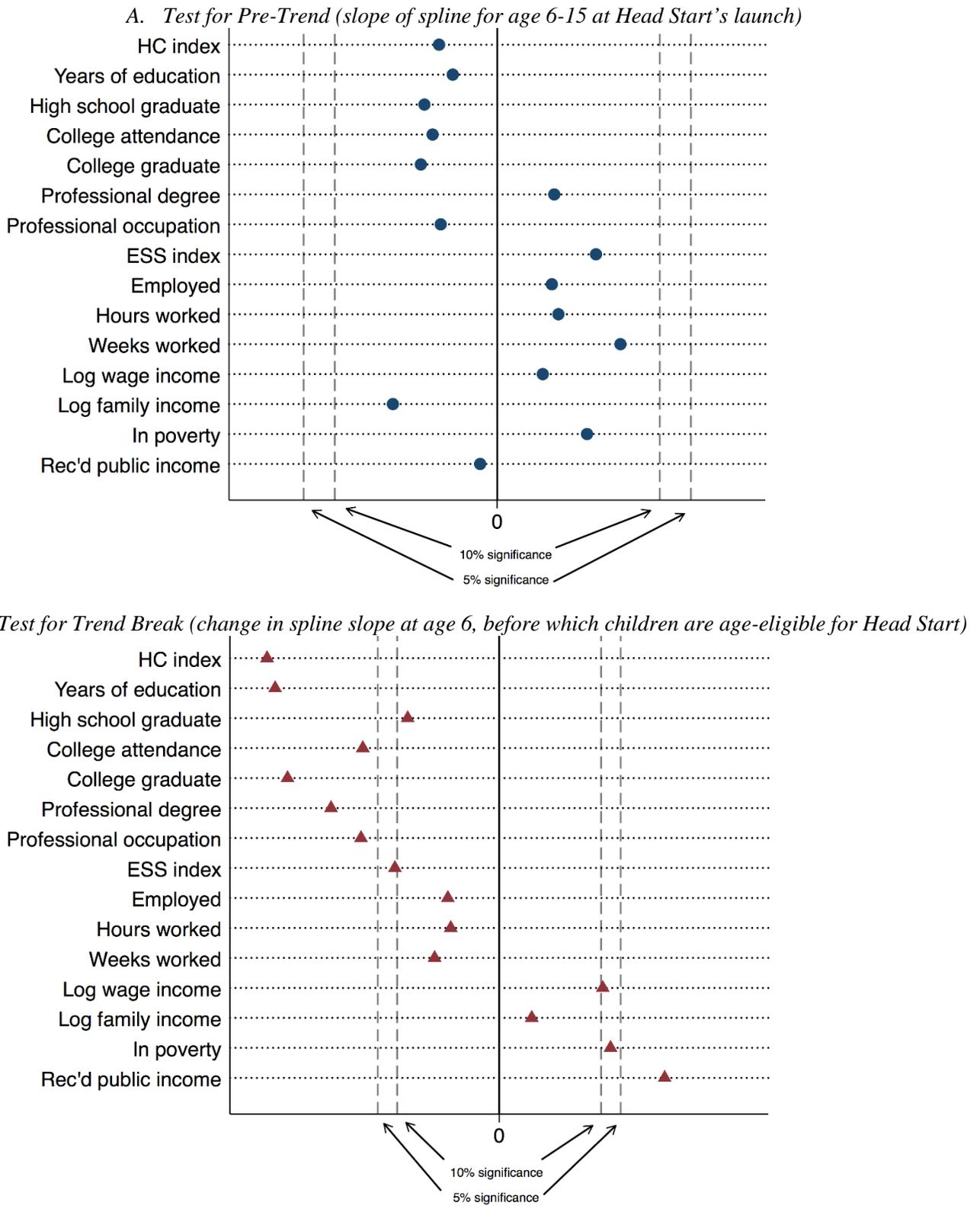
Notes: Dependent variable are binary variables for whether a county received a grant for the indicated program in the indicated year. Data on federal grants and programs are drawn from the NARA.

Figure 4. The Effect of Head Start on Adult Human Capital



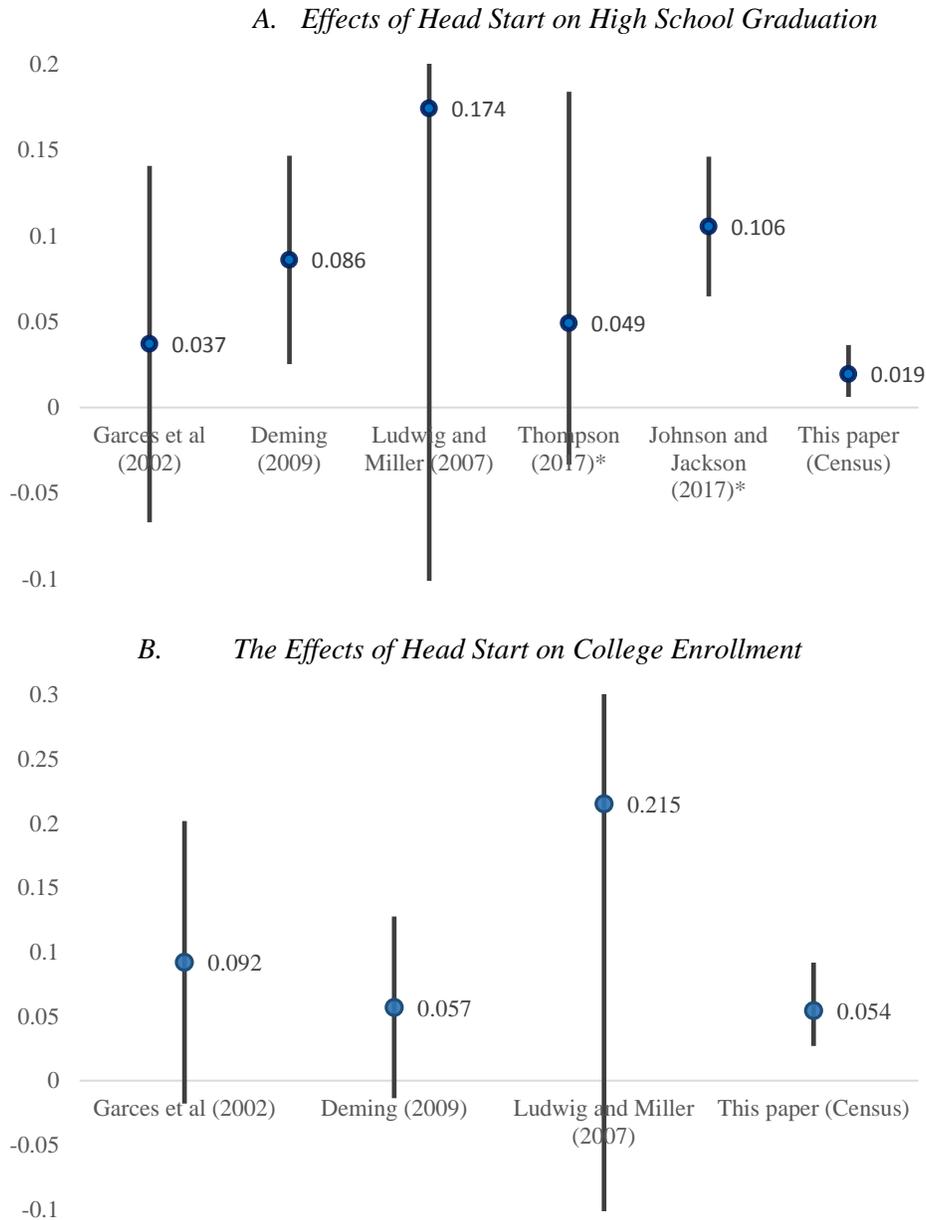
Notes: The figures plot event-study estimates of ϕ for different outcomes using the specification in equation (1). Standard errors clustered at the county level. Dashed lines show 95-percent, point-wise confidence intervals for each estimate.

Figure 5. Visual Representation of Test Statistics Evaluating Pre-Trends and Trend Breaks in Head Start's Effects



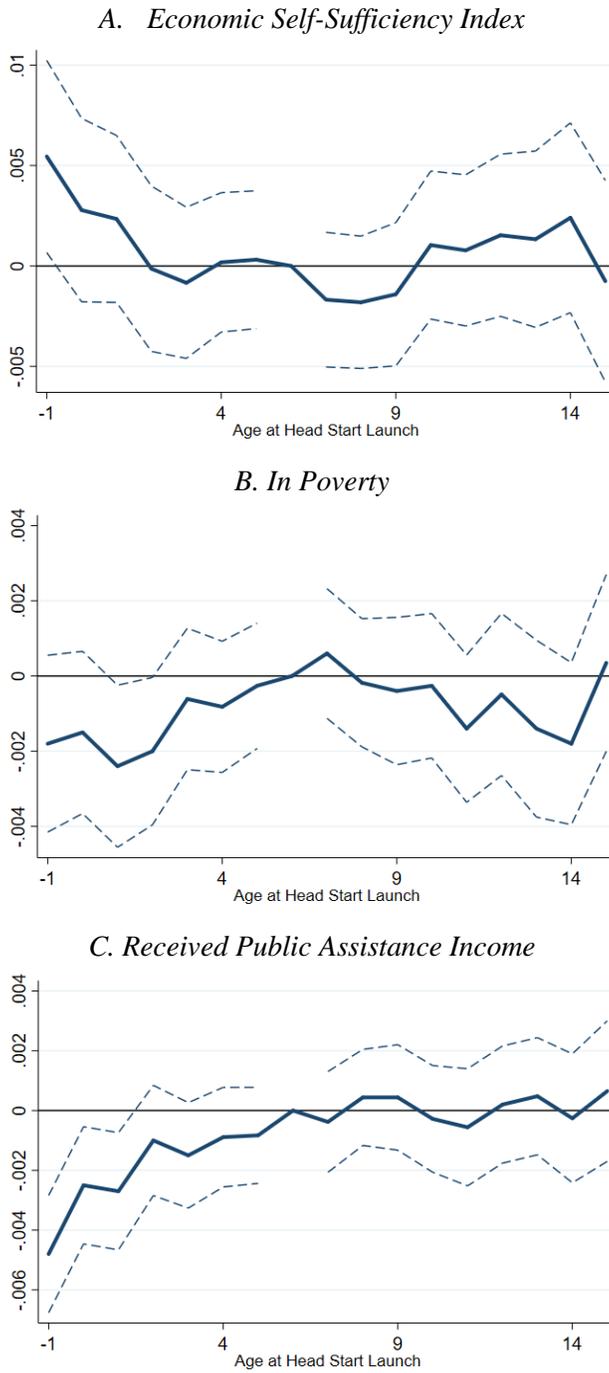
Notes: The figure plots the t-statistic on the slope of the spline for ages 6-15 (panel A) or the F-test for the trend-break at age 6 (panel B). Dashed lines show the threshold for statistical significance at the 10 and 5 percent levels. Compare these to columns 4 and 5 of Tables 1 and 3.

Figure 6. The Magnitude of Head Start’s Effects on Education across Studies



Notes: Circles indicate the reported or derived ATET from different studies. For sibling fixed effect studies, the ATET is directly reported in the papers. Because we cannot resample from data used in other Head Start papers, we calculate the ATETs for other papers using a parametric bootstrap procedure using 10,000 independent draws from normal distributions with means and standard deviations equal to the point estimates and standard errors from the reduced-form and first-stage estimates (Efron & Tibshirani, 1993). Because Johnson and Jackson (2017) does not report a standard error on the first stage, the confidence interval reported for this study in Panel A does not include this first-stage uncertainty. We limited the y-axis range so that the confidence intervals for most studies could be read from the figure. The confidence intervals for Ludwig and Miller (2007) fall outside the y-axis range and are [-0.54,1.47] in panel A and [-0.67,1.82] in panel B. Bars indicate the reported 95-percent confidence interval for sibling fixed-effect models or constructed for the ITT studies as described in the text. See Appendix for more details on the exact figures used. *Johnson and Jackson (2017) and Thompson (2017) sample likely eligible samples of the PSID and NLSY79: individuals born to parents in the bottom quartile of the income distribution, and parents with no college education, respectively.

Figure 7. The Effect of Head Start on Adult Economic Self-Sufficiency



Notes: See Figure 4 notes. Note that In Poverty and Received Public Assistance are reverse coded when included in the Economic Self-Sufficiency Index.

Table 1. The Effect of Head Start on Adult Human Capital

Dependent Variables	(1) Control Mean (std.)	(2) Event Study at -1 (s.e.) [BH p-val]	(3) Spline at -1 (s.e.) [BH p-val]	(4) Test for pre-trend (s.e.) [BH p-val]	(5) F-test of trend break at age 6 (p-val) [BH p-val]	(6) ATET [95% CI]	(7) ATET % increase
Human capital index	0.014 (0.198)	0.015 (0.003)	0.015 (0.003)	-0.0002 (0.0003)	14.02 (0.00)	0.101 [.057,.163]	
Completed high school/GED	0.92 (0.078)	0.003 (0.001)	0.003 (0.001)	-0.0001 (0.0001)	2.18 (0.14)	0.0189 [.005,.038]	2.1%
Attended some college	0.62 (0.140)	0.008 (0.002)	0.008 (0.002)	-0.0002 (0.0003)	4.84 (0.03)	0.054 [.027,.092]	8.7%
Completed 4 year college	0.29 (0.122)	0.008 (0.002)	0.009 (0.002)	-0.0002 (0.0002)	11.66 (0.00)	0.054 [.025,.094]	18.6%
Prof. or doc. degree	0.028 (0.037)	0.002 (0.001)	0.002 (0.000)	0.0000 (0.0001)	7.36 (0.01)	0.014 [.006,.024]	50.0%
Years of schooling	13.57 (0.695)	0.043 (0.011)	0.049 (0.010)	-0.0005 (0.0012)	13.06 (0.00)	0.291 [.144,.49]	2.1%
Has a professional job	0.35 (0.121)	0.007 (0.002)	0.007 (0.002)	-0.0001 (0.0002)	4.97 (0.03)	0.0489 [.022,.085]	14.0%

Notes: In column 1, the control mean and standard deviation are calculated using the cohorts ages 6 and 7 at the time Head Start was launched. Column 2 presents the estimated intention-to-treat (ITT) effect evaluated at birth cohort of full exposure (-1, see Figure 2). Column 3 presents the ITT spline estimate evaluated at -1. Column 4 presents the pre-trend estimate for the spline segment for age 6 and older at implementation. Column 5 presents the F-statistic and p-value for the test of a trend-break in the spline at age 6. The ATET estimate in column 6 divides the ITT effect at -1 by the estimate of receiving a Head Start grant on school enrollment at school age 5, 0.149 (s.e. 0.022) for the full sample and 0.151 (s.e. 0.022) for men and 0.145 (s.e. 0.022) for women; see Appendix Table A4). Column 7 computes the percentage increase implied by the ATET relative to the control mean (the ratio of Column 6 to Column 1) for components of the index. The BH p-values presented in columns 2, 4, and 5 in brackets use the Bonferroni-Holm method to account for multiple hypotheses testing of individual outcomes within an index.

Table 2. The Effect of Head Start on Adult Human Capital by Sex

Dependent Variables	(1) Control Mean (std.)	(2) Event Study at -1 (s.e.) [BH p-val]	(3) Spline at -1 (s.e.) [BH p-val]	(4) Test for pre-trend (s.e.) [BH p-val]	(5) F-test of trend break at age 6 (p-value) [BH p-val]	(6) ATET [95% CI]	(7) ATET % increase
<i>A. Men</i>							
Human capital index	0.014 (0.247)	0.021 (0.004)	0.02 (0.003)	-0.0005 (0.0004)	11.79 (0.00)	0.1360 [.081, .215]	
Completed high school/GED	0.91 (0.104)	0.004 (0.002)	0.003 (0.001)	-0.0001 (0.0002)	1.92 (0.17)	0.0250 [.005, .049]	2.7%
Attended some college	0.59 (0.176)	0.011 (0.003)	0.01 (0.002)	0.0000 (0.0003)	5.67 (0.02)	0.0740 [.036, .125]	12.5%
Completed 4 year college	0.29 (0.150)	0.012 (0.003)	0.011 (0.002)	-0.0003 (0.0003)	9.99 (0.00)	0.0769 [.041, .126]	26.5%
Prof. or doc. degree	0.034 (0.052)	0.003 (0.001)	0.003 (0.001)	0.0000 (0.0001)	4.73 (0.03)	0.0199 [.008, .035]	58.5%
Years of schooling	13.5 (0.878)	0.062 (0.013)	0.063 (0.011)	-0.0012 (0.0015)	12.15 (0.00)	0.4129 [.228, .668]	3.1%
Has a professional job	0.34 (0.153)	0.01 (0.003)	0.009 (0.002)	-0.0006 (0.0003)	2.54 (0.11)	0.0649 [.031, .11]	19.1%
<i>B. Women</i>							
Human capital index	0.015 (0.228)	0.01 (0.004)	0.012 (0.003)	0.0000 (0.0004)	6.50 (0.01)	0.0659 [.014, .132]	
Completed high school/GED	0.93 (0.087)	0.002 (0.001)	0.002 (0.001)	-0.0001 (0.0001)	1.42 (0.23)	0.0140 [-.003, .035]	1.5%
Attended some college	0.65 (0.163)	0.005 (0.002)	0.006 (0.002)	-0.0004 (0.0003)	1.20 (0.27)	0.0370 [.005, .077]	5.7%
Completed 4 year college	0.29 (0.147)	0.005 (0.003)	0.008 (0.002)	-0.0002 (0.0003)	4.63 (0.03)	0.0320 [-.004, .076]	11.0%
Prof. or doc. degree	0.022 (0.042)	0.001 (0.001)	0.001 (0.001)	0.0001 (0.0001)	3.78 (0.05)	0.0080 [-.001, .018]	36.4%
Years of schooling	13.6 (0.795)	0.024 (0.013)	0.038 (0.012)	-0.0003 (0.0014)	4.91 (0.03)	0.1659 [-.013, .379]	1.2%
Has a professional job	0.37 (0.150)	0.005 (0.003)	0.005 (0.002)	0.0004 (0.0003)	4.35 (0.04)	0.0350 [.0, .078]	9.5%

Notes: See Table 1 notes.

Table 3. The Effect of Head Start on Adult Economic Self-Sufficiency

Dependent Variables	(1) Control Mean (std.)	(2) Event Study at -1 (s.e.) [BH p-val]	(3) Spline at -1 (s.e.) [BH p-val]	(4) Test for pre-trend (s.e.) [BH p-val]	(5) F-test of trend break at age 6 [p-val] [BH p-val]	(6) ATET [95% CI]	(7) ATET % increase
Self-sufficiency index	0.024 (0.154)	0.0055 (0.0024)	0.0048 (0.0022)	0.0003 (0.0002)	4.24 (0.04)	0.037 [.005,.076]	
Worked last year	0.86 (0.098)	0.00082 (0.0014)	0.00091 (0.0013)	0.0000 (0.0002)	0.22 (0.64)	0.006 [-.013,.026]	0.7%
Weeks worked last year	41.1 (5.281)	0.047 (0.075)	0.055 (0.0721)	0.0057 (0.0092)	0.61 (0.43)	0.32 [-.68,1.4]	0.8%
Usual hours works per week	35.72 (4.907)	0.0334 (0.080)	0.056 (0.0791)	0.0106 (0.0085)	1.09 (0.30)	0.22 [-.84,1.4]	0.6%
Log labor income	10.6 (0.268)	0.0055 (0.0035)	0.0061 (0.0031)	0.00018 (0.0004)	2.8 (0.09)	0.037 [-.009,.09]	
Log family income/poverty	5.86 (0.258)	0.0059 (0.0043)	0.0054 (0.0041)	-0.0004 (0.0004)	0.28 (0.60)	0.04 [-.017,.104]	
In poverty*	0.1 (0.084)	-0.0018 (0.0012)	-0.002 (0.0011)	0.0001 (0.0001)	3.23 (0.07)	-0.012 [-.03,0.004]	-12.0%
Rec'd public assistance*	0.11 (0.093)	-0.0048 (0.0010)	-0.0037 (0.0008)	0.0000 (0.0001)	7.13 (0.01)	-0.032 [-.052,-.018]	-29.1%

Notes: *In poverty and received public program income are reverse-coded when used in the self-sufficiency index. See also Table 1 notes.

Table 4. The Effect of Head Start on Adult Economic Self-Sufficiency by Sex

Dependent Variables	(1) Control Mean (std.)	(2) Event Study at -1 (s.e.) [BH p-val]	(3) Spline at -1 (s.e.) [BH p-val]	(4) Test for pre-trend (s.e.) [BH p-val]	(5) F-test of trend break at age 6 [p-val] [BH p-val]	(6) ATET [95% CI]	(7) ATET % increase
<i>A. Men</i>							
Self-sufficiency index	0.031 (0.201)	0.0046 (0.0031)	0.00553 (0.0025)	-0.0001 (0.0003)	4.24 (0.21)	0.0299 [-.01,.077]	
Worked last year	0.91 (0.108)	0.0029 (0.0014)	0.00301 (0.0011)	0.0000 (0.0001)	2.33 (0.13)	0.0189 [.001,.041]	2.1%
Weeks worked last year	44.79 (6.031)	0.153 (0.0780)	0.168 (0.0637)	-0.0015 (0.0078)	2.46 (0.12)	1.013 [.002,2.221]	2.3%
Usual hours works per week	41.31 (6.040)	0.148 (0.0882)	0.1722 (0.0798)	0.0039 (0.0080)	2.94 (0.09)	0.98 [-.161,2.317]	2.4%
Log labor income	10.88 (0.301)	0.0033 (0.0043)	0.00238 (0.0036)	-0.0008 (0.0005)	0.33 (0.57)	0.0219 [-.035,.083]	
Log family income/poverty	5.91 (0.302)	0.0035 (0.0048)	0.00245 (0.0041)	-0.0006 (0.0005)	0.07 (0.79)	0.023 [-.04,.092]	
In poverty*	0.07 (0.095)	-0.00033 (0.0013)	-0.000637 (0.0011)	-0.0001 (0.0001)	0.01 (0.90)	-0.002 [-.02,.016]	-2.9%
Rec'd public assistance*	0.11 (0.110)	-0.0046 (0.0015)	-0.0046 (0.0011)	0.0002 (0.0001)	9.93 (0.00)	-0.029 [-.056,-.01]	-26.4%
<i>B. Women</i>							
Self-sufficiency index	0.019 (0.18)	0.0058 (0.0029)	0.00406 (0.0027)	0.0005 (0.0004)	4.24 (0.10)	0.0399 [.001,.088]	
Worked last year	0.8 (0.13)	-0.0018 (0.0020)	-0.00168 (0.0018)	0.0001 (0.0003)	0.15 (0.70)	-0.012 [-.043,.016]	-1.5%
Weeks worked last year	37.66 (6.80)	-0.097 (0.1100)	-0.0931 (0.1015)	0.0091 (0.0146)	0.03 (0.87)	-0.667 [-2.31,.858]	-1.8%
Usual hours works per week	30.49 (5.81)	-0.11 (0.0970)	-0.092 (0.0950)	0.0117 (0.0129)	0.00 (0.95)	-0.745 [-2.24,.599]	-2.4%
Log labor income	10.31 (0.34)	0.0062 (0.0051)	0.007 (0.0048)	0.00084 (0.0005)	3.94 (0.04)	0.043 [-.026,.122]	
Log family income/poverty	5.81 (0.32)	0.0083 (0.0054)	0.0084 (0.0050)	-0.00034 (0.0005)	0.69 (0.41)	0.057 [-.015,.143]	
In poverty*	0.12 (0.11)	-0.0028 (0.0017)	-0.00273 (0.0015)	-0.0003 (0.0002)	4.21 (0.04)	-0.018 [-.047,.004]	-15.0%
Rec'd public assistance*	0.12 (0.12)	-0.0049 (0.0014)	-0.0029 (0.0011)	0.0002 (0.0001)	0.67 (0.41)	-0.034 [-.06,-.014]	-28.3%

Notes: *In poverty and received public program income are reverse-coded when used in the self-sufficiency index. See also Table 1 notes.

Table 5. Heterogeneity in the Effect of Head Start, by Local Programs and Economic Circumstances

	Intent-to-Treat (ITT) Effect			Effect on preschool enrollment (s.e.)		Average Treatment Effect on Treated Children (ATET)	
	Above median	Below median	F-test of diff (p-val)	Above median	Below median	Above median	Below median
	<i>A. Human capital index</i>						
Medicaid exposure	0.021 (0.004)	0.009 (0.004)	4.095 (0.043)	0.123 (0.024)	0.175 (0.030)	0.171	0.052
CHC exposure	0.020 (0.003)	0.010 (0.003)	9.258 (0.002)	0.203 (0.034)	0.109 (0.025)	0.100	0.090
Food Stamps exposure	0.011 (0.003)	0.020 (0.004)	4.515 (0.034)	0.122 (0.026)	0.167 (0.032)	0.092	0.117
Poorest 300th counties	0.009 (0.011)	0.015 (0.003)	0.320 (0.572)	0.073 (0.026)	0.125 (0.023)	0.125	0.123
Predicted economic growth	0.023 (0.004)	0.010 (0.003)	11.948 (0.001)	0.139 (0.033)	0.125 (0.025)	0.166	0.078
<i>B. Self-sufficiency index</i>							
Medicaid exposure	0.003 (0.004)	0.005 (0.002)	0.336 (0.562)	0.123 (0.024)	0.175 (0.030)	0.023	0.031
CHC exposure	0.006 (0.003)	0.003 (0.003)	1.708 (0.191)	0.203 (0.034)	0.109 (0.025)	0.030	0.028
Food Stamps exposure	0.002 (0.003)	0.007 (0.002)	3.547 (0.060)	0.122 (0.026)	0.167 (0.032)	0.017	0.039
Poorest 300th counties	0.004 (0.010)	0.004 (0.002)	0.000 (0.991)	0.073 (0.026)	0.125 (0.023)	0.060	0.035
Predicted economic growth	0.006 (0.003)	0.003 (0.003)	1.014 (0.314)	0.139 (0.033)	0.125 (0.025)	0.041	0.027

Notes: ATETs are constructed by dividing the group-specific ITT estimate of Head Start's effect on long-run outcomes by the group-specific estimated first stage. Results for the 300 poorest counties are reported in the column for “above median” with results for other counties reported in “below median.”

[\[Click here for Online Appendices\]](#)