Online Appendix

The Long-Run Impacts of Head Start on Human Capital and Economic Self-Sufficiency

April 26, 2021

1. Census Data

The project's primary data source is the 2000 Census and 2001 to 2018 ACS combined with the SSA's Numident file, accessed through project 1284 in the University of Michigan Research Data Center (RDC). The advantage of these data is that they link a rich set of productivity outcomes for cohorts potentially benefitting from War on Poverty programs (those who are ages 25 to 54 in the Census/ACS) with information on their access to Head Start programs in childhood using Numident information on county of birth. The 2000 Census long-form contains information on 16.7 percent of the U.S. population; the 2001 to 2018 ACSs contain information for around 14 percent of the U.S. population. The number of Numident-linked, unique individuals in these combined data sources represent about one-quarter of the U.S. population. In addition, we use the 1970 restricted long-form Census that contains information on school enrollment for children. Unfortunately, these data cannot be linked to the Numident because they have not yet been PIK'd by the Census.

In addition to the sample restrictions above, we limit our sample to individuals with non-missing information for all of the components of our human capital and economic self-sufficiency indices. The only exceptions are individuals with 0 wage or family income, where we keep these observations and compute relevant sample sizes required as part of disclosure. We implement this restriction because it allows us to maintain a consistent sample for our estimated effects on the indices and each of their components, while minimizing the burden on Census Bureau review processes and the risk of disclosing sensitive information. What might not be clear to many who haven't used the Research Data Center (RDC) is that disclosure requires us to calculate the sample size underlying each estimate as well as "implicit sample" sizes for each estimate. Implicit samples are defined as the difference between any two samples used in the analysis. So, if we estimate a regression using employment as our outcome variable, and then estimate a separate regression using log wages as the outcome (where log wages are defined only among people who are employed), disclosure from the RDC requires not only that the samples for each of these regressions meet a certain size threshold, but also that the difference between the two samples meet a size threshold. Different rates of missing values across outcomes, therefore, introduce many different implicit samples and also may limit disclosure of results. In addition, different specifications may imply different samples and implicit samples due to the availability of covariates. The potential combinations of outcomes, models, and subgroups leads these sample and implicit sample combinations to grow very quickly and ultimately preclude disclosure of the results if the Census deems the multiple implicit samples could create risks of indirect disclosure of respondent information. For example, the analysis of employment and log wages described above would create 2 samples and 1 implicit sample. If we added an analysis of the same variables in a sample restricted to college graduates, we would gain 2 samples and 4 implicit samples. An additional sample restriction would add 2 more samples and 7 more implicit samples.

Restricting our sample to individuals with non-missing information for all of the components of our human capital and economic self-sufficiency indices reduces the number of observations from 30.8 million individuals to 22.48 million—a loss of roughly 25% of the observations. In Appendix section 8, we show the results are robust to this sample restriction.

2. SSA's Numident: Data on County and Date of Birth

Links from the Census/ACS to place and date of birth are important for studying the long-term impacts of Head Start, as place and date of birth provide crucial information on exposure of individuals to these programs in early childhood. For the Census/ACS files we use the survey-internal PIK code to match individuals with the SSA's Numident file. The Numident place-of-birth variable is a string variable detailing, in most cases, the city and state of birth. In previous work, Isen, Rossin-Slater, and Walker (2013) developed a matching algorithm to connect this string variable to the Census Bureau database of places, counties, and minor civil divisions as well as the United States Geological Survey's Geographic Names Information System (GNIS) file. We also make use of code that was developed for a similar purpose by Black, Sanders, Taylor, and Taylor (2015). Using both sources, we constructed a crosswalk between the NUMIDENT place of birth string variable and (standard) county FIPS codes, with over 90 percent of individuals matched to their counties of birth. Taylor, Stuart, and Bailey (2016)'s technical memorandum has been posted with the Census Bureau and contains this information.

The Census/ACS data have the benefit of providing a wide range of outcomes of interest, including individual earnings, but also program participation, disability, living arrangements, and family and household variables such as income and poverty. Additionally, we observe individuals in all states and we observe individuals regardless of whether they are employed. The Census/ACS data also have limitations. They are repeated cross-sections and information is self-reported (and, so, measured with error). They provide no information on childhood living circumstances or parental characteristics. This prevents us from constructing a "high-impact" sample of individuals who were more likely to benefit from Head Start (as done in studies using longitudinal surveys), based on measures such as parental income or education.

3. Data on School Age Entry Cutoffs

Information on school-age entry cutoffs is taken from Bedard and Dhuey (2012) and supplemented using our own research for two states. According to state legislative documents, the school-entry age cutoff for the entire state of Texas was September 1st starting in 1969.[1] The state of Kansas' cutoff date was January 1st before 1965.[2] State cohorts missing information from Bedard and Dhuey (2012) and our research include all individuals born in Colorado, Georgia, Indiana, Massachusetts, New Jersey, and Utah. In addition, we are missing information for Montana for cohorts born before 1979, Rhode Island for cohorts born before 1967, South Carolina for cohorts born before 1978, Texas for cohorts born before 1969, Washington for cohorts born before 1977, and West Virginia for cohorts born before 1972. For states where we are missing information on school age entry cutoffs before a certain year, we assume the dates for the school age entry cutoffs for younger cohorts (1979 and after) apply to older cohorts. For instance, because Bedard and Dhuey (2012) and our research did not turn up state school-entry age cutoffs for Montana before 1979, we assume that the cutoff in 1979 applies to all cohorts prior to 1979. For states where we could not identify school-entry cutoffs at any date for the cohorts in our sample, we impute the cutoff using the year-specific average cutoff date among all states with non-missing cutoffs. This approximation adds some

[1] http://www.heinonline.org/HOL/Page?men_tab=srchresults&handle=hein.ssl/sstx0163&size=2&collection=ssl&id=718
[2] http://www.heinonline.org/HOL/Page?men_tab=srchresults&handle=hein.ssl/ssks0074&size=2&collection=ssl&id=477; http://kslegislature.org/li_2014/b2013_14/statute/072_000_0000_chapter/072_011_0000_article/072_011_0007_section/072_011_0007_k/

measurement error to our measure of school age, but also allows us to greatly expand our sample to all states but Alaska and Hawaii. Our results are robust to dropping states and cohorts where we do not observe the exact cutoff date.

4. Data on Head Start

To study the long-run impacts of access to Head Start, we also use additional county-level sources of data to account for potentially confounding local programs and the economy. These data include information from the following sources: Bailey and Goodman-Bacon (2015) collected data on the OEO's community programs from the National Archives Community Action Program (NACAP) files, as well as from some administrative sources. For Community Health Centers, some information was hand-entered from annual Public Health Service (PHS) Reports. The resulting database contains information on (1) the county where a program delivered services, which allows each federal grant to be linked to county-level mortality rates; (2) the date that each county received its *first program services* grant (this excludes planning grants), which provides the year that programs began operating; and (3) some information on program grants between 1978 and 1980 from the National Archives Federal Outlays (NAFO) files. We supplement these data with information on the legal services program from Cunningham (2013) and Food Stamps from Hoynes and Schanzenbach (2009).

5. Data on Head Start Launch Dates

Our main policy variable is the availability of the Head Start program, which was rolled out across counties during the War on Poverty. Bailey and Duquette (2014) and Bailey and Goodman-Bacon (2015) have compiled information from the National Archives and Records Administration on changes in Head Start funding between 1965 and 1980, which we use in this study. To verify their accuracy, these data have been compared to federal government directories of Head Start programs (Office of Child Development, 1973; Project Head Start, 1971, 1978). The following tables and figures supplement the analysis and description in the paper with more information on the roll-out of the program.

Table A1. Share of Counties and Children under 6 in County with Head Start, 1965-1980

Fiscal Year HS Launched	# of Counties	Percent of counties	Cumulative percent of counties	Percent of kids age under 6 in 1970	Cumulative percent of kids age under 6 in 1970
1966	536	17.53%	17.57%	54.71%	55.22%
1967	217	7.10%	24.66%	10.57%	65.79%
1968	607	19.86%	44.52%	16.06%	81.85%
1969	41	1.34%	45.86%	0.71%	82.56%
1970	45	1.47%	47.33%	0.87%	83.43%
1971	10	0.33%	47.66%	0.17%	83.61%
1972	30	0.98%	48.64%	0.51%	84.12%
1973	7	0.23%	48.87%	0.11%	84.23%
1974	9	0.29%	49.17%	0.15%	84.38%
1975	16	0.52%	49.69%	0.29%	84.67%
1976	7	0.23%	49.92%	0.08%	84.75%
1977	4	0.13%	50.05%	0.16%	84.90%
1978	9	0.29%	50.34%	0.31%	85.21%
1979	3	0.10%	50.44%	0.08%	85.29%
No HS<1980	1515	49.56%	100.00%	14.71%	100.00%
Total	3057				

Table A2. 1960 County Characteristics and Head Start's Launch, 1965-1980

	(1)	(2)	(3)	(4)	(5)
	Head Start launches before 1980	No Head Start before 1980	First Head Start grant in		
			1965-1966	1967-1968	1969-1980
# of Counties	1542	1515	537	824	181
County Population	87,460	15,326	161,162	52,056	29,971
% Urban	72.21	31.02	79.44	61.4	42.35
% Rural Farm	5.72	24.44	3.42	9.01	16.38
% Nonwhite	10.77	11.2	11.33	9.45	12.31
% Population Aged 0-4	11.53	11.15	11.51	11.56	11.58
% Population Aged 65+	8.95	10.89	8.74	9.29	9.66
Median Family Income (1959)	5712.77	4145.36	5984.64	5311.17	4550.95
Active MD per 1,000 Pop	0.69	0.02	0.96	0.23	0.04
AMR, All Ages	955.45	925.51	963.51	941.67	935.72
Infant Mortality Rate	25.6	27.28	25.36	25.77	28.01
Any Med Students (1969)	0.25	0.01	0.35	0.1	0
% Family Income <$3000	20.41	35.85	18.02	23.85	31.39
% Family Income $10,000+	15.15	7.44	16.73	12.74	9.04
% w/ 12+ Yrs Schooling (Age 25+)	41.81	34.18	43.04	40.01	36.37
% w/ <4 Yrs Schooling (Age 25+)	8.22	10.8	7.79	8.74	10.94
Gov'ts Exp ($000s) per 1,000 Pop (1957)	143.29	128.68	152.01	128.96	117.34
Labor Force Participation	0.38	0.36	0.38	0.37	0.36
% Labor Force Unemployed	5.32	4.8	5.38	5.21	5.24

Notes: All values are population-weighted means, with the exception of average county population in row 2. Characteristics are for 1960 unless otherwise specified. All variables are taken from the 1960 County and City Databooks (Haines, Inter-university Consortium for, & Social, 2010) and 1990 Area Resource Files (US DHHS 1994) except the following. Medicare variables are for 1966, taken from the County-level Medicare File (US SSA 1969-1977; US HFA 1978-1980). Data on Medicare expenditures were shared by Almond, Hoynes, and Schanzenbach (2011). We also use the 1959 to 1988 Vital Statistics Multiple-Cause of Death Files (US DHHS and NCHS 2007) to compute mortality rates.

Table A3. Regression Analysis of County Characteristics and Head Start's Launch, 1965-1980

	(1)	(2)	(3)
% Nonwhite	0.008	0.006	0.005
	[0.006]	[0.007]	[0.007]
% Population Aged 0-4	0.014	0.053	0.059
	[0.060]	[0.063]	[0.061]
% Population Aged 65+	0.002	0.021	-0.001
	[0.033]	[0.035]	[0.035]
Active MD per 1000 Pop	-0.755***	-0.540***	0.395***
	[0.192]	[0.182]	[0.145]
AMR, All Ages	0.000	-0.000	-0.000
	[0.001]	[0.001]	[0.001]
AMR, 1960-65 Change	-0.000	0.000	0.000
	[0.001]	[0.001]	[0.001]
Infant Mortality Rate	0.000	-0.000	-0.003
	[0.005]	[0.006]	[0.006]
Any Med Students (1969)	-0.238	-0.439**	-0.401**
	[0.187]	[0.211]	[0.190]
% Family Income <$3,000	0.028	0.017	0.023
	[0.027]	[0.028]	[0.027]
% Family Income $10,000+	-0.090**	-0.090*	-0.064
	[0.042]	[0.048]	[0.046]
% w/ 12+ Yrs Schooling (Age 25+)	0.013	0.020	0.012
	[0.014]	[0.017]	[0.017]
% w/ <4 Yrs Schooling (Age 25+)	0.001	0.007	0.011
	[0.017]	[0.020]	[0.020]
Gov'ts Exp ($000s) per 1000 Pop (1957)	0.002	0.003	0.001
	[0.002]	[0.002]	[0.002]
Labor Force Participation	5.325***	5.417***	2.863
	[1.789]	[1.813]	[1.797]
% Labor Force Unemployed	-0.068**	-0.061**	-0.040
	[0.027]	[0.028]	[0.027]
% Labor Force Male	0.054***	0.050***	-0.006
	[0.017]	[0.018]	[0.019]
Median Family Income (1959)	0.000	0.000	0.001
	[0.000]	[0.000]	[0.000]
Log Population			-0.646***
			[0.110]
Additional Covariates		S,U	S,U
R-squared	0.065	0.113	0.157
P-value from test of joint significance of covariates			0.23

Notes: We estimate each regression by ordinary least squares. The dependent variable is year of Head Start's launch. Heteroskedasticity robust standard errors are beneath the point estimates in brackets. The regressions exclude counties that never received Head Start funding and are unweighted. Characteristics are for 1960 unless otherwise specified. P-value is from an F-test of joint significance of all covariates except state fixed effects (S) and urban group categories (U, $0, 0 < u < 25, 25 \leq u < 50, 50 \leq u < 75, 75 \leq u \leq 100$, where u is the share of a county's population living in an urban area), log population, and median family income. See also Table A2 notes.

Table A4. Age at Launch by Cohort and Year Head Start Launched

	Year of Head Start's Launch (Fall Term)													
Birth Cohort	**1965**	**1966**	**1967**	**1968**	**1969**	**1970**	**1971**	**1972**	**1973**	**1974**	**1975**	**1976**	**1977**	**1978**
1950	15													
1951	14	15												
1952	13	14	15											
1953	12	13	14	15										
1954	11	12	13	14	15									
1955	10	11	12	13	14	15								
1956	9	10	11	12	13	14	15							
1957	8	9	10	11	12	13	14	15						
1958	7	8	9	10	11	12	13	14	15					
1959	6	7	8	9	10	11	12	13	14	15				
1960	5	6	7	8	9	10	11	12	13	14	15			
1961	4	5	6	7	8	9	10	11	12	13	14	15		
1962	3	4	5	6	7	8	9	10	11	12	13	14	15	
1963	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1964	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1965	0	1	2	3	4	5	6	7	8	9	10	11	12	13
1966	-1	0	1	2	3	4	5	6	7	8	9	10	11	12
1967	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11
1968	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10
1969	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9
1970	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8
1971	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
1972	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
1973	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5
1974	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4
1975	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3
1976		-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2
1977			-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1
1978				-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
1979					-10	-9	-8	-7	-6	-5	-4	-3	-2	-1
1980						-10	-9	-8	-7	-6	-5	-4	-3	-2

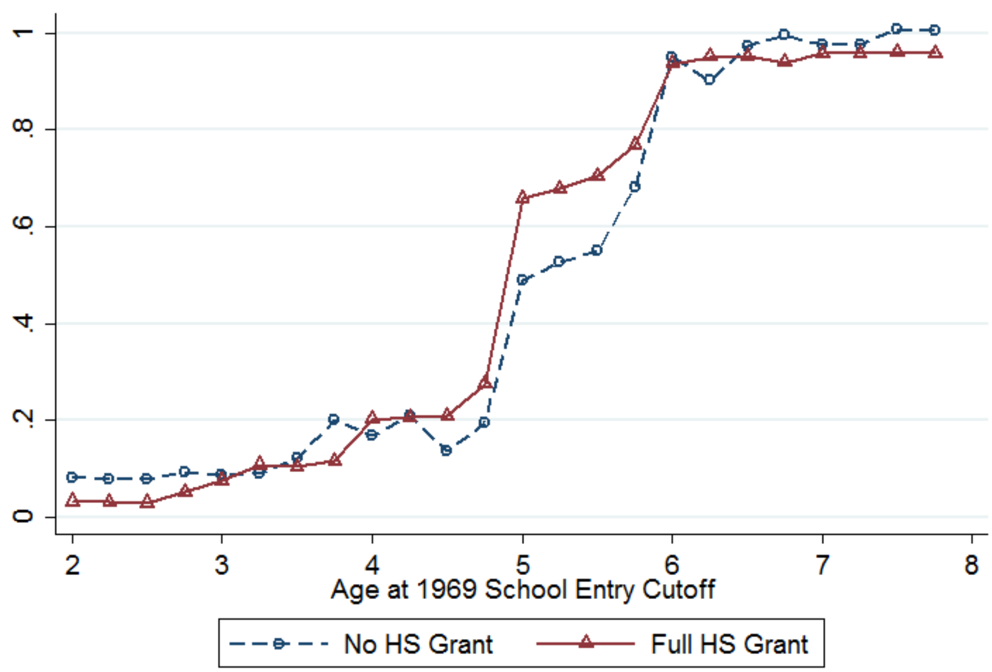
Notes: Table documents the age at Head Start's launch for each birth cohort and Head Start launch date in our data. Specified age applies to children born before the school-entry cutoff for the school-year that began in the fall of the given year of Head Start's launch. Noteworthy is that our sample is compositionally balanced from ages -2 to 14. Age 15 is unbalanced because in counties that adopted Head Start in the 1965-66 school year and had school-age cutoffs in the fall, this school-age cohort included children born in late 1949, which is the birth year before our sample starts.

6. The Effect of a Head Start Launch on Head Start Enrollment

Figure A1 shows the unadjusted enrollment gap in the public IPUMS data between counties that had a Head Start program in 1970 versus those that did not. Notably, 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). Five-year-old children were 17 percentage points less likely to be enrolled in school (48.9 versus 65.9). In the public data, these gaps are 5.9 percentage points among four-year-olds and 21.3 percentage points among five-year-olds when looking only at children of mothers with less than a high school education.

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.

Figure A1. 1970 School Enrollment by Mother's Education



Notes: The figure plots the predicted school enrollment by age for children in counties with Head Start versus those in counties without Head Start in 1970. These predictions come from a linear-probability model using a dependent variable is equal to one if a child was enrolled in school on February 1, 1970. Because the most detailed level of geography available in the public-use data is county group, availability of Head Start is operationalized as the population-weighted share of counties in an individual's county group that had Head Start by the 1969-1970 school year. The sample is limited to states where the school-entry cutoff falls at the beginning of a quarter and includes children between school age 2 and 7 using the school-entry age cutoff date in 1969. Source: Authors' calculations using the public 1970 Census (Ruggles, Genadek, Grover, & Sobek, 2015).

We explore these gaps in more detail using the *restricted* 1970 Census, which allows us to use a 1 in 6 sample of the U.S. population and county of residence (rather than county group). Using exact county rather than county group is more analogous to the long-term outcomes in the Census/ACS analysis. We compare school enrollment by a child's age in 1970 after adjusting for different county characteristics using either covariates or county fixed effects using the following specification:

$$School\ Enrollment_{ic} = \boldsymbol{Z}_c'\boldsymbol{\beta_0} + \boldsymbol{A}_i'\boldsymbol{\beta_1} + \boldsymbol{A}_i'HeadStart_c\boldsymbol{\beta_2} + \varepsilon_{ic},$$

where $\boldsymbol{A}_i$ has elements indicating the child's age in quarters relative to the school entry cut-off (e.g., 4, 4.25, 4.5, 4.75); *HeadStart* is a dummy variable equal to 1 if the county had a Head Start program funded before 1970 (and is zero for places receiving their program in fiscal year 1971 or later). The set of covariates, $\boldsymbol{Z}_c$, includes either (1) $\theta_{s(c)}$, which captures state fixed effects to account for age-invariant, state-level factors that determine the local supply of preschools as well as 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) or (2) county-level fixed effects (π_c). The point estimates of interest are the elements of $\boldsymbol{\beta_2}$, which, after regression-adjusting for county characteristics, capture differences in school enrollment rates of children ages 3 to 5 in counties with Head Start.

Table A5. Regression-Adjusted Relationship between Head Start and Enrollment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Dependent variable: 1=School Enrollment						
	All children			Boys	Girls	Nonwhite	White
Head Start x age 3	0.038	0.038	0.038	0.037	0.039	0.032	0.039
	(0.0058)	(0.0058)	(0.0059)	(0.0065)	(0.0070)	(0.0115)	(0.0062)
Head Start x age 4	0.088	0.089	0.089	0.093	0.084	0.099	0.086
	(0.0088)	(0.0087)	(0.0088)	(0.0094)	(0.0095)	(0.0154)	(0.0094)
Head Start x age 5	0.148	0.149	0.149	0.151	0.145	0.214	0.132
	(0.0217)	(0.0217)	(0.0217)	(0.0220)	(0.0222)	(0.0275)	(0.0228)
Observations	830,000	830,000	830,000	420,000	410,000	156,000	675,000
Control mean	0.52	0.56	0.52	0.52	0.52	0.40	0.55
State FE	Y	Y					
County controls		Y					
County FE			Y	Y	Y	Y	Y

Notes: Sample is limited to children in states where we observe a school age entry cutoff, and where the school entry cutoff coincides with the beginning or end of a calendar quarter. Access to Head Start is measured as equal to 1 if a child lives in a county with a Head Start program in the 1969-70 school year. Observation counts are rounded per disclosure requirements. Omitted category is children age 7.75. Standard errors in brackets are clustered by county. Source: Authors calculations using the restricted 1970 Census.

The regression-adjusted, preschool enrollment gaps are summarized in Table A5. School enrollment was 29 percent higher for all five-year-olds (0.149/0.52), 29 percent higher for boys (0.151/0.52) and 28 percent higher for girls (0.145/0.52). Although we are unable to assess many potential threats to the internal validity of this cross-sectional research design, the high degree of robustness of these estimates to the inclusion of different covariates in columns 2 and 3 is encouraging.

Our best estimate of the effect of a Head Start program from the 1970 Census on a birth cohort's exposure to Head Start is around 14.9 percentage points (Table A5, column 3). It is 0.151 for men (column 4) and 0.145 for women (column 5). By construction, Census estimates should *omit* summer Head Start. An additional advantage of using the Census estimates is that they provide standard errors, which we use in our parametric bootstrap to scale our ITT estimates into ATETs.

Census estimates accord well with administrative data and are comparable to other studies. Administrative data suggest that in 1971, the average Head Start program served about 10 percent of resident 4-year-olds, which compares very well to around the 9 percent increase in school attendance in 1970 in the Census. The similarity of these administrative numbers and Census estimates suggests that crowd-out is minimal. To the extent that interested readers believe the estimate of the first stage should be higher or lower, they can deflate or inflate our ATET estimates accordingly.

We use the estimate of 0.149 from the Census to transform the ITT effects from the spline specifications into average treatment-effects-on-the-treated (ATET). We also construct confidence intervals using a parametric bootstrap procedure with 10,000 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.[3] Rather than impose independence between the magnitude of the Head Start take-up estimates (first stage) and ITT estimates (reduced form), we sample counties with replacement, estimate separate regressions for the first stage and reduced form, and repeat the procedure 1,000 times to compute the correlation between first-stage and reduced-form effects. This produces an estimated correlation of 0.07, which modestly reduces the 95-percent confidence intervals for these estimates.

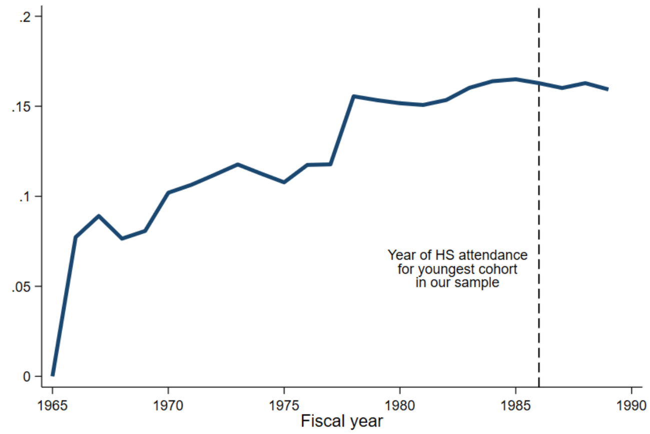
7. The Evolution of Head Start Capacity and Quality

We do not have data on individual program enrollment by year, but Figure A2 (next page) shows that Head Start attendance was increasing through around 1978—roughly 10 years after the average program began. This fact fits well with the ten years implied by the difference between age 5 and -5—the second knot point at which point the intention-to-treat effects level off. Because the number of locations with Head Start barely increased after 1970 (Appendix Table A1), the growth in enrollment implies increasing local capacity—the likelihood that more children in an area benefitted the longer the programs were in existence.

Also, the growth in program capacity is complemented by improvements in quality with program age (described in the text) as well as increased funding from other sources. The first mandatory Head Start performance standards for local programs were adopted in 1975 (GAO report HEHS-98-65, page 5). Guidelines had been developed earlier but were not mandatory. Moreover, the share of programs citing OEO as the sole source of support decreased from 75% in 1968-69 to 59% in 1969-70 (Project Head Start 1969-1970: A descriptive report of programs and participants, page 22).

[3] This procedure follows Efron and Tibshirani (1993) (pp. 53-6) and Johnston and DiNardo (1997) (pp. 365-6).

Figure A2. Changes in Head Start Enrollment per 4-Year Old, 1965-1990

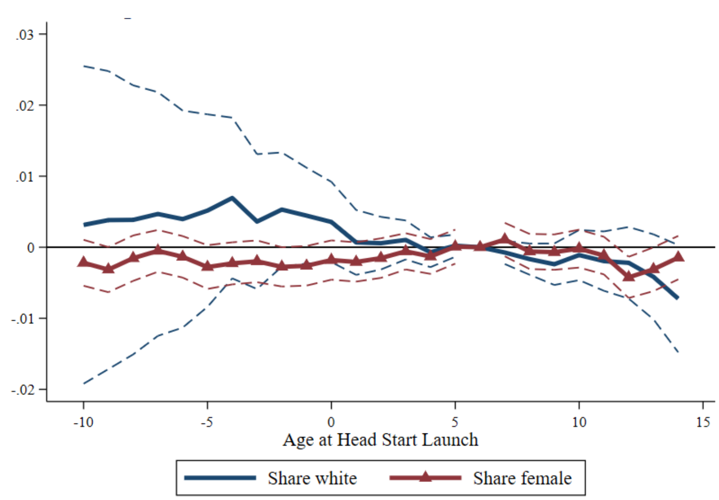


Notes: The enrollment rate is constructed as total enrollment in full-year Head Start divided by the population of 4-year-olds in counties that had implemented Head Start by a given year. Enrollment data comes from the Office of Head Start, (accessed May 5, 2020 here). Population data comes from the CDC SEER file. Before 1969, the first year of SEER population data, the county-level of population of 4-year-olds is imputed using the 1969 population of 5-8-year-olds. National Archives data on Head Start grants by county is used to calculate the population of 4-year-olds in counties that had implemented Head Start.

8. Additional Balance and Validity Tests of the Research Design

This section outlines additional tests of the validity of the research design. Figure A3 shows event-study estimates from our main estimating equation using share female or share white as the dependent variable. These figures find little evidence of imbalance in these demographic characteristics. Although the female share falls slightly and the share white rises slightly for younger cohorts exposed to Head Start, the evolution is smooth and there is no evidence of a trend break at age 6 that would confound our estimates. Taking these point estimates very seriously would suggest that a one-half percentage-point increase in the share white (the estimate at -5 in Figure A3) would explain an increase in college completion of 0.0003 through a change in sample composition alone (as shown in column 1 of Table 3, 31 percent of whites but only 25 percent of nonwhites have a 4-year degree, so 0.005 x 0.31 – 0.005 x 0.25 = 0.0003). This figure is less than 2 percent the size of our actual estimate of 1.7 percentage points (see Table 1). College completion rates are higher for females in our control group than their male counterparts (see Table 2), so a decrease in the share female would make our estimates smaller.

Figure A3. Event Study Estimates of the Effect of Head Start on Race and Sex Composition

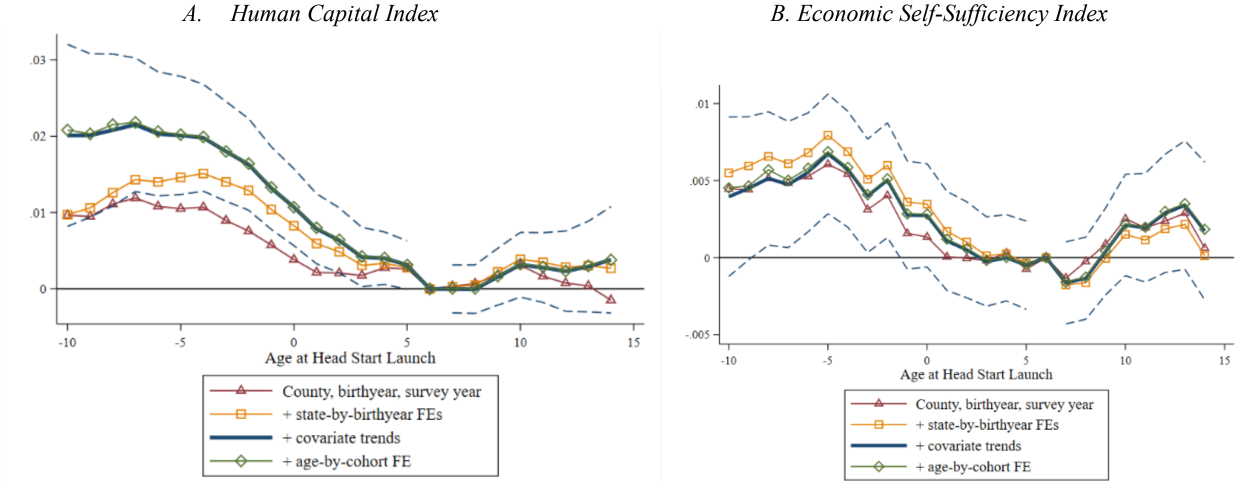


Notes: See Figure 4 notes.

Figure A4 provides evidence on the sensitivity of our treatment effects to the inclusion of flexible controls for age and other covariates. The estimated effects on human capital increase when we add state-by-year-of-birth fixed effects, effectively controlling non-parametrically for state-level policies and other state shocks to outcomes. They increase again when we add 1960 covariates interacted with linear trends, highlighting the importance of accounting for the demographic trends that affected predominantly urban Head Start counties more severely at this time period. In particular, 1960 covariates interacted with linear trends account for the *secular trends toward worse outcomes* in poor, urban areas over the 1960s and 1970s. However, the inclusion of age-by-year-of-birth fixed effects has virtually no effect on our estimates. Our estimated effects on economic self-sufficiency are quite similar across all specifications.

Figure A4. Robustness of Indices to Control for Age and Other Covariates

A. Human Capital Index *B. Economic Self-Sufficiency Index*



Notes: See Figure 4 notes.

We also examine selection into our analysis sample due to the "full information" restriction discussed in section 1. We test for differential selection into our analysis sample (what we call the "full information" sample, which omits individuals with missing values) by regressing an indicator for being in the full information sub-sample (=1) versus in the entire sample (30.8 million individuals with valid PIKs in the Census and ACS) on the covariates in our preferred specification. Figure A5 (next page) shows these estimates alongside our estimates for the HC and ESS indices. The selection plot is both flat and the data reject a trend break at age 6 (F-statistic of 0.59, p-value 0.44), suggesting that selection is not driving our results.

In addition, Figure A6 (next page) shows the robustness of our main estimates to using *all* of the 30.8 million individuals—not just the individuals with non-missing values. Reassuringly, the estimates are nearly identical in our non-missing value sample to the full sample. We think these results provide reassuring evidence that allocated or missing responses do not introduce serious concern about selection bias.

Figure A5. Event-Study Estimates of Selection into the "Full Information" Sample

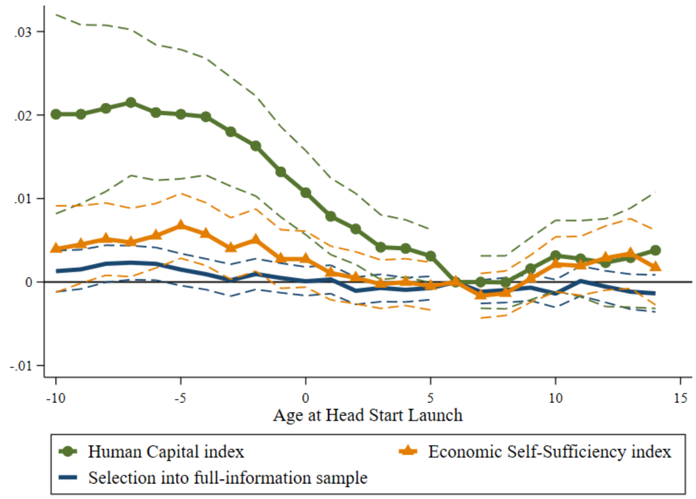
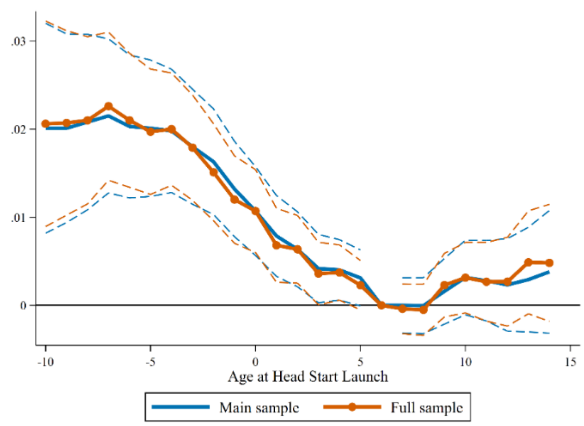
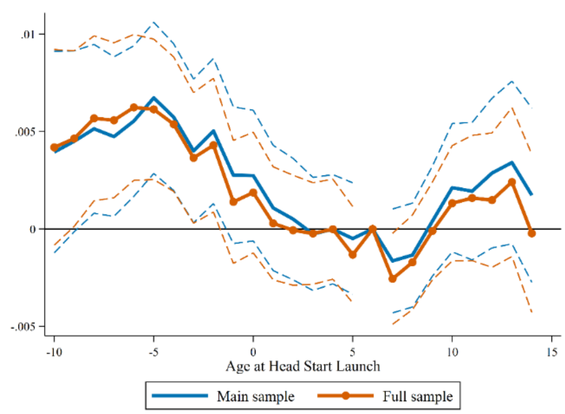


Figure A6. Effect of Head Start on Human Capital and Economic Self-Sufficiency

A. Human Capital Index



B. Economic Self-Sufficiency Index



Notes: Figures show event-study estimates of the effect of Head Start on an index of human capital and economic self-sufficiency as described in section 3 of the paper. The main sample (blue line, no markers) is the analysis sample used in the paper (the estimates replicate Figure 4A and Figure 6A). The "Main sample" includes all individuals born in the United States between 1950-1980 who appear in the 2000 long-form Census or 2001-2018 ACS and can be traced via PIK to their county of birth in the Social Security NUMIDENT file, and who have non-missing responses to all components of the human capital and economic self-sufficiency indices. The "Full sample" (orange line, circle markers) drops the restriction that requires non-missing information on all components of the indices. Dashed lines show 95-percent confidence intervals constructed using standard errors clustered at the county level.

Table A6 examines whether differential mortality drives selection into our analysis sample. Because mortality by 2000 is very rare (our oldest cohorts are 50 and our youngest 20 in 2000) and the event-study estimates very noisy, we summarize differential mortality using the spline specification using the same format as our main paper tables. The results suggest limited role for selection. The slight pre-trend suggests that younger cohorts with access to Head Start were more likely to survive to 2000 on average, and the test of the trend break age 6 is statistically insignificant. Moreover, the ATET of -0.28 is statistically insignificant and small relative to the baseline mean of 98, representing a change of -0.3 percent. In short, little evidence suggests that differential survival affects selection into our analysis sample.

Table A6. The Effect of Head Start on Mortality

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variables	Control mean (s.d.)	ITT estimate (s.e.)	Slope of pre-trend (s.e.)	Test of trend break at age 6 [p-val]	Slope of post-trend (s.e.)	ATET [95% CI]
Survived to 2000	98	-0.041	-0.0072	1.5	0.0089	-0.28
	(13)	(0.033)	(0.0024)	[0.22]	(0.0033)	[-0.79,0.16]

Notes: See Table 1 notes.

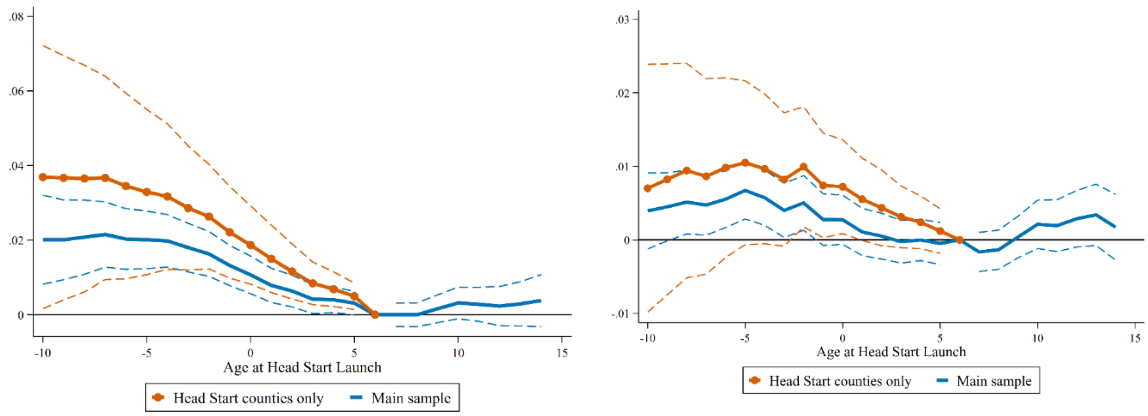
Next we examine the robustness of our results to dropping never-treated counties from regressions. Our motivation for including untreated counties in our main estimating equations is that they help us estimate the state-year fixed effects and, therefore, partial out policy effects or economic shocks common to all counties in the state that vary over time. This is especially important because Head Start launched over a short period. In addition, as demonstrated by Borusyak and Jaravel (2017), event studies with no control group (the never-treated counties in our case) suffer from underidentification of a linear component of the pre-trend. Excluding the control group can induce collinearity between cohort time and event-time. In short, including all counties can be particularly helpful in our context, because the control group is quite large (there are 1,515 never treated counties versus 1,542 treated counties) and Head Start launched quickly (roll-out varies by less than 7 years on average within states). Accordingly, including untreated counties helps a lot with precision.

Figure A7 (next page) shows the robustness of our results to limiting our sample to individuals born in counties that received Head Start between 1965 and 1980 using our main indices as outcomes and the specification in equation (1). To account for the underidentification problem in these specifications, we adopt a parametric event-study approach and replace our event-study dummies for values 7 through 14 with a linear trend in event time. This approach allows us to pin down both year and event-time effects. It is also consistent with our approach to measuring pre-trends throughout the paper, as this specification melds our event-study and three-part spline approaches.

The blue lines replicate our estimates of the human capital and economic self-sufficiency indices shown in Figures 4 and 6, respectively. The estimated effects from our sample with Head Start counties only are noticeably larger but follow a similar pattern, with a phase-in region during the period in which the program was increasing in scope and quality followed by a flattening. Furthermore, the confidence intervals contain our main estimates in all cases; the width of the confidence intervals points to the importance of our comparison counties in our research design, a result of the relatively fast pace at which Head Start rolled out.

Figure A7. Robustness of the Human Capital Index to Dropping non-Head Start Counties

A. Human Capital Index | *B. Economic Self-Sufficiency Index*



To further inform the interpretation of our results, we also ask whether the effect of Head Start in early-adopting communities was different from the effect in late-adopting communities. We might expect to find heterogeneous effects if later-adopting counties were able to scale up—both in terms of quality and enrollment—faster than counties that adopted Head Start in the program's chaotic early years. Alternatively, children in late-adopting counties might experience different economic and social conditions that serve as complements or substitutes for Head Start's programming and alter the magnitude of the treatment effects accordingly (see the heterogeneity analysis in section VII, Table 7). However, if the positive effect of Head Start were driven by only one of these groups, it might raise concerns that our estimates are driven by unobserved confounders that are not picked up by our analysis of the internal validity of our research design (see section IV.B and the preceding analysis in this section).

To test for heterogeneity by date of adoption, we modify spline specification (2) to allow for different effects in early- and late-adopting counties. The results are shown in Table A7 (next page). Column 1 displays estimates for the human capital index. The first two rows provide estimates of the main effect: The slope of the spline from ages -5 to 6 is a highly statistically significant -0.00209 (note that a negative slope implies a positive treatment effect), while we see no statistically significant evidence of a pre-trend in the slope of the spline from ages 6 to 14. We omit the section of the spline from ages -10 to -5 because late-adopting counties are highly imbalanced in this region. The test for heterogeneity is embedded in the second two rows, where we show the estimated slope of each spline interacted with an indicator for launching after the 1969-70 school year. Each coefficient is statistically indistinguishable from 0, suggesting there is no evidence of heterogeneity in Head Start's effect on human capital. Column 2 provides analogous estimates for the economic self-sufficiency index. In this case, we do find evidence of heterogeneity. However, the negative point estimate on the interaction term suggests that while the main effect on economic self-sufficiency is positive, children from late-adopting counties saw even *larger* impacts.

In summary, this analysis suggests that Head Start had similar effects on the human capital accumulation of children in early- and late-adopting counties, but that children from the late-adopting counties were more successful at translating these human capital gains into improvements in their labor-market outcomes and other markers of self-sufficiency. On the other hand, we find no evidence to suggest that our results are driven by unobserved confounders.

Table A7. The Effect of Head Start on Adult Human Capital and Economic Self-Sufficiency, by Early and Late Program Adopters

Outcome	(1) Human capital index	(2) Economic self-sufficiency index
Trend Break Component (Ages -5 to 6)	-0.00209***	-0.000663***
	(0.000331)	(0.000162)
Pre-trend Component (Age 6 to 14)	0.000695	0.000557**
	(0.000434)	(0.000270)
Late Launch x Trend Break Component	-0.000166	-0.00189**
	(0.00133)	(0.000786)
Late Launch x Pre-Trend Component	-0.000987	0.000871
	(0.00112)	(0.000723)

Notes: Late Launch takes the value of 1 for programs beginning after the 1969-70 school year, and 0 otherwise. All specifications include county fixed effects, state-by-birthyear fixed effects, and 1960 covariates interacted with linear year of birth trends.

Finally, Tables A8 and A9 (p. 18-19) examine how different measures of access to Head Start and research designs with our Census/ACS data yield different estimates of the program's effects. Table A8 provides estimates of the effect on our human capital and economic self-sufficiency indices, while Table A9 provides estimates of the effect on high school completion and college attendance. We construct these alternative measures of Head Start access to facilitate a direct comparison of the magnitudes in the literature as best we can, using researcher replication packages when available.

To construct the conventional difference-in-difference estimator used in column 2, we code Head Start exposure as 1 if an individual's county of birth received a significant amount of Head Start funding at any point prior to their age-6 school year, and 0 otherwise. Following Barr and Gibbs (2017), we code a county as receiving a Head Start grant only if it receives an amount greater than the 10^{th} percentile among funded county-year observations. Column 3 uses a measure of exposure to Head Start that captures the share of the three-year period between ages 3 and 5 in which Head Start was present in an individual's county of birth (possible values are 0, 1/3, 2/3, and 1). Columns 4 and 5 measure Head Start access using real funding per capita as in Thompson (2018) and Johnson and Jackson (2019), respectively. Thompson (2018) measures access as average per-year, per-capita Head Start spending in the three years when an individual was between 3 and 6 years old. Johnson and Jackson (2019) measure access as funding per poor 4-year-old, so total funding in each county and year is divided by the product of the county-year population of 4-year-olds and the county-level child poverty rate from the 1970 Census (see Johnson and Jackson 2019, appendix E, page 10). To make the estimates in columns 4 and 5 more easily comparable to each other and to the estimates in columns 2-3, we present estimates multiplied by the average Head Start funding per capita. Each of these estimates can therefore be interpreted as the effect of exposure to an average level of Head Start funding relative to children without access. We are not able to use the family fixed effects estimator with our data, because we do not observe siblings in the Census/ACS data for most adults.

There are several key take-aways from this analysis:

First, a more standard differences-in-differences approach (DD, cf. Barr and Gibbs 2017) takes a variance-weighted average of treatment effects across all years after implementation, including the early years in which programs were small and of lower quality (see our event studies). The DD estimators in columns 1 and 2 yield an ATET of around 1.2 to 1.3 percentage points for high school completion (panel B of columns 2 and 3 in Table A9, divided by our first-stage estimate of 0.149). In contrast, our spline estimate, which captures the effect of a full-capacity and mature Head Start program at -5 relative to unexposed cohorts, is 2.4 percentage points (column 1 of panel B of Table A9, divided by 0.149). For college attendance, DD yields an ATET of around 3.6 to 4 percentage points for college attendance versus 5.4 percentage points using our spline. Using our data, the DD approach tends to yield **smaller** estimates than what we obtain using the spline.

Second, a comparison of our estimates (where the measure of Head Start access is based on the roll-out design) to those which use variation in funding levels constructed following the procedures of Thompson (2018) and Johnson and Jackson (2019) (Tables A8-A9, columns 4-5) highlights some weaknesses of the National Archives grant data. Although not directly comparable to columns 1-3, scaling the coefficients by the average Head Start funding experienced by the average 4-year old shows that the effect of dollars appear **smaller**. We believe this attenuation reflects two factors: (1) Head Start grant funding records contain sizable measurement error in grant amounts,[4] and (2) grant amounts are not highly correlated with program capacity or quality. On (2), Figure A8 (subsequent page) shows a comparison of these measures, averaged over our definition of event time among ever-funded counties. These four measures are correlated, but initial grants tend to be very large because they help build infrastructure development (building or remodeling a Head Start building, whereas later grants tend to fund greater capacity and program quality). In addition, a substantial but unknown portion of Head Start grants in the early years funded summer Head Start programs, which were very large but short-lived and not likely to drive long-run effects even by the program's architects. Smaller grants sizes in later years do not mean that the full-year Head Start program was shrinking.

Finally, a comparison of Tables A8-A9 panels A and C to panels B and D highlight the influence of state-by-birth-year fixed effects on all estimates. Including state-by-birth-year fixed effects (which is difficult for papers that use longitudinal data due to small samples sizes) generally decreases the magnitudes of the estimates from DD and funding per capita approaches. However, including state-by-birth-year fixed effects tends to **increase** the estimates when using the spline. Human capital estimates are more sensitive to this specification choice, perhaps because state policies surrounding education were changing much more for our cohorts of interest (Cascio, Gordon, Lewis, & Reber, 2010). This pattern also relates to the fact that the spline is able to account for parallel pre-trends not holding exactly. After adjusting for state-by-birth-year fixed effects, Head Start counties are more likely to experience an *off-setting* pre-trend (i.e., a trend that works against a positive treatment effect). Our spline specification adjusts for these pre-trends whereas the other estimators do not.

---

[4] Measurement error in this grant data is well documented. Ludwig and Miller (2007) note that "the accuracy of these data [on Head Start funding from the National Archives] is less than perfect given poor documentation and some obvious errors. In the end, only data from 1968 and 1972 were usable, in the sense that the electronic data matched published Head Start and other federal spending figures at the national and state levels. Even here measurement error arises from complications such as providers that run Head Start programs in multiple counties but are listed as receiving federal funds only in the county with the organization's headquarters." Barr and Gibbs (2017) document similar problems and went to great lengths to clean the National Archives data. Measurement error in National Archives grant data for other War on Poverty programs has led other papers to use a roll-out design rather than funding per capita (Bailey, 2012; Bailey & Goodman-Bacon, 2015; Bailey, Hoynes, Rossin-Slater, & Walker, 2020; Bailey, Malkova, & McLaren, 2018). Relatedly, Barr and Gibbs (2017) treatment variable "turns on" once a county gets a grant that is larger than the 10th percentile in the Head Start grant per capita distribution.

In summary, Tables A8 and A9 underscore the reasoning for our approach using a roll-out design and a spline to capture phased implementation. Our approach (1) minimizes the impact of measurement error in access to Head Start (by using start dates rather than historical grant) while explicitly (2) allowing for the gradual implementation of Head Start and (3) accounting for differential (off-setting) pre-trends and the state-level changes in education policy and economic circumstances with state-by-birth-year fixed effects. Using other approaches generally yields much **smaller** estimates.

On the other hand, Figure 5 shows that the literature's estimates of the ATET are generally **larger** than our estimates. There are several reasons why this could be: (1) It is possible that our estimates of program take-up from the 1970 Census are too large (see discussion on pages 17-18). Although our estimates are in line with others in the literature, using a smaller scaling factor could increase the magnitudes of our ATETs. (2) Treatment effects may be heterogeneous, and the LATEs in the literature are specific to different sets of compliers. Some studies explicitly focus on specific groups (e.g., Barr and Gibbs focus on women or Johnson and Jackson focus only on a high impact sample). (3) Attrition in longitudinal surveys may play a role in other studies, whereas it does not in the Census/ACS (see Appendix Figure A6). (4) Limitations in the data and research designs in other studies may lead to biased and imprecise estimates. We do not know which combination of the explanations for differences holds, so provide information here so that readers are informed and can judge for themselves.

Table A8. The Sensitivity of Estimates to using Different Measures of Head Start Access and Controls for the Human Capital Index and the Economic Self-Sufficiency Index

	(1) This paper (spline at -5)	(2) Barr and Gibbs (2017) difference-in-differences	(3) Exposure model (share of age 3-5 with access)	(4) Thompson (2018) funding per capita	(5) Johnson and Jackson (2019) funding per capita
		Human Capital Index			
A. Excluding state-by-birth-year FEs					
Effect of Head Start	1.9	1.3	1.4	0.11	0.11
	(0.74)	(0.41)	(0.50)	(0.055)	(0.069)
B. Including state-by-birth-year FEs					
Effect of Head Start	2.7	1.1	1.1	0.016	0.041
	(0.54)	(0.28)	(0.29)	(0.050)	(0.049)
		Economic Self-Sufficiency Index			
C. Excluding state-by-birth-year FEs					
Effect of Head Start	1.3	0.41	0.26	0.0058	0.013
	(0.47)	(0.17)	(0.18)	(0.025)	(0.019)
D. Including state-by-birth-year FEs					
Effect of Head Start	1.4	0.29	0.24	-0.026	-0.0032
	(0.35)	(0.12)	(0.12)	(0.026)	(0.014)

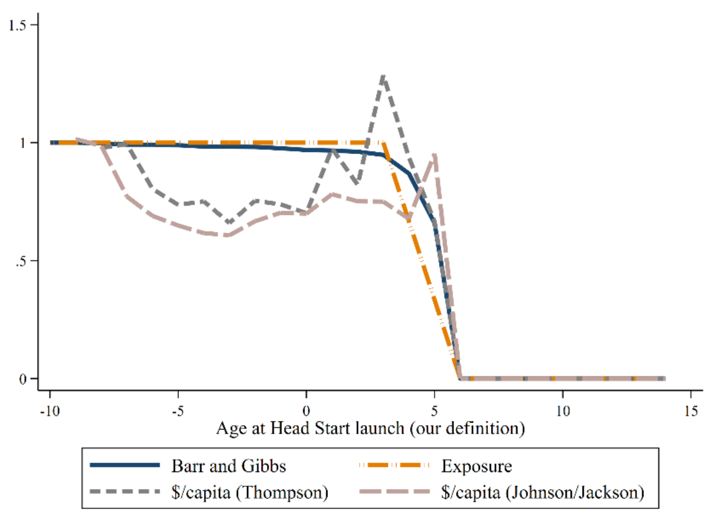
Notes: There are 22,480,000 individuals (rounded for disclosure) in each regression. Estimates are presented in percentage-point units. Panels A and C present the results from specifications that exclude state-by-year-of-birth fixed effects as is necessary in longitudinal samples. Panels B and D presents the results from specifications that include state-by-year-of-birth fixed effects. The only difference across columns is the specification of the variable used to measure access to Head Start. Column 1 repeats the ITT estimate from our spline specification, as shown in the Table 1 of the paper. Columns 2-5 use the definitions of Head Start exposure as described in the text and in Figure A8.

Table A9. The Sensitivity of Estimates to using Different Measures of Head Start Access and Controls for High School/GED Completion and College Attendance

	(1) This paper (spline at -5)	(2) Barr and Gibbs (2017) difference-in-differences	(3) Exposure model (share of age 3-5 with access)	(4) Thompson (2018) funding per capita	(5) Johnson and Jackson (2019) funding per capita
	Completed High School or GED				
A. Excluding state-by-birth-year FEs					
Effect of Head Start	0.059	0.28	0.32	0.014	0.018
	(0.16)	(0.092)	(0.13)	(0.016)	(0.0095)
B. Including state-by-birth-year FEs					
Effect of Head Start	0.36	0.19	0.18	-0.020	-0.00013
	(0.17)	(0.061)	(0.076)	(0.016)	(0.0067)
	Attended some college				
C. Excluding state-by-birth-year FEs					
Effect of Head Start	0.092	0.61	0.88	0.027	0.049
	(0.50)	(0.28)	(0.42)	(0.031)	(0.040)
D. Including state-by-birth-year FEs					
Effect of Head Start	0.80	0.54	0.60	-0.0066	0.0056
	(0.41)	(0.14)	(0.16)	(0.031)	(0.028)

Notes: See Table A8 notes.

Figure A8. Comparison of Measures of Head Start Access



Notes: We construct the Barr and Gibbs (2017) measure of Head Start exposure as 1 if an individual's county of birth received a significant amount of Head Start funding (specifically, an amount greater than the 10^{th} percentile among funded county-year observations) at any point prior to their age-6 school year, and 0 otherwise. Exposure variable captures the share of the three-year period between ages 3 and 5 in which Head Start was present in an individual's county of birth (possible values are 0, 1/3, 2/3, and 1). Measures of $/capita for Thompson (2018) and Johnson and Jackson (2019) are taken from their replication packages. Figure shows unweighted average among ever-funded counties by event time.

9. Additional Estimates

The following tables present estimates by race and sex by race for both human capital and economic self-sufficiency.

Table A10. The Effect of Head Start on Adult Human Capital, by Sex and Race

	(1) Control mean (s.d.)	(2) ITT estimate (s.e.)	(3) Slope of pre-trend (s.e.)	(4) Test of trend break at age 6 [p-val]	(5) Slope of post-trend (s.e.)	(6) ATET [95% CI]	(7) ATET % change
A.White females							
Human capital		2.0	0.050	16	-0.048	15	
		(0.44)	(0.044)	[<0.001]	(0.095)	[7.7,25]	
Human capital (pre-trend adjusted)		2.5			-0.098	20	
		(0.64)			(0.096)	[9.0,33]	
Subindices (pre-trend adjusted)							
Completed high school/GED	94	0.22	-0.0082	1.5	-0.025	1.7	1.8%
	(24)	(0.18)	(0.012)	[0.23]	(0.031)	[-0.77,5.0]	
Attended some college	68	0.44	-0.038	1.0	-0.0053	3.4	5.0%
	(47)	(0.43)	(0.029)	[0.31]	(0.060)	[-3.2,11]	
Completed college	32	1.7	0.030	15	-0.037	13	0.39
	(47)	(0.43)	(0.031)	[<0.001]	(0.061)	[6.0,24]	
Prof. or doc. degree	2.5	0.30	0.013	7.1	-0.013	2.3	93%
	(16)	(0.11)	0.00719	[0.0078]	(0.013)	[0.66,4.8]	
Years of schooling	14	0.098	0.0025	19	-0.0033	0.76	5.5%
	(2.4)	(0.0226)	(0.0016)	[<0.001]	(0.0033)	[0.38,1.3]	
Has a professional job	38	1.8	0.079	21	-0.081	14	37%
	(49)	(0.40)	(0.027)	[<0.001]	(0.052)	[7.7,25]	
B. White males							
Human capital		2.4	-0.0017	14	0.017	18	
		(0.41)	(0.048)	[<0.001]	(0.081)	[11,29]	
Human capital (pre-trend adjusted)		2.4			0.019	18	
		(0.63)			(0.086)	[8.5,31]	
Subindices (pre-trend adjusted)							
Completed high school/GED	92	0.33	-0.0032	2.0	-0.0063	2.4	2.6%
	(28)	(0.23)	(0.014)	[0.15]	(0.034)	[-0.96,6.3]	
Attended some college	62	1.0	-0.0078	3.8	0.012	7.5	12%
	(49)	(0.52)	(0.033)	[0.052]	(0.061)	[-0.96,17]	
Completed college	31	1.6	0.0027	13	-0.0093	12	38%
	(46)	(0.43)	(0.033)	[<0.001]	(0.055)	[5.0,20]	
Prof. or doc. degree	3.5	0.44	0.011	7.0	0.012	3.2	94%
	(18)	(0.17)	(0.012)	[0.0083]	(0.017)	[0.85,6.4]	
Years of schooling	14	0.082	0.00034	13	0.00051	0.61	4.5%
	(2.5)	(0.022)	(0.0017)	[<0.001]	(0.0031)	[0.27,1.1]	
Has a professional job	35	1.0	-0.027	6.6	0.024	7.4	21%
	(48)	(0.39)	(0.029)	[0.010]	(0.047)	[1.5,15]	

Notes. Column 6 scales by an estimated takeup of 0.130 (s.e. 0.023) for white females and 0.135 (s.e. 0.023) for white males. For the subindices, For the subindices, ITT effects (column 2) and post-trends (column 5) are adjusted to net out estimated linear pre-trend (column 3). See also Table 1 notes.

Table A10. The Effect of Head Start on Adult Human Capital, by Sex and Race (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Control mean (s.d.)	ITT estimate (s.e.)	Slope of pre-trend (s.e.)	Test of trend break at age 6 [p-val]	Slope of post-trend (s.e.)	ATET [95% CI]	ATET % change
C. Nonwhite females							
Human capital		0.71	0.29	8.5	-0.065	3.4	
		(0.7)	(0.093)	[0.0036]	(0.17)	[-4.3,11]	
Human capital (pre-trend adjusted)		3.9			-0.36	19	
		(1.3)			(0.20)	[5.1,34]	
Subindices (pre-trend adjusted)							
Completed high school/GED	88	1.2	0.12	3.5	-0.090	5.9	6.7%
	(33)	(0.65)	(0.046)	[0.060]	(0.081)	[0.41,13]	
Attended some college	62	1.4	0.089	2.5	-0.10	6.9	11%
	(49)	(0.91)	(0.062)	[0.11]	(0.13)	[-1.5,16]	
Completed college	26	2.5	0.14	11	-0.24	12	46%
	(44)	(0.75)	(0.052)	[0.0010]	(0.12)	[4.5,21]	
Prof. or doc. degree	2.3	0.30	0.021	1.4	-0.049	1.4	61%
	(15)	(0.25)	(0.015)	[0.24]	(0.037)	[-0.8,4.1]	
Years of schooling	13	0.12	0.0091	5.6	-0.0094	0.56	4.2%
	(2.5)	(0.049)	(0.0035)	[0.018]	(0.0071)	[0.058,1.1]	
Has a professional job	31	1.7	0.15	4.5	-0.14	8	27%
	(46)	(0.80)	(0.053)	[0.033]	(0.10)	[0.8,16]	
D. Nonwhite males							
Human capital		0.94	0.20	4.5	0.034	4.2	
		(0.8)	(0.095)	[0.034]	(0.16)	[-2.9,12]	
Human capital (pre-trend adjusted)		3.2			-0.17	14	
		(1.5)			(0.19)	[1.4,30]	
Subindices (pre-trend adjusted)							
Completed high school/GED	86	0.63	0.069	0.75	-0.022	2.8	3.3%
	(34)	(0.73)	(0.050)	[0.39]	(0.083)	[-3.8,10]	
Attended some college	56	0.94	0.057	0.91	-0.10	4.2	8%
	(50)	(0.99)	(0.065)	[0.34]	(0.13)	[-5.5,14]	
Completed college	24	1.8	0.10	5.4	-0.089	8	34%
	(43)	(0.78)	(0.051)	[0.020]	(0.11)	[1.2,16]	
Prof. or doc. degree	2.9	0.57	0.029	3.2	-0.036	2.5	89%
	(17)	(0.32)	(0.020)	[0.074]	(0.039)	[-0.28,5.6]	
Years of schooling	13	0.13	0.0098	7.0	-0.0072	0.60	4.5%
	(2.6)	(0.050)	(0.0034)	[0.0080]	(0.0066)	[0.15,1.1]	
Has a professional job	27	0.75	0.032	0.61	-0.0025	3.4	13%
	(44)	(0.96)	(0.060)	[0.44]	(0.13)	[-5.3,12]	

Notes: Column 6 scales by an estimated takeup of 0.208 (s.e. 0.029) for nonwhite females and 0.223 (s.e. 0.029) for nonwhite males. For the subindices, ITT effects (column 2) and post-trends (column 5) are adjusted to net out estimated linear pre-trend (column 3). See also Table 1 notes.

Table A11. The Effect of Head Start on Adult Economic Self-Sufficiency, by Race

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Control mean (s.d.)	ITT estimate (s.e.)	Slope of pre-trend (s.e.)	Test of trend break at age 6 [p-val]	Slope of post-trend (s.e.)	ATET [95% CI]	ATET % change
A. White							
ESS index		0.68	0.069	15	0.0072	5.1	
		(0.18)	(0.027)	[<0.001]	(0.041)	[2.3,8.9]	
ESS index (pre-trend adjusted)		1.4			-0.062	10.9	
		(0.38)			(0.046)	[4.8,19]	
Subindices (pre-trend adjusted)							
Worked last year	86	0.77	0.042	10	-0.086	5.9	6.8%
	(35)	(0.24)	(0.017)	[0.0013]	(0.029)	[2.7,12]	
Weeks worked last year	41	0.39	0.023	8.6	-0.029	3.0	7.2%
	(19)	(0.13)	(0.0094)	[0.0035]	(0.017)	[1.0,5.7]	
Usual hours works per week	36	0.51	0.029	15	-0.023	3.8	10.7%
	(18)	(0.13)	(0.0085)	[<0.001]	(0.013)	[1.8,7.2]	
Log labor income	11	0.0037	-0.000015	0.47	0.0029	0.028	
	(0.97)	(0.0054)	(0.00041)	[0.49]	(0.00071)	[-0.049,0.12]	
Log family income/poverty	5.9	0.0086	0.00033	2.0	0.00045	0.065	
	(0.89)	(0.0061)	(0.00040)	[0.16]	(0.00094)	[-0.034,0.17]	
In poverty*	10	-0.33	-0.021	4.4	-0.0095	-2.5	-27%
	(29)	(0.16)	(0.010)	[0.036]	(0.026)	[-5.4,-0.33]	
Received public assistance*	11	-0.57	-0.0064	14	0.0019	-4.3	-40%
	(31)	(0.16)	(0.010)	[<0.001]	(0.019)	[-7.8,-2.0]	
B. Nonwhite							
ESS index		0.32	0.025	0.56	0.14	1.5	
		(0.40)	(0.052)	[0.46]	(0.078)	[-2.2,5.4]	
ESS index (pre-trend adjusted)		0.60			0.12	2.8	
		(0.80)			(0.092)	[-4.7,11]	
Subindices(pre-trend adjusted)							
Worked last year	78	-0.27	-0.014	0.23	0.0009	-1.3	-1.6%
	(41)	(0.57)	(0.037)	[0.63]	(0.061)	[-7.1,4.8]	
Weeks worked last year	37	-0.096	-0.0051	0.091	0.027	-0.45	-1.2%
	(22)	(0.32)	(0.021)	[0.76]	(0.032)	[-3.4,2.4]	
Usual hours works per week	32	-0.054	-0.0070	0.042	0.014	-0.25	-0.8%
	(19)	(0.26)	(0.017)	[0.84]	(0.029)	[-2.9,2.2]	
Log labor income	10	0.0097	0.00066	0.48	0.0010	0.045	
	(0.99)	(0.014)	(0.00094)	[0.49]	(0.0016)	[-0.086,0.17]	
Log family income/poverty	5.5	0.0012	0.00042	0.007	0.0014	0.0058	
	(1.05)	(0.015)	(0.0010)	[0.93]	(0.0021)	[-0.13,0.15]	
In poverty*	21	-0.52	-0.036	0.96	-0.046	-2.4	-12%
	(41)	(0.53)	(0.036)	[0.33]	(0.066)	[-7.6,2.5]	
Received public assistance*	18	0.25	0.0039	0.23	-0.063	1.2	6.5%
	(38)	(0.52)	(0.033)	[0.63]	(0.050)	[-4.2,5.9]	

Notes: Column 6 scales by an estimated take up of 0.132 (s.e. 0.023) for whites and 0.214 (s.e. 0.028) for nonwhites (Appendix Table A5). For the subindices, ITT effects (column 2) and post-trends (column 5) are adjusted to net out estimated linear pre-trend (column 3). See also Table 1 notes.

Table A12. The Effect of Head Start on Adult Economic Self-Sufficiency, by Sex and Race

	(1) Control mean (s.d.)	(2) ITT estimate (s.e.)	(3) Slope of pre-trend (s.e.)	(4) Test of trend break at age 6 [p-val]	(5) Slope of post-trend (s.e.)	(6) ATET [95% CI]	(7) ATET % change
A. White females							
ESS index		0.73	0.10	13	-0.0018	5.6	
		(0.23)	(0.038)	[<0.001]	(0.054)	[1.8,11]	
ESS index (pre-trend adjusted)		1.9			-0.11	14	
		(0.52)			(0.070)	[6.0,26]	
Subindices (pre-trend adjusted)							
Worked last year	80	0.89	0.063	5.5	-0.16	6.9	8.6%
	(40)	(0.38)	(0.027)	[0.019]	(0.051)	[1.7,15]	
Weeks worked last year	37	0.43	0.034	4.4	-0.073	3.3	8.9%
	(21)	(0.21)	(0.015)	[0.037]	(0.028)	[0.28,7.3]	
Usual hours works per week	30	0.46	0.035	5.5	-0.064	3.5	12%
	(18)	(0.19)	(0.013)	[0.019]	(0.024)	[0.43,7.6]	
Log labor income	10	0.0090	0.00054	1.2	0.0015	0.069	
	(1.0)	(0.0083)	(0.00058)	[0.28]	(0.0011)	[-0.044,0.22]	
Log family income/poverty	5.8	0.010	0.00042	1.6	0.00056	0.078	
	(0.93)	(0.0081)	(0.00048)	[0.21]	(0.0012)	[-0.054,0.22]	
In poverty*	12	-0.50	-0.033	4.9	-0.0023	-3.8	-33%
	(32)	(0.23)	(0.015)	[0.027]	(0.036)	[-8.1,-0.67]	
Received public assistance*	11	-0.28	0.011	1.7	0.019	-2.2	-19%
	(32)	(0.22)	(0.014)	[0.19]	(0.025)	[-5.9,1.2]	
B. White males							
ESS index		0.49	-0.00342	1.0	0.057	3.6	
		(0.20)	(0.032)	[0.31]	(0.055)	[0.71,7.6]	
ESS index (pre-trend adjusted)		0.45			0.060	3.3	
		(0.45)			(0.054)	[-3.3,11]	
Subindices (pre-trend adjusted)							
Worked last year	92	0.44	0.010	4.9	-0.018	3.3	3.6%
	(28)	(0.20)	(0.013)	[0.028]	(0.022)	[0.082,7.2]	
Weeks worked last year	45	0.21	0.0052	3.4	0.015	1.6	3.5%
	(16)	(0.11)	(0.0078)	[0.0657]	(0.013)	[-0.088,3.5]	
Usual hours works per week	41	0.39	0.015	10	0.023	2.9	7.0%
	(16)	(0.12)	(0.0080)	[0.001]	(0.014)	[1.1,5.4]	
Log labor income	11	-0.0066	-0.00087	0.87	0.0047	-0.049	
	(0.86)	(0.0071)	(0.00051)	[0.35]	(0.00091)	[-0.17,0.052]	
Log family income/poverty	5.9	0.0043	0.000017	0.44	0.00047	0.032	
	(0.84)	(0.0066)	(0.00048)	[0.51]	(0.00095)	[-0.065,0.14]	
In poverty*	7.4	-0.064	-0.0042	0.12	-0.017	-0.47	-6.4%
	(26)	(0.18)	(0.011)	[0.72]	(0.025)	[-3.3,2.3]	
Received public assistance*	10	-0.82	-0.022	16	-0.013	-6.1	-59%
	(30)	(0.21)	(0.014)	[<0.001]	(0.025)	[-11,-3.1]	

Notes: Column 6 scales by an estimated takeup of 0.130 (s.e. 0.023) for white females and 0.135 (s.e. 0.023) for white males. For the subindices, ITT effects (column 2) and post-trends (column 5) are adjusted to net out estimated linear pre-trend (column 3). See also Table 1 notes.

Table A12. The Effect of Head Start on Adult Economic Self-Sufficiency, by Sex and Race (Continued)

	(1) Control mean (s.d.)	(2) ITT estimate (s.e.)	(3) Slope of pre-trend (s.e.)	(4) Test of trend break at age 6 [p-val]	(5) Slope of post-trend (s.e.)	(6) ATET [95% CI]	(7) ATET % change
C. Nonwhite females							
ESS index		0.53	0.076	1.7	0.055	2.5	
		(0.50)	(0.069)	[0.19]	(0.10)	[-2.5,7.6]	
ESS index (pre-trend adjusted)		1.4			-0.022	6.6	
		(1.0)			(0.12)	[-4.0,17]	
Subindices (pre-trend adjusted)							
Worked last year	75	0.17	0.029	0.044	-0.17	0.8	1.1%
	(43)	(0.79)	(0.053)	[0.83]	(0.083)	[-6.2,8.6]	
Weeks worked last year	35	0.11	0.0094	0.069	-0.033	0.54	1.6%
	(23)	(0.43)	(0.028)	[0.79]	(0.044)	[-3.5,4.8]	
Usual hours works per week	29	0.19	0.015	0.29	-0.049	0.9	3.1%
	(19)	(0.35)	(0.023)	[0.59]	(0.038)	[-2.6,4.5]	
Log labor income	10	0.027	0.0018	1.9	-0.00051	0.13	
	(1.0)	(0.019)	(0.0012)	[0.16]	(0.0024)	[-0.043,0.34]	
Log family income/poverty	5.4	0.015	0.0015	0.58	0.0017	0.071	
	(1.1)	(0.019)	(0.0013)	[0.45]	(0.0027)	[-0.13,0.26]	
In poverty*	25	-1.3	-0.096	2.8	-0.0040	-6.1	-25%
	(43)	(0.75)	(0.049)	[0.092]	(0.095)	[-14,0.41]	
Received public assistance*	19	0.34	0.012	0.26	-0.075	1.6	8.6%
	(39)	(0.67)	(0.044)	[0.61]	(0.068)	[-4.5,8.6]	
D. Nonwhite males							
ESS index		-0.27	-0.055	0.65	0.22	-1.2	
		(0.50)	(0.071)	[0.42]	(0.10)	[-6.0,3.3]	
ESS index (pre-trend adjusted)		-0.87			0.27	-3.9	
		(1.1)			(0.12)	[-14,5.7]	
Subindices (pre-trend adjusted)							
Worked last year	83	-0.83	-0.064	1.4	0.19	-3.7	-4.5%
	(38)	(0.70)	(0.047)	[0.23]	(0.073)	[-11,3.3]	
Weeks worked last year	39	-0.38	-0.024	1.0	0.093	-1.7	-4.3%
	(20)	(0.37)	(0.025)	[0.31]	(0.040)	[-5.1,1.6]	
Usual hours works per week	35	-0.40	-0.035	1.3	0.072	-1.8	-5.1%
	(19)	(0.35)	(0.023)	[0.25]	(0.037)	[-5.2,1.3]	
Log labor income	11	-0.014	-0.00062	0.52	0.00089	-0.062	
	(0.95)	(0.019)	(0.0013)	[0.47]	(0.0020)	[-0.24,0.10]	
Log family income/poverty	5.6	-0.015	-0.00090	0.68	0.00048	-0.07	
	(0.97)	(0.019)	(0.0013)	[0.41]	(0.0023)	[-0.24,0.09]	
In poverty*	16	0.61	0.049	0.87	-0.084	2.7	17%
	(37)	(0.66)	(0.045)	[0.35]	(0.075)	[-2.8,9.1]	
Received public assistance*	16	0.40	0.0086	0.29	-0.029	1.8	11%
	(37)	(0.74)	(0.049)	[0.59]	(0.072)	[-5.5,8.4]	

Notes: Column 6 scales by an estimated takeup of 0.208 (s.e. 0.029) for nonwhite females and 0.223 (s.e. 0.029) for nonwhite males. For the subindices, ITT effects (column 2) and post-trends (column 5) are adjusted to net out estimated linear pre-trend (column 3). See also Table 1 notes.

10. Additional Estimates Without Pre-Trend Adjustment

The following tables present estimates of the effect on incarceration and heterogeneous effects on the human capital and economic self-sufficiency indices without adjustments for pre-trends.

Table A13. The Effect of Head Start on Incarceration in Adulthood

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Control mean (s.d.)	ITT estimate (s.e.)	Slope of pre-trend (s.e.)	Test of trend break at age 6 [p-val]	Slope of post-trend (s.e.)	ATET [95% CI]	ATET % change
Full Sample	1.2	-0.036	-0.0021	0.92	-0.0003	-0.24	-19.9%
	(11)	(0.030)	(0.0046)	[0.34]	(0.0066)	[-0.67,0.20]	
White females	0.17	0.024	0.00080	0.76	0.0025	0.18	107.0%
	(4.1)	(0.018)	(0.0025)	[0.38]	(0.0041)	[-0.066,0.49]	
White males	1.3	-0.062	-0.0095	2.47	0.0047	-0.46	-35.4%
	(11)	(0.052)	(0.0073)	[0.12]	(0.011)	[-1.4,0.27]	
Nonwhite females	0.50	0.020	-0.014	0.46	-0.0173	0.10	19.3%
	(7.1)	(0.090)	(0.013)	[0.50]	(0.018)	[-0.75,1.0]	
Nonwhite males	6.5	0.36	-0.018	0.050	0.058	1.6	25.0%
	(25)	(0.35)	(0.055)	[0.82]	(0.069)	[-1.9,5.2]	

Notes: Dependent variable is a binary variable for being incarcerated at the time of observation. Incarceration is measured using group quarters residence in 2000 Census and 2005-2018 American Community Survey. All estimates are shown in percentage point units. Column 6 scales by an estimated takeup of 0.149 (s.e. 0.022) for full sample, 0.130 (s.e. 0.023) for white females, 0.135 (s.e. 0.023) for white males, 0.208 (s.e. 0.029) for nonwhite females, and 0.223 (s.e. 0.029) for nonwhite males. See also Table 1 notes.

Table A14. Heterogeneity in the Effect of Head Start, by Local Programs and Circumstances

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Intention-to-treat effect			Effects on pre-school enrollment		Average treatment effect on treated children (ATET)	
	Above median	Below median	F-test of difference [p-value]	Above median	Below median	Above median	Below median
A. Human capital index							
Medicaid exposure	3.0	1.4	5.1	12	18	24	8.2
	(0.55)	(0.40)	[0.02]	(2.4)	(3.0)	[15,43]	[3.2,15]
CHC exposure	2.5	1.7	2.6	20	11	12	16
	(0.45)	(0.39)	[0.10]	(3.4)	(2.5)	[7.7,20]	[8.7,28]
Food Stamps exposure	1.4	2.8	9.8	12	17	11	17
	(0.40)	(0.45)	[0.0017]	(2.6)	(3.2)	[4.7,20]	[11,29]
Poorest 300 counties	2.7	2.1	0.28	7.3	13	37	17
	(1.1)	(0.36)	[0.59]	(2.6)	(2.3)	[4.8,128]	[10,27]
Predicted economic	2.5	1.8	2.304	14	13	18	14
growth	(0.48)	(0.39)	[0.13]	(3.3)	(2.5)	[10,34]	[8.1,24]
B. Economic self-sufficiency index							
Medicaid exposure	0.79	0.68	0.10	12	18	6.5	3.9
	(0.29)	(0.22)	[0.75]	(2.4)	(3.0)	[1.9,13]	[1.3,7.3]
CHC exposure	1.1	0.52	5.9	20	11	5.2	4.8
	(0.20)	(0.20)	[0.016]	(3.4)	(2.5)	[2.9,9]	[1.1,11]
Food Stamps exposure	0.43	1.1	8.3	12	17	3.5	6.4
	(0.20)	(0.21)	[0.0039]	(2.6)	(3.2)	[0.20,8.0]	[3.6,12]
Poorest 300 counties	0.52	0.76	0.08641	7.3	13	7.1	6.1
	(0.83)	(0.17)	[0.77]	(2.6)	(2.3)	[-21,41]	[3.1,11]
Predicted economic	0.85	0.65	0.63	14	13	6.1	5.2
growth	(0.22)	(0.20)	[0.43]	(3.3)	(2.5)	[2.9,13]	[2.0,11]
C. Incarceration							
Medicaid exposure	-0.084	0.0039	0.052	12	18	-0.68	0.022
	(0.05)	(0.039)	[0.82]	(2.4)	(3.0)	[-1.5,0.12]	[-0.43,0.49]
CHC exposure	-0.055	-0.019	0.44	20	11	-0.27	-0.17
	(0.036)	(0.033)	[0.51]	(3.4)	(2.5)	[-0.64,0.088]	[-0.84,0.48]
Food Stamps exposure	0.024	-0.093	0.83	12	17	0.20	-0.555
	(0.035)	(0.035)	[0.36]	(2.6)	(3.2)	[-0.3,0.93]	[-1.1,-0.12]
Poorest 300 counties	0.67	-0.050	2.8	7.3	13	9.1	-0.40
	(0.22)	(0.029)	[0.094]	(2.6)	(2.3)	[2.7,33]	[-0.93,0.10]
Predicted economic	-0.075	-0.0054	0.65	14	13	-0.54	-0.043
growth	(0.039)	(0.034)	[0.42]	(3.3)	(2.5)	[-1.3,0.040]	[-0.56,0.53]

Notes: Point estimates and standard errors are in percentage-point units. 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."

11. Cost-Benefit Analysis of Head Start with the NLSY-79

A full accounting of the costs and benefits of Head Start is outside the scope of this paper. However, for comparison purposes, we compute the cumulative benefits of Head Start on economic opportunity through the program's cumulative effects on earnings *potential*. This is important in our context because Head Start appears to influence men and women's work effort, making the sample of wage earners selected.

To circumvent the effects of this selection, we use a potential earnings framework following Neal and Johnson (1996) and is directly comparable with Deming (2009). Like Deming, we use the *National Longitudinal Survey of Youth 1979* (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.

We use observations on respondents' labor market wage income between 2002 and 2014, when they are between ages 35 and 57 years old, adjusted to be in 2013 dollars. We implement the following regression:

$$\ln\left(Wages_i\right) = \boldsymbol{HC}_i'\boldsymbol{\beta_1} + \boldsymbol{ESS}_i'\boldsymbol{\beta_2} + \boldsymbol{X}_i'\boldsymbol{\beta_3} + \varepsilon_i,$$

Where $\boldsymbol{X_i}$ is a vector of race, gender, age, survey year dummy variables and a quadratic in the respondent's standardized and age-normalized AFQT score (as in Deming 2009, Neal and Johnson 1996). In some specifications, $\boldsymbol{X_i}$ includes Deming's covariates of age-19 outcomes as regressors to account for potential sources of omitted variables bias, which has a negligible effect on our calculations. We also include new outcomes that were contained in our human capital (***HC***) and economic self-sufficiency indices (***ESS***). Note, however, that we omit components of the ***ESS*** that are directly related to log wages such as poverty and log family income. The resulting regression coefficients capture the importance of each index component after accounting flexibly for AFQT. Table A15 shows that the results are generally very similar with and without Deming's covariates (columns 1-3 versus columns 4-6). Panel A summarizes the internal rate of return (IRR) using the estimated cost of Head Start per student of around $5,400 and the realized human capital and self-sufficiency gains at ages 25 to 64. We present the regression estimates underlying the calculations in panel B.

In terms of the human capital and self-sufficiency returns only, we find an IRR to Head Start of 13.7 percent averaged over men and women (Panel A, column 1). The IRR is nearly identical for men and women (columns 2-3). Note, that this approach calculates <u>only some of the private benefits</u> that accrue to individuals and does not include benefits through improvements in outcomes not measured here. For instance, the extent to which more education engenders better health, longevity, or well-being is ignored in these calculations.

The human capital and self-sufficiency calculations capture a portion of the benefit that accrues to students who attended Head Start. In addition, the program delivered a fiscal externality as a result of increases in labor supply and reductions in public-assistance spending. The 2000 Census and 2001-2013 ACS suggest that the average amount received by public assistance recipients between ages 25 and 54 was $8,700 per year in 2013 dollars, which is 15 percent smaller than in the Survey of Income and Program Participation (SIPP) of $9,967 due to misreporting. Much of this cost is driven by receipt of Social Security Disability Insurance, consistent with the effect being driven by men. Using the SIPP calculation, the IRR to the government on putting one child through Head Start is 2.3% (3.5% for men, -0.1% for women) if the *only* returns to Head Start were in savings in public assistance expenditures. This is likely an underestimate because a substantial cost of SSDI is receipt of Medicare (Autor & Duggan 2006), which is unlikely to be included in survey respondents' tally of income from public sources.

To estimate the additional tax revenue generated by Head Start, we use data from the NLSY. We first estimate the selection effect on wage income due to Head Start beneficiaries entering the labor force using two approaches: (1) lower truncation and (2) assuming that new entrants are representative of the overall income distribution of Head Start participants. The first approach implies that the actual wage gains were 18.2% for Head Start participants. The second is more conservative, implying that the wage gains were only 5.2%. We then apply these gains to the earnings of Head Start participants and use the NBER's Taxsim program to estimate the average payroll tax rate and individual income tax rate among NLSY respondents who report attending Head Start as children. This calculation depends in large part on the earnings of labor-market entrants. Our preferred estimate suggests that the average Head Start participant pays an extra $576 annually in federal taxes as a result of the program. A less conservative estimate based on lower truncation suggests an increase in tax revenue per student of $2,341. Combining the flow of savings from reduced public assistance with our estimate of the revenue from additional labor income, the internal rate of return to the federal government's expenditure per student ranges from 5.4 to 9.1%. This calculation assumes the fiscal externality due to wage effects among individuals whom Head Start induced to work are zero or even negative.

Table A15. The Effect of Human Capital and Self-Sufficiency on Adult Earnings Potential

	(1)	(2)	(3)	(4)	(5)	(6)
	A. Ages 14 to 22			*B. Deming's sample: Ages 14 to 19*		
	Total	Male	Female	Total	Male	Female
A. Internal rate of return (IRR)						
HC and ESS only	13.65%	12.86%	13.32%	13.78%	13.16%	13.43%
B. Regression estimates						
High School graduation	0.138***	0.0913*	0.217***	0.164***	0.116**	0.233***
	(0.0354)	(0.0479)	(0.0530)	(0.0414)	(0.0549)	(0.0648)
College completion	0.152***	0.215***	0.121***	0.158***	0.211***	0.135***
	(0.0317)	(0.0487)	(0.0410)	(0.0389)	(0.0594)	(0.0508)
Prof or doc degree	0.0166	0.0358	0.00868	-0.00138	0.0228	0.00267
	(0.0359)	(0.0588)	(0.0444)	(0.0429)	(0.0698)	(0.0537)
Professional job	0.249***	0.222***	0.256***	0.270***	0.225***	0.288***
	(0.0158)	(0.0246)	(0.0202)	(0.0184)	(0.0281)	(0.0242)
Years of schooling	0.0304***	0.0270**	0.0268***	0.0313***	0.0340***	0.0219*
	(0.00719)	(0.0111)	(0.00929)	(0.00876)	(0.0132)	(0.0116)
Weeks worked last year	0.0352***	0.0335***	0.0359***	0.0355***	0.0342***	0.0359***
	(0.000763)	(0.00116)	(0.00101)	(0.000918)	(0.00134)	(0.00124)
Usual weekly hours	0.0199***	0.0156***	0.0237***	0.0198***	0.0153***	0.0240***
	(0.000659)	(0.000844)	(0.00103)	(0.000765)	(0.000965)	(0.00122)
AFQT score	0.190***	0.183***	0.202***	0.154***	0.153***	0.156***
	(0.0106)	(0.0150)	(0.0151)	(0.0124)	(0.0170)	(0.0181)
AFQT squared	-0.0476***	-0.0358***	-0.0655***	-0.0372***	-0.0357***	-0.0427***
	(0.00733)	(0.0101)	(0.0105)	(0.00833)	(0.0113)	(0.0121)
Idle				0.0436***	0.0445***	0.0482***
				(0.00923)	(0.0144)	(0.0120)
Crime				0.0146	0.0179*	-0.00151
				(0.00928)	(0.0105)	(0.0196)
Teen pregnancy				-0.00256	-0.00500	0.00424
				(0.00831)	(0.0146)	(0.0101)
Health				0.0105	0.0110	0.0112
				(0.00856)	(0.0136)	(0.0109)
Constant	7.843***	7.872***	6.877***	7.818***	7.772***	6.912***
	(0.0947)	(0.139)	(0.120)	(0.112)	(0.161)	(0.147)
Observations	36,536	18,252	18,284	25,508	13,004	12,504
# Individuals	7,327	3,653	3,674	5,053	2,578	2,475
R-squared	0.409	0.356	0.400	0.414	0.366	0.405

Notes: The dependent variable is log wage income. Control variables not reported in the table include race, gender, and birth and survey year fixed effects. High school completion, college completion, and professional or doctoral degree indicate completed years of education is greater or equal to 12, 16, or 18, respectively. Standard errors (in parentheses) are clustered at the individual level to account for longitudinal dependence in the data.

12. Additional References Not in Main Text

Cunningham, J. (2013). *More Lawyers, More Crime? The Impact of the OEO Legal Services Program on Crime*. Portland State University Working Paper. Retrieved from http://paa2015.princeton.edu/uploads/150684

Haines, M. R., Inter-university Consortium for, P., & Social, R. (2010). *Historical, Demographic, Economic, and Social Data: The United States, 1790-2002*. Retrieved from: http://doi.org/10.3886/ICPSR02896.v3

Office of Child Development. (1973). *Directory of Full Year Head Start Programs, February 1973*. Washington, D.C.: U.S. Department of Health, Education and Welfare

Project Head Start. (1971). *Directory of Full Year Head Start Programs*. Washington, D.C.: U.S. Department of Health, Education, and Welfare

Project Head Start. (1978). *Directory of Full Year Head Start Programs, Summer Head Start Programs, Head Start Parent and Child Centers*. Department of Health, Education, and Welfare, Office of Human Development Services, Administration for Children, Youth, and Families