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

August 27, 2018

[click here for most updated draft]

Abstract

This paper evaluates the long-run effects of Head Start on human capital and economic self-sufficiency using large-scale, restricted 2000-2013 Census-ACS data linked to date and place of birth in the Numident. Our research design exploits the county-level rollout of Head Start between 1965 and 1980 together with variation in eligibility captured by state-level school-entry age cutoffs. We find that the human capital index of children induced to participate in Head Start increased by 10 percent of a standard deviation. Participating children achieved 0.29 more years of schooling, reflecting 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. Head Start also raised the index of economic self-sufficiency by 4 percent of a standard deviation, decreasing adult poverty by 12 percent and public assistance receipt by 29 percent. Our estimates imply substantial, 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 (April 2017)

In 1965, the U.S. began a new experiment in the provision of public pre-school 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 "prep school for poor kids," aptly named "Head Start" (Levitan 1969), which aimed to help millions of children escape poverty.

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.[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 (Westinghouse Learning Corporation 1969, Currie 2001, Duncan and Magnuson 2013, Phillips et al. 2017). Convincing evidence regarding the Head Start's long-term effects in reducing poverty has remained even 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. Whether Head Start achieved its broad goal of increasing life opportunities for children remains an open question.

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

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

(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 2012, Bailey and Danziger 2013, Bailey and Duquette 2014, Bailey and Goodman-Bacon 2015). 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* amounts. 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-*in*eligible (children 6 and older). Similar to the assumptions of a regression kink design (Card et al. 2015), our key identifying assumption is 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.

Our results suggest that Head Start increased the human capital and economic self-sufficiency of disadvantaged children. An aggregated index of human capital rose by 10 percent of a standard deviation among participants relative to the comparison group. Children participating in Head Start 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. Overall, these results suggest that Head Start achieved its goal of increasing children's economic opportunities and reducing poverty. A final analysis quantifies the large, long-term internal rates of return to dollars spent on Head Start in the 1960s and 1970s.

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

In the 1960s, the (then) revolutionary idea that preschool could improve children's cognitive development 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 and Mission

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. 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, Head Start had begun in another roughly 900 counties, and federal expenditures on the program reached $326 million, or $1.9 billion in 2012 dollars (OEO 1970). This early expansion ensured that by 1970, Head Start existed in roughly half of U.S. counties nationwide and covered 83 percent of U.S. children under age six. Because much of this history is described elsewhere, our quantitative description of the program's roll-out appears in our Appendix.

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

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

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).

B. A Stylized Example of the Expected Effects by a Cohort's Age at Head Start's Launch

Head Start's phased implementation implies a pattern in the expected relationship between adult outcomes and a child's age at the time of the program's launch. Figure 1 plots the expected patterns. Note, this is only intended as a stylized example and will not be assumed as part of our later analysis.

First, if we assume there is no effect of Head Start on children who were over age 5 when it 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 in Figure 1.

---

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

Second, if Head Start has a 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.[4] Program quality also increased over time with better hiring and training of teachers, curriculum development, and the implementation of auxiliary services (e.g., health).[5] 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 1 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. 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—only that the child's potential access to the program is much higher. Figure 1 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

[4] 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, OEO 1966, OEO 1967, OEO 1968, OEO 1970). Enrollment in summer Head Start was much higher, but we expect the summer program to 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.

[5] 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.

line). In Figure 1's stylized example, children ages -1 or older 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 relaxing two assumptions implies a slightly different 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 et al. 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.

Neither economic theory nor the history of Head Start provides more guidance about the expected effects beyond these basic patterns. Our approach to estimating these relationships, therefore, is to use an event-study research design, which allows the data to characterize these relationships rather than assuming them *ex ante*.

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

Previous evaluations of Head Start's long-term effects provide suggestive evidence of the program's long-term effects on human capital and economic self-sufficiency.[6] One pioneering approach was the use of family fixed effects with longitudinal data. Building on work by Currie and Thomas (1995), Garces et al. (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

[6] See reviews of studies of Head Start's short-term effects (Currie 2001, Cascio and Schanzenbach 2013, Duncan and Magnuson 2013, Gibbs et al. 2014).

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 provide unreliable estimates of Head Start's effects (Grosz et al. 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.[7] 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

[7] 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).

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 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, which limits sub-group analyses of the lower-income children heavily represented in the program.[8] 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 between 2000 and 2013 (ages 25 to 54).[9] We collapse these data to means by birth year, survey year, and county of birth and weight our regressions using the number of observations in each cell. To minimize disclosure concerns at the Census Bureau, we use only observations with non-allocated and non-missing values for all outcomes.

[8] 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).
[9] We find no evidence that Head Start affected survival to 2000 (see Appendix).

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 Head Start spending. This refinement (1) addresses the potential endogeneity of Head Start funding decisions to the program's *performance* and (2) sidesteps issues of measurement error in the National Archives grant amounts (Barr and Gibbs 2017).[10] Second, we examine changes in outcomes for children who were *age-eligible* for Head Start when it launched (ages 3-5) relative to those who were age-ineligible (ages 6+), 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 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 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 information on our sample and support for this research design.)

We implement these comparisons within the following event-study framework,

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

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, $\boldsymbol{Z}_c$ interacted with a linear trend in year of birth, b (Hoynes et al. 2011, Bailey 2012, Bailey and Goodman-Bacon 2015).[12]

---

[10] Thompson (2017) also tries this strategy but notes that his estimates in the NLSY are statistically insignificant.
[11] Appendix Tables 12-15 documents how the estimated effects of Head Start change in our Census/ACS-Numident dataset using (1) alternative measures of access to Head Start from the literature 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.

HeadStart is a binary variable equal to 1 if a child was born in a county that received a Head Start grant before 1980. $\boldsymbol{Age}$ 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 (1965−1980) to 30 (1980−1950) 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 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 et al. 2004).[14]

Based on Figure 1's predictions, we also summarize the event-study estimates using a three-part spline with knots at age 6 and -1, which we implement by replacing the $\boldsymbol{Age}_{bs(c)}$ in equation (1) with components of the spline in age, $a = T_c^* - b$:

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

where $\boldsymbol{D}_{cb}'$ is a vector of dummy variables, $1(-10 \le a \le -1)$, $1(-1 \le a \le 6)$, $1(6 \le a \le 15)$, $1(a < -10)$, 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 event-study, it has the added benefits of permitting two tests of our research design: a parsimonious method to test for a pre-trend (captured in the slope of the segment for $1(6 \le a \le 15)$) and a formal trend-break test between components $1(6 \le a \le 15)$ and $1(-1 \le a \le 6)$—a test in the spirit of a regression kink design (for which we do not have sufficient data).

Our outcomes of interest are summary measures of human capital and economic self-sufficiency, which permit tests of co-movements of related adult outcomes and limit the number of statistical tests

[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, 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 et al. 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.

(Kling et al. 2007).[15] 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.

These ITT effects average over all children in a county, regardless of whether they participated in Head Start. To compare our estimates to the literature, we generate average treatment-effects-on-treated children (ATET) by dividing by 0.149, our best estimate of the effect of Head Start launch on program enrollment using both administrative and 1970 Census data (see Appendix for details.) We construct confidence intervals 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 and Tibshirani 1993). Because we cannot resample from data used in other Head Start papers, we also use this method to recalculate confidence intervals for other estimates in the literature for comparison.

IV. HEAD START'S EFFECTS ON HUMAN CAPITAL

Figure 2 plots the event-study estimates for all outcomes in the human capital index for the set of compositionally balanced county-birth-cohorts,[16] or individuals ages 15 to as young as -1 when Head Start launched. Similar to Figure 1, the x-axis is the child's age at the time Head Start launched, and the solid line plots the event-study estimates. Consistent with the stylized example in Figure 1, the human capital

[15] Following Kling et al. (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.

[16] See Appendix Table A2 for a table describing program roll-out by cohort and age.

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.

Table 1 complements this visual impression, summarizing both the event-study and spline estimates at -1 as well as formal tests for a pre-trend (effects at age 6 and higher) and trend break at age 6. 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 6). Across outcomes, column 3's ITT-spline estimates are identical to those in column 2 to the hundredth. Supporting the impression in Figure 2, column 4 shows no evidence of a pre-trend in human capital outcomes, failing to reject zero for each outcome. Finally, column 5 shows that an F-test rejects the null hypothesis of no trend-break in the adult human capital index at age 6 at the 1-percent level.

Some of the most commonly studied outcomes in the preschool literature relate to years of education, including both high school graduation and college enrollment. Table 1 shows 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 3A shows that it is roughly half the size of Garces et al. (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.

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 et al. (2002) and one quarter the size of Ludwig and Miller (2007) (Figure 3B). (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 trend-break in Figure 2C, 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).[17] 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. 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 does not appear driven by a pre-trend.[18]

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

[17] 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.

[18] 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

children and mostly black, and Perry's participants had low IQs.[19] In contrast, Head Start was not exclusively for poor, African-American, or low-IQ children.[20] 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 1 stratifies 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.[21]

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

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

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

[20] The OEO encouraged interactions between poor and less disadvantaged children, allowing 15 percent (later 10 percent) of participants to come from non-poor families. Roughly two-thirds of children in the 1969-1970 full-year programs had mothers with less than a high school degree. Mothers of about 7 percent of children had attended or graduated from college.

[21] Our appendix reports estimates of Head Start's effects on human capital by race and sex. Effects for nonwhites are generally imprecise, because this group represents only 15 percent of the overall sample. The overall patterns suggest that Head Start's effects are largest among white men (13 percent of a std.) and smaller effects among white and non-white women (5-6 percent of a std., respectively). Effects for non-white men were small or negative and imprecise.

self-sufficiency. Table 2 shows that an index of economic self-sufficiency aggregated over both sexes increased by 4 percent of a standard deviation. 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. However, Table 2 shows little effect of Head Start on labor-force participation or wage income. This null result may reflect the fact that men's and women's work effort (and potentially selection) changed in offsetting ways. 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).

Panels A and B of Table 2 are 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. 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.[22] 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.[23]

The pattern is different for women. The self-sufficiency index increased by 4 percent of a standard

---

[22] 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).

[23] We find no evidence of decreases in incarceration among men (see Appendix).

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).[24]

VI. 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 (Cunha and Heckman 2007, Heckman et al. 2010, Almond and Currie 2011, Duncan and Magnuson 2013). 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.

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 earnings *potential* (Neal and Johnson 1996, Deming 2009). (Using potential earnings circumvents problems with selection.) The NLSY79, which allows for controls for ability, 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. 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*; see Appendix).

---

[24] Our appendix reports estimates of Head Start's effects on self-sufficiency by race and sex. Effects for nonwhites are generally statistically insignificant. The overall patterns suggest that Head Start's effects on self-sufficiency are largest among white men and women (3 and 4 percent of a std., respectively). See also footnote 21.

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 (Bound et al. 2001, Meyer et al. 2015), 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.

VII. REFERENCES

Almond, D. and J. Currie (2011). "Human Capital Development before Age Five." Handbook of labor economics 4: 1315-1486.

Almond, D., H. W. Hoynes and D. W. 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, M. A. (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, M. (1987). "Computing Robust Standard Errors for within-Groups Estimators." Oxford Bulletin of Economics and Statistics 49(4): 431-434.

Bailey, M. 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, M. J. and S. Danziger (2013). The Legacies of the War on Poverty. In Legacies of the War on Poverty, edited by Bailey, M. J. and S. Danziger. New York: Russell Sage Foundation.

Bailey, M. J. and N. 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, M. J. and A. 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 L. N. Masse (2007). "Comparative Benefit-Cost Analysis of the Abecedarian Program and Its Policy Implications." Economics of Education Review 26(1): 113-125.

Barr, A. and C. R. Gibbs (2017). "Breaking the Cycle? Intergenerational Effects of an Anti-Poverty Program in Early Childhood." Retrieved Available at https://curry.virginia.edu/sites/default/files/files/EdPolicyWorks_files/61_Anti_Proverty_Effects_ECE.pdf.

Bertrand, M., E. Duflo and S. Mullainathan (2004). "How Much Should We Trust Differences-in-Differences Estimates?" Quarterly Journal of Economics 119(1): 249-275.

Bound, J., C. Brown and N. Mathiowetz, Eds. (2001). Measurement Error in Survey Data. Handbook of Econometrics. Amsterdam: Elsevier.

Bound, J. and G. 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, S. J. and E. P. Edwards (1972). History and Theory of Early Childhood Education. Belmont, CA: Wadsworth.

Cameron, A. C., J. B. Gelbach and D. L. Miller (2011). "Robust Inference with Multiway Clustering." Journal of Business & Economic Statistics 29(2): 238-249.

Card, D., D. S. Lee, Z. Pei and A. Weber (2015). "Inference on Causal Effects in a Generalized Regression Kink Design." Econometrica 83(6): 2453-2483.

Carneiro, P. and R. 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.

Cascio, E. U. (2009). "Maternal Labor Supply and the Introduction of Kindergartens into American Public Schools." Journal of Human Resources 44(1): 140-170.

Cascio, E. U. and D. W. Schanzenbach (2013). "The Impacts of Expanding Access to High-Quality Preschool Education." Brookings Papers on Economic Activity: 127-178.

Cunha, F. and J. Heckman (2007). "The Technology of Skill Formation." American Economic Review 97: 31-47.

Currie, J. (2001). "Early Childhood Education Programs." Journal of Economic Perspectives 15(2): 213-238.

Currie, J. and D. Thomas (1995). "Does Head Start Make a Difference? ." American Economic Review 85(3): 341-364.

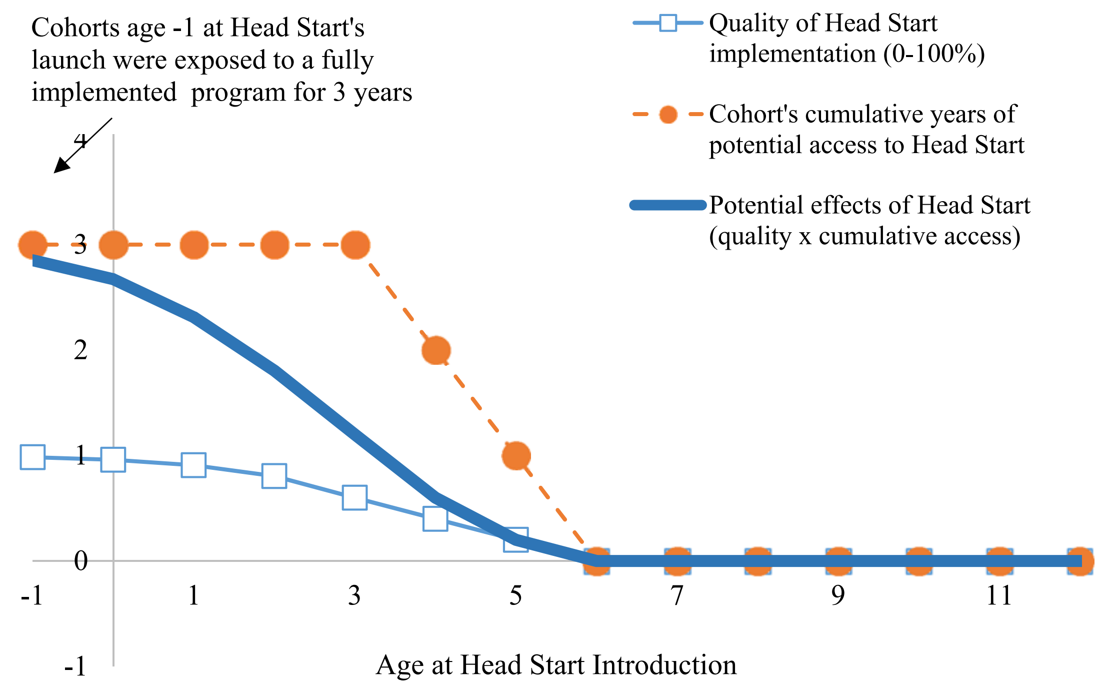
Deming, D. (2009). "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head

Start " American Economic Journal: Applied Economics 1(3): 111-134.
Duncan, G. J. and K. Magnuson (2013). "Investing in Preschool Programs." Journal of Economic Perspectives 27(2): 109-132.
Efron, B. and R. J. Tibshirani (1993). An Introduction to the Bootstrap. New York: Chapman & Hall.
Fosburg, L. B., N. N. Goodrich, M. K. Fox, P. Granahan, J. Smith, J. H. Himes and M. 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, edited by. Cambridge, MA: Abt Associates.
Frisvold, D. E. (2015). "Nutrition and Cognitive Achievement: An Evaluation of the School Breakfast Program." Journal of Public Economics 124: 91-104.
Garces, E., D. Thomas and J. Currie (2002). "Longer Term Effects of Head Start." American Economic Review 92(4): 999-1012.
Gibbs, C., J. Ludwig and D. L. Miller (2014). Does Head Start Do Any Lasting Good? In Legacies of the War on Poverty, edited by Bailey, M. J. and S. Danziger. New York, NY: Russell Sage Foundation. 39-65.
Goodman-Bacon, A. J. (2018). "Public Insurance and Mortality: Evidence from Medicaid Implementation." Journal of Political Economy 126(1): 216-262.
Griliches, Z. (1979). "Sibling Models and Data in Economics: Beginnings of a Survey " Journal of Political Economy 87(5, Part 2): S37-S64.
Grosz, M. Z., D. L. Miller and N. a. Shenhav (2017). "All in the Family: Assessing the External Validity of Family Fixed Effects Estimates and the Long-Term Impact of Head Start." Retrieved July 31, 2017. Available at http://conference.nber.org/confer//2017/SI2017/CH/Miller_Shenhav_Grosz.pdf.
Hechinger, F. M., Ed. (1966). Pre-School Education Today: New Approaches to Teaching Three-, Four-, and Five-Year-Olds. New York: Doubleday.
Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev and A. Yavitz (2010). "The Rate of the Return to the High Scope Perry Preschool Program " Journal of Public Economics 94(1): 114-128.
Hoynes, H. W., M. E. Page and A. H. Stevens (2011). "Can Targeted Transfers Improve Birth Outcomes? Evidence from the Introduction of the Wic Program." Journal of Public Economics 95(7): 813-827.
Johnson, R. and C. K. Jackson (2017). "Reducing Inequality through Dynamic Complementarity: Evidence from Head Start and Public School Spending." NBER Working Paper No. 23489 June.
Kling, J. R., J. B. Liebman and L. F. Katz (2007). "Experimental Analysis of Neighborhood Effects." Econometrica 75(1): 83-119.
Levine, R. A. (1970). The Poor Ye Need Not Have with You: Lessons from the War on Poverty. Cambridge: MIT Press.
Levitan, S. A. (1969). The Great society's poor Law: A New approach to poverty. Baltimore: Johns Hopkins Press.
Ludwig, J. and D. 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, B. D., W. K. C. Mok and J. X. Sullivan (2015). "Household Surveys in Crisis." Journal of Economic Perspectives 29(4).
Neal, D. and W. Johnson (1996). "The Role of Premarket Factors in Black-White Wage Differences." Journal of Political Economy 104(5): 869-895.
North, A. F. (1979). Health Services in Head Start. In Project Head Start: A Legacy of the War on Poverty, edited by Zigler, E. and J. 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. 397, H. D. N.
OEO (1968). 4th Annual Report: As the Seed Is Sown.
OEO (1970). Annual Report Fiscal Years 1969-70.
Phillips, D. A., M. W. Lipsey, K. A. Dodge, R. Haskins, D. Bassok, M. R. Burchinal, G. J. Duncan, M. Dynarski, K. A. Magnuson and W. Christina (2017). "The Current State of Scientific Knowledge on Pre-Kindergarten Effects." Retrieved April 18, 2017. Available at https://www.brookings.edu/wp-

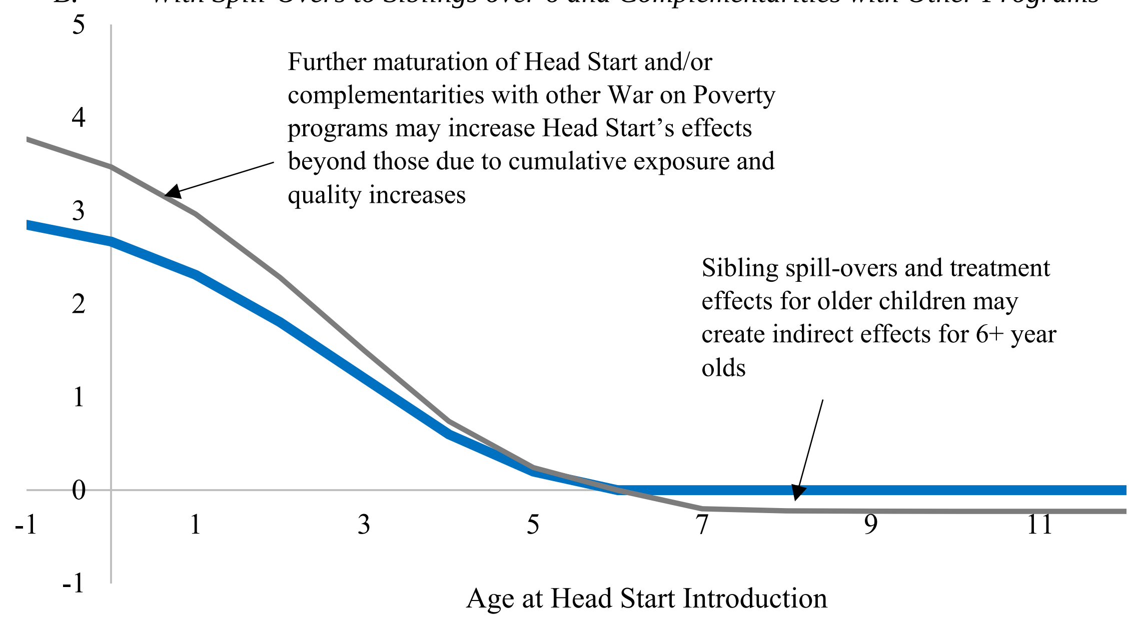
content/uploads/2017/04/duke_prekstudy_final_4-4-17_hires.pdf.
Thompson, O. (2017). "Head Start's Long Run Impact: Evidence Form the Program's Introduction." Journal of Human Resources 53(4).
United States Senate Committee on Labor and Public Welfare (1964). Economic Opportunity Act of 1964. Congress, t. Report no. 1218.
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).

Figure 1. 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

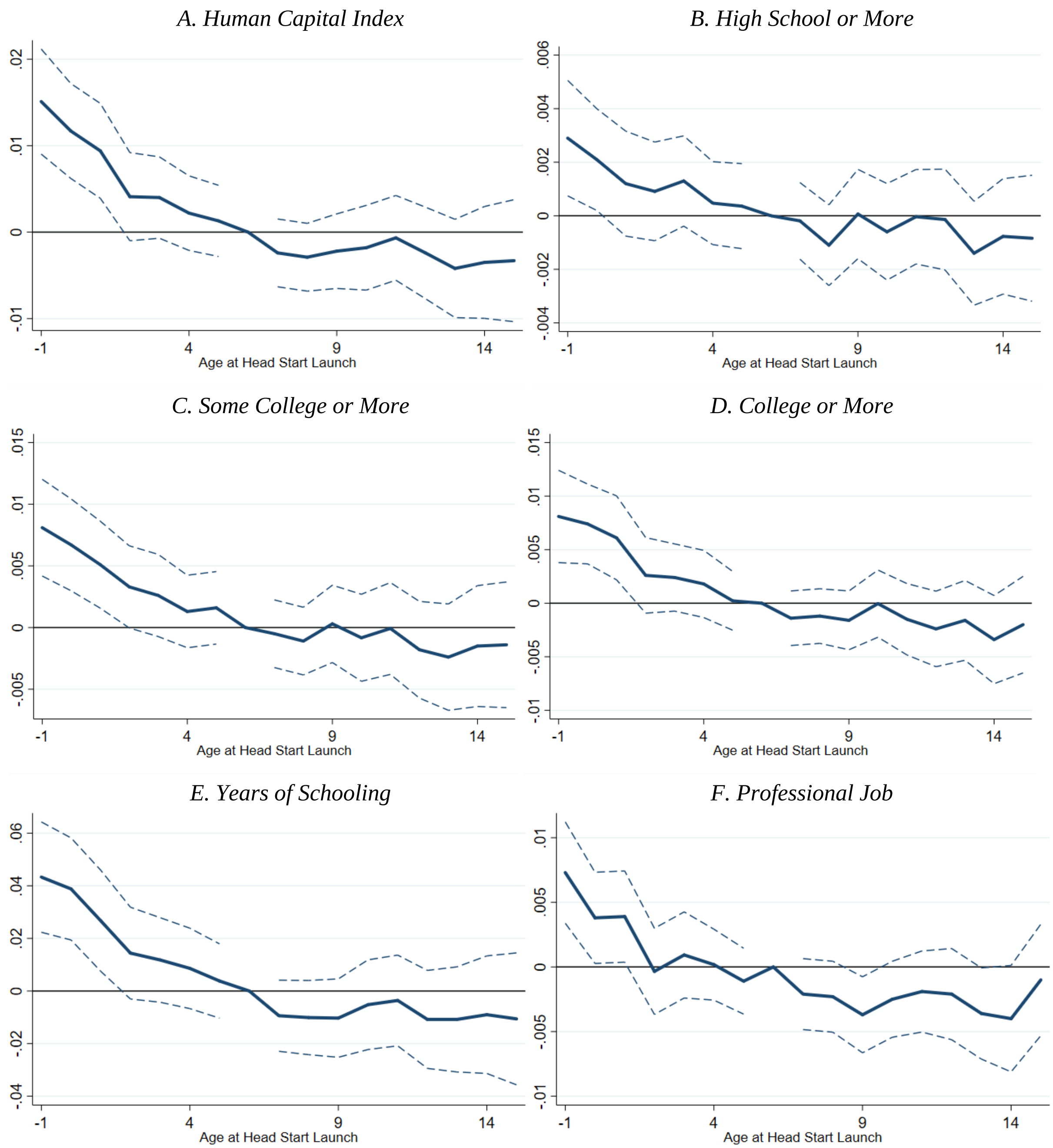


B. With Spill-Overs to Siblings over 6 and Complementarities with Other Programs



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.

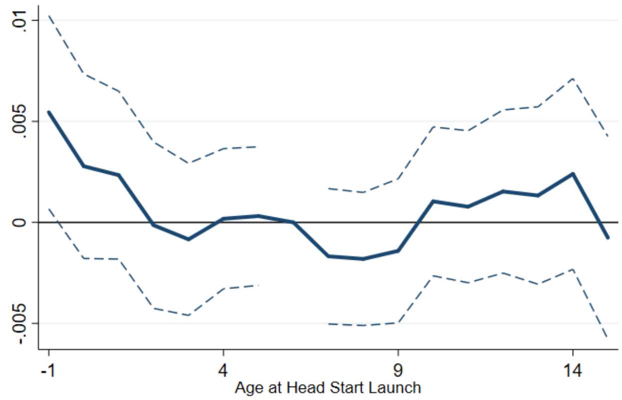
Figure 2. The Effect of Head Start on Adult Human Capital



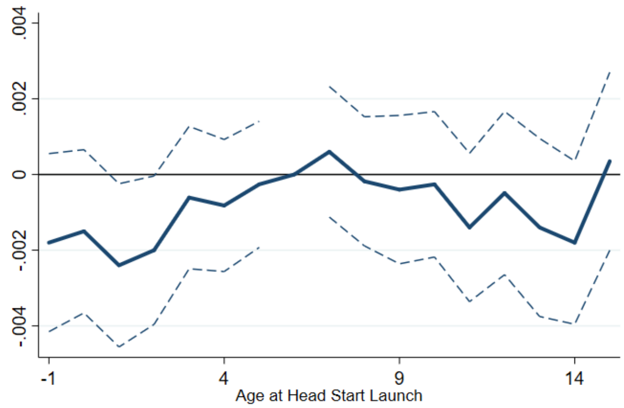
Notes: The figures plot event-study estimates of φ for different outcomes using the specification in equation (1). Standard errors clustered at the county level. Dashed lines show 95-percent, point-wise confidence intervals for each estimate.

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

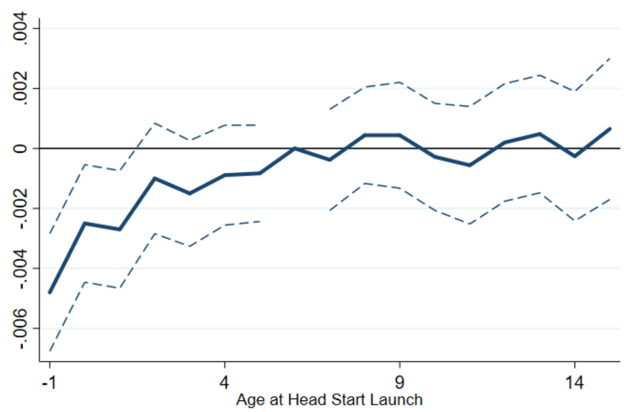
A. Economic Self-Sufficiency Index



B. In Poverty



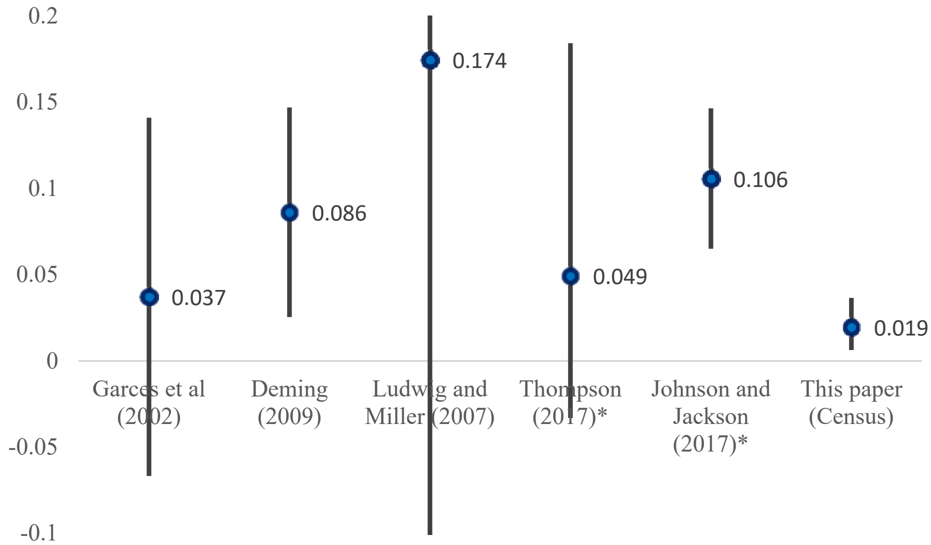
C. Received Public Assistance Income



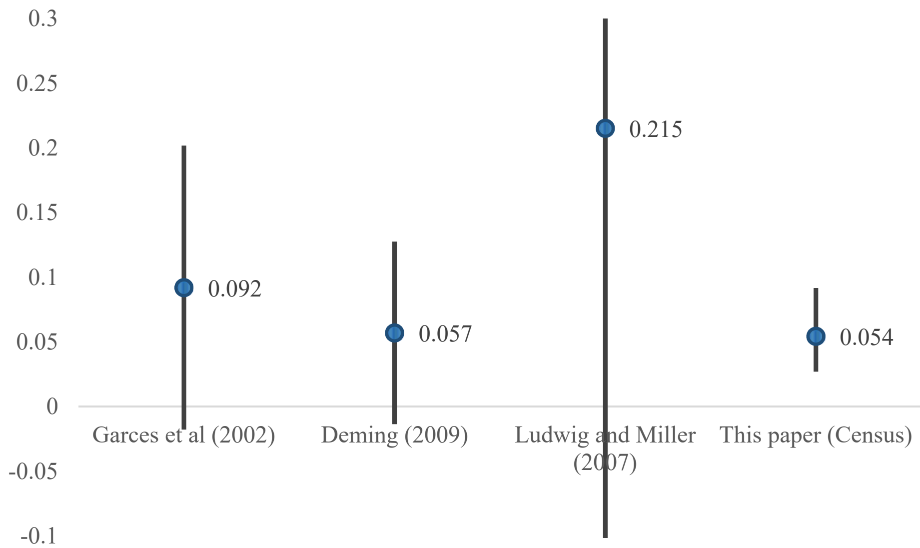
Notes: See figure 2 notes. Note that In Poverty and Received Public Assistance are reverse coded when included in the Economic Self-Sufficiency Index.

Figure 4. The Magnitude of Head Start's Effects on Education across Studies

A. Effects of Head Start on High School Graduation



B. The Effects of Head Start on College Enrollment



Notes: Circles indicate the reported or derived ATET from different studies. For sibling fixed effect studies, the ATET is directly reported in the papers. For other studies, we scale the reported ITT by the reported first-stage estimate and constructed 95-percent confidence intervals using the standard error of the first stage as described in the text. 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.

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

Dependent Variables	(1) Control Mean (std.)	(2) Event Study at -1 (s.e.)	(3) Spline at -1 (s.e.)	(4) Test for pre-trend (s.e.)	(5) F-test of trend break at age 6 [p-value]	(6) ATET [95% CI]	(7) ATET % increase
A. Men and Women							
Human capital index	0.014	0.015	0.015	-0.0002	14.02	0.1010	
	(0.198)	(0.003)	(0.003)	(0.0003)	[0.00]	[.057,.163]	
Completed high school/GED	0.92	0.003	0.003	-0.0001	2.18	0.0189	2.1%
	(0.078)	(0.001)	(0.001)	(0.0001)	[0.14]	[.005,.038]	
Attended some college	0.62	0.008	0.008	-0.0002	4.84	0.0540	8.7%
	(0.140)	(0.002)	(0.002)	(0.0003)	[0.03]	[.027,.092]	
Completed 4 year college	0.29	0.008	0.009	-0.0002	11.66	0.0540	18.6%
	(0.122)	(0.002)	(0.002)	(0.0002)	[0.00]	[.025,.094]	
Prof. or doc. degree	0.028	0.002	0.002	0.0000	7.36	0.0140	50.0%
	(0.037)	(0.001)	(0.000)	(0.0001)	[0.01]	[.006,.024]	
Years of schooling	13.57	0.043	0.049	-0.0005	13.06	0.2910	2.1%
	(0.695)	(0.011)	(0.010)	(0.0012)	[0.00]	[.144,.49]	
Has a professional job	0.35	0.007	0.007	-0.0001	4.97	0.0489	14.0%
	(0.121)	(0.002)	(0.002)	(0.0002)	[0.03]	[.022,.085]	
B. Men							
Human capital index	0.014	0.021	0.020	-0.0005	11.79	0.1360	
	(0.247)	(0.004)	(0.003)	(0.0004)	[0.00]	[.081,.215]	
Completed high school/GED	0.91	0.004	0.003	-0.0001	1.92	0.0250	2.7%
	(0.104)	(0.002)	(0.001)	(0.0002)	[0.17]	[.005,.049]	
Attended some college	0.59	0.011	0.010	0.0000	5.67	0.0740	12.5%
	(0.176)	(0.003)	(0.002)	(0.0003)	[0.02]	[.036,.125]	
Completed 4 year college	0.29	0.012	0.011	-0.0003	9.99	0.0769	26.5%
	(0.150)	(0.003)	(0.002)	(0.0003)	[0.00]	[.041,.126]	
Prof. or doc. degree	0.034	0.003	0.003	0.0000	4.73	0.0199	58.5%
	(0.052)	(0.001)	(0.001)	(0.0001)	[0.03]	[.008,.035]	
Years of schooling	13.5	0.062	0.063	-0.0012	12.15	0.4129	3.1%
	(0.878)	(0.013)	(0.011)	(0.0015)	[0.00]	[.228,.668]	
Has a professional job	0.34	0.010	0.009	-0.0006	2.54	0.0649	19.1%
	(0.153)	(0.003)	(0.002)	(0.0003)	[0.11]	[.031,.11]	
C. Women							
Human capital index	0.015	0.010	0.012	0.0000	6.50	0.0659	
	(0.228)	(0.004)	(0.003)	(0.0004)	[0.01]	[.014,.132]	
Completed high school/GED	0.93	0.002	0.002	-0.0001	1.42	0.0140	1.5%
	(0.087)	(0.001)	(0.001)	(0.0001)	[0.23]	[-.003,.035]	
Attended some college	0.65	0.005	0.006	-0.0004	1.20	0.0370	5.7%
	(0.163)	(0.002)	(0.002)	(0.0003)	[0.27]	[.005,.077]	
Completed 4 year college	0.29	0.005	0.008	-0.0002	4.63	0.0320	11.0%
	(0.147)	(0.003)	(0.002)	(0.0003)	[0.03]	[-.004,.076]	
Prof. or doc. degree	0.022	0.001	0.001	0.0001	3.78	0.0080	36.4%
	(0.042)	(0.001)	(0.001)	(0.0001)	[0.05]	[-.001,.018]	
Years of schooling	13.6	0.024	0.038	-0.0003	4.91	0.1659	1.2%
	(0.795)	(0.013)	(0.012)	(0.0014)	[0.03]	[-.013,.379]	
Has a professional job	0.37	0.005	0.005	0.0004	4.35	0.0350	9.5%
	(0.150)	(0.003)	(0.002)	(0.0003)	[0.04]	[0,.078]	

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 1). 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.

Table 2. The Effect of Head Start on Adult Economic Self-Sufficiency

Dependent Variables	(1) Control Mean (std.)	(2) Event Study at -1 (s.e.)	(3) Spline at -1 (s.e.)	(4) Test for pre-trend (s.e.)	(5) F-test of trend break at age 6 [p-value]	(6) ATET [95% CI]	(7) ATET % increase
A. Men and Women							
Self-sufficiency index	0.024	0.0055	0.0048	0.0003	4.24	0.037	
	(0.154)	(0.0024)	(0.0022)	(0.0002)	[0.04]	[.005,.076]	
Worked last year	0.86	0.00082	0.00091	0.0000	0.22	0.0060	0.7%
	(0.098)	(0.0014)	(0.0013)	(0.0002)	[0.64]	[-.013,.026]	
Weeks worked last year	41.1	0.047	0.055	0.0057	0.61	0.32	0.8%
	(5.281)	(0.075)	(0.0721)	(0.0092)	[0.43]	[-.68,1.4]	
Usual hours works per week	35.72	0.0334	0.056	0.0106	1.09	0.22	0.6%
	(4.907)	(0.080)	(0.0791)	(0.0085)	[0.30]	[-.84,1.4]	
Log labor income	10.60	0.0055	0.0061	0.00018	2.8	0.037	
	(0.268)	(0.0035)	(0.0031)	(0.00039)	[0.09]	[-.009,.09]	
Log family income/poverty	5.86	0.0059	0.0054	-0.0004	0.28	0.04	
	(0.258)	(0.0043)	(0.0041)	(0.0004)	[0.60]	[-.017,.104]	
In poverty*	0.1	-0.0018	-0.0020	0.0001	3.23	-.012	-12.0%
	(0.084)	(0.0012)	(0.0011)	(0.0001)	[0.07]	[-.03,0.004]	
Rec'd public assistance*	0.11	-0.0048	-0.0037	0.0000	7.13	-.032	-29.1%
	(0.093)	(0.0010)	(0.0008)	(0.0001)	[0.01]	[-.052,-.018]	
B. Men							
Self-sufficiency index	0.031	0.0046	0.00553	-0.0001	4.24	0.0299	
	(0.201)	(0.0031)	(0.0025)	(0.0003)	[0.21]	[-.01,.077]	
Worked last year	0.91	0.0029	0.00301	0.0000	2.33	0.0189	2.1%
	(0.108)	(0.0014)	(0.0011)	(0.0001)	[0.13]	[.001,.041]	
Weeks worked last year	44.79	0.153	0.168	-0.0015	2.46	1.013	2.3%
	(6.031)	(0.0780)	(0.0637)	(0.0078)	[0.12]	[.002,2.221]	
Usual hours works per week	41.31	0.148	0.1722	0.0039	2.94	0.9800	2.4%
	(6.040)	(0.0882)	(0.0798)	(0.0080)	[0.09]	[-.161,2.317]	
Log labor income	10.88	0.0033	0.00238	-0.0008	0.33	0.0219	
	(0.301)	(0.0043)	(0.0036)	(0.0005)	[0.57]	[-.035,.083]	
Log family income/poverty	5.91	0.0035	0.00245	-0.0006	0.07	0.023	
	(0.302)	(0.0048)	(0.0041)	(0.0005)	[0.79]	[-.04,.092]	
In poverty*	0.07	-0.00033	-0.000637	-0.0001	0.01	-.002	-2.9%
	(0.095)	(0.0013)	(0.0011)	(0.0001)	[0.90]	[-.02,.016]	
Rec'd public assistance*	0.11	-0.0046	-0.0046	0.0002	9.93	-.029	-26.4%
	(0.110)	(0.0015)	(0.0011)	(0.0001)	[0.00]	[-.056,-.01]	
C. Women							
Self-sufficiency index	0.019	0.0058	0.00406	0.0005	4.24	0.0399	
	(0.18)	(0.0029)	(0.0027)	(0.0004)	[0.10]	[.001,.088]	
Worked last year	0.8	-0.0018	-0.00168	0.0001	0.15	-.012	-1.5%
	(0.13)	(0.0020)	(0.0018)	(0.0003)	[0.70]	[-.043,.016]	
Weeks worked last year	37.66	-0.097	-0.0931	0.0091	0.03	-.667	-1.8%
	(6.8)	(0.11)	(0.1015)	(0.0146)	[0.87]	[-2.31,.858]	
Usual hours works per week	30.49	-0.11	-0.092	0.0117	0.00	-.745	-2.4%
	(5.81)	(0.097)	(0.095)	(0.0129)	[0.95]	[-2.24,.599]	
Log labor income	10.31	0.0062	0.007	0.00084	3.94	0.0430	
	(0.34)	(0.0051)	(0.0048)	(0.00049)	[0.04]	[-.026,.122]	
Log family income/poverty	5.81	0.0083	0.0084	-0.00034	0.69	0.057	
	(0.32)	(0.0054)	(0.0050)	(0.00047)	[0.41]	[-.015,.143]	
In poverty*	0.12	-0.0028	-0.00273	-0.0003	4.21	-.018	-15.0%
	(0.11)	(0.0017)	(0.0015)	(0.0002)	[0.04]	[-.047,.004]	
Rec'd public assistance*	0.12	-0.0049	-0.0029	0.0002	0.67	-.034	-28.3%
	(0.12)	(0.0014)	(0.0011)	(0.0001)	[0.41]	[-.06,-.014]	

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

[Click here for Online Appendices]