EdWorkingPaper No. 19-141

Breaking the Cycle?Intergenerational Effects of an Anti-Poverty Program in Early Childhood

Andrew C. Barr
Texas A&M University

Chloe Gibbs
University of Notre Dame

Despite substantial evidence that resources and outcomes are transmitted across generations, there has been limited inquiry into the extent to which anti-poverty programs actually disrupt the cycle of bad outcomes. We explore how the effects of the United States' largest early childhood program, Head Start, transfer across generations. We leverage the rollout of this federally funded, means-tested preschool program to estimate the effect of early childhood exposure among mothers on their children's long-term outcomes. We find evidence of intergenerational transmission of effects in the form of increased educational attainment, reduced teen pregnancy, and reduced criminal engagement in the second generation.

VERSION: October 2019

Breaking the Cycle?

Intergenerational Effects of an Anti-Poverty Program in Early Childhood *

Andrew Barr† & Chloe R. Gibbs‡

August 2019

Abstract

Despite substantial evidence that resources and outcomes are transmitted across generations, there has been limited inquiry into the extent to which anti-poverty programs actually disrupt the cycle of bad outcomes. We explore how the effects of the United States' largest early childhood program, Head Start, transfer across generations. We leverage the rollout of this federally funded, means-tested preschool program to estimate the effect of early childhood exposure among mothers on their children's long-term outcomes. We find evidence of intergenerational transmission of effects in the form of increased educational attainment, reduced teen pregnancy, and reduced criminal engagement in the second generation.

Keywords: intergenerational, early childhood, Head Start, long-term

*We are grateful for financial support from the University of California, Davis Center for Poverty Research's Small Grants Competition, made possible in part by grant number AE00102 from the U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Analysis. We thank Maria Rosales-Rueda and Terri Sabol for helpful discussant comments and seminar participants at the University of Chicago Harris School, Stanford University Center for Education Policy Analysis, Institute for Research on Poverty Summer Research Workshop (Wisconsin), UC Davis, and W.E. Upjohn Institute for Employment Research for useful feedback. We also thank Martha Bailey, Marianne Bitler, Jason Lindo, Jens Ludwig, Bhash Mazumder, and Doug Miller for insightful conversations. Sreeraahul Kancherla, David Melnick, and Keanan Gleason provided excellent research assistance. Any remaining errors are our own. The views expressed in publications resulting from supported research do not necessarily reflect the official policies of the U.S. Department of Health and Human Services.

†Department of Economics, Texas A&M University, abarr@tamu.edu

‡Department of Economics, University of Notre Dame, chloe.gibbs@nd.edu.

1 Introduction

> *"I believe this response reflects a realistic and a wholesome awakening in America. It shows that we are recognizing that poverty perpetuates itself. Five and six year old children are inheritors of poverty's curse and not its creators. Unless we act these children will pass it on to the next generation, like a family birthmark."*
>
> – President Lyndon B. Johnson, May 18, 1965

The effects of poverty are pernicious and persistent across generations. Those born to parents in the lowest quintile of the income distribution are twice as likely to end up there as children born to middle-income parents, and the intergenerational correlations in income, educational attainment, female headship, receipt of government assistance, and risky behavior are quite high.[1] Family, school, and neighborhood contexts collectively shape children's trajectories and generate correspondence between their parents' outcomes and their own. These linkages are particularly acute for minorities, potentially contributing to the early emergence and persistence of achievement gaps by race/ethnicity.[2]

Societal investments in education may disrupt the transmission of poverty across generations by increasing educational attainment and labor market attachment and decreasing engagement in risky behavior. Early childhood in particular is a critical developmental period and an opportunity for especially effective intervention. Indeed, multiple studies indicate that interventions in the preschool and early school years can have substantial effects on schooling attainment, labor market success, and other measures of health and well-being into adulthood.[3] And yet, we know almost

[1] A variety of recent estimates suggest intergenerational correlations in income of 0.3 to 0.6 (Black and Devereux 2011, Chetty, Hendren, Kline, Saez and Turner 2014b, Mazumder 2005, Solon 1999), in education levels of 0.4 to 0.5 (Hertz et al. 2008), in female headship of 0.2 (Page 2004), and in welfare use of 0.3 (Page 2004). Similarly, Duncan and coauthors (2005) review and provide a variety of evidence indicating positive intergenerational correlations in early pregnancy, drug use, and other measures of delinquent or risky behavior.

[2] While there is some dispute about the magnitude of these gaps, there is consistent evidence that cognitive test-score gaps by race and ethnicity exist at formal school entry and remain throughout the schooling years (Fryer and Levitt 2004, Fryer and Levitt 2006, Murnane, Willett, Bub and McCartney 2006), and while race/ethnicity achievement gaps have narrowed in recent decades, gaps by socioeconomic status are pronounced (Reardon and Portilla 2016).

[3] Long-term evidence from the Abecedarian Project, Perry Preschool Project, Head Start, and the Project STAR class-size reduction intervention all suggest large positive effects on participants (Campbell et al. 2014, Chetty et al. 2011; Deming 2009, Dynarski et al. 2013, Garces et al. 2002, Heckman et al. 2013, Schweinhart et al. 2005).

nothing about whether these benefits carry over to the next generation. In other words, do these needs-targeted early childhood programs truly break the cycle of poverty?

We answer this question in the context of the Head Start program, providing the first evidence of the intergenerational effects of an early childhood intervention in the United States. The Head Start program, funded and administered through the U.S. Department of Health and Human Services, has been an integral part of the conversation about early childhood intervention for the 50 years of its existence. Easily the largest early childhood education program in the United States, annual Head Start enrollment has grown from 400,000 during the early years of the program to nearly a million participants today. Across multiple datasets and different study designs, a large body of quasi-experimental evidence consistently indicates that participation in the Head Start program yields long-term benefits, particularly for early cohorts of program participants (Carneiro and Ginja 2014, Deming 2009, Garces, Thomas and Currie 2002, Ludwig and Miller 2007).[4]

To explore intergenerational spillovers, we capitalize on differential exposure to Head Start induced by variation in the early rollout of the program. We generate Head Start availability measures using data we extracted and compiled from the National Archives and Records Administration (NARA).[5] We link these measures with the National Longitudinal Survey of Youth–1979 Cohort (NLSY79) and the NLSY79 Children and Young Adults Survey (CNLSY) in order to compare the children of mothers who differ in terms of Head Start availability. We are interested in effects on the second generation's long-term outcomes, including educational attainment, teen pregnancy, and criminal engagement.

We find a significant impact of Head Start availability, between 0.2 to 0.3 standard deviations, on a summary index of long-term outcomes for the second generation. These estimates are robust to the inclusion of a variety of flexible controls for within-county and within-state variation across birth cohorts, including county by birth year trends, as well as direct controls for the time-varying availability of other War on Poverty programs within counties. The legitimacy of the geographic

[4]Estimates of the effect of Head Start on more recent cohorts of participants is less clear. Results of the National Head Start Impact Study, the first large-scale, randomized controlled study of the program showed initial impacts on cognitive and non-cognitive skills for Head Start participants, but these effects faded almost entirely by the first and third grades (Puma et al. 2005, Puma et al. 2010, Puma et al. 2012). However, quasi-experimental evidence indicates meaningful effects on long-term outcomes despite the fade out of test-score impacts (Deming 2009).

[5]See Appendix C for details on construction of the NARA data.

rollout strategy is bolstered by estimates that demonstrate no relationship between Head Start availability and the outcomes of children unlikely to have been eligible for the program.

Our findings are further supported by a second strategy that leverages variation in Head Start exposure generated by the provision of Head Start grant-writing assistance to the poorest 300 counties (Ludwig and Miller 2007). Because funding was contingent on a successful grant proposal, this grant-writing assistance resulted in substantially higher levels of funding and participation in the poorest 300 counties. Using a difference-in-differences approach, we compare the difference in outcomes of children of women who were born too early to participate in Head Start (prior to 1961) with the children of women born later across counties that did and did not receive grant-writing assistance. We find large improvements in second-generation outcomes that are consistent with our primary strategy.

Our findings indicate that societal investments in early childhood education can disrupt the intergenerational transmission of the effects of poverty. Indeed, when comparing children of mothers more or less likely to have grown up in poverty, our estimates suggest that Head Start closes most of the gap in a summary index of long-term outcomes for the second generation.

2 A Path Out of Poverty

There is substantial evidence documenting the path dependency of socioeconomic status. At each developmental stage from infancy through adulthood, children from the highest family income quintile are dramatically more likely to attain educational and economic indicators of lifetime well-being than those from the lowest quintile (Sawhill, Winship and Grannis 2012). While children growing up in disadvantage fall behind at each stage, those who successfully meet such benchmarks from early childhood to adolescence are more likely to attain a middle class existence (Sawhill et al. 2012). These effects carry over to the next generation, resulting in high intergenerational correlations in income, educational attainment, and risky behavior.

Evidence on the lack of intergenerational mobility in the United States has led to considerable interest in understanding why the resources, behaviors, and outcomes of parents are so strongly related to those of their children (Auten, Gee and Turner 2013, Chetty et al. 2014b, Corak 2013, Lee and Solon 2009). Furthermore, there is increased attention on the question of whether interventions that improve these behaviors and outcomes might carry over to the affected individuals' children. The existing evidence on the collective importance of childhood contexts—families, schools, and neighborhoods—in determining long-term outcomes suggests that childhood interventions may be influential (Chetty, Hendren, Kline and Saez 2014a).

Despite the clear importance of this question, the data requirements necessary to answer it convincingly have resulted in limited investigations. While we are unaware of any study to focus on the long-term intergenerational effects of a large-scale anti-poverty program, several have estimated the intergenerational effects of increases in educational attainment in adolescence and beyond. Increases in college access or attainment have resulted in improved birth outcomes and reduced grade retention in the next generation (Currie and Moretti 2003, Maurin and McNally 2008, Page 2009). The evidence at the middle and high school levels is more mixed, with positive intergenerational effects of additional schooling generated by compulsory schooling changes in the U.S. and Great Britain in the 1960s and 70s and no effect in Norway (Black, Devereux and Salvanes 2005, Oreopoulos, Page and Stevens 2006, Chevalier 2007). Recent evidence on the intergenerational spillovers of new school construction in Indonesia suggests that increasing a mother's educational attainment, in particular, produces large improvements in her children's test scores and years of schooling (Akresh, Halim and Kleemans 2018, Mazumder, Rosales-Rueda and Triyana 2019).

Taken together, the evidence for the intergenerational effects of education is promising. However, we know little about the intergenerational effects of interventions in early childhood, despite the large estimated effects of these types of programs on adult outcomes. Interest in the potential intergenerational spillovers of investments made in one generation has resulted in increasing attention from researchers very recently. Rossin-Slater and Wust (2019) explore the intergenerational impact of, and interaction between, a Danish preschool program and a nurse home-visiting program in infancy. They find educational attainment effects in the first generation that persist in the second generation. A recent follow-up of the experimental Perry Preschool Project study also documents substantial, long-term effects for the children of program participants (Heckman and

Karapakula 2019).[6]

Indeed, there is a substantial body of empirical evidence demonstrating that early childhood programs can generate improvements in participants' outcomes over the long-term, and, furthermore, evidence that early skills are important predictors of subsequent academic attainment and labor market success (Chetty et al. 2011, Duncan et al. 2007, Dynarski, Hyman and Schanzenbach 2013). Specifically, evidence from the Abecedarian Project, Perry Preschool Project, Head Start, and Project STAR collectively suggests that interventions in the preschool and elementary years can have substantial effects on schooling attainment, labor market success, and other measures of well-being into adulthood (Chetty et al. 2011, Deming 2009, Schweinhart et al. 2005). Recent evidence documents improvements in health, reductions in behavioral problems, and increases in rates of college-going (Campbell et al. 2014, Carneiro and Ginja 2014, Dynarski et al. 2013).

We contribute to this conversation by providing some of the first evidence—and the first evidence in a U.S. context, to our knowledge—on whether the effects of early childhood programs transfer across generations. The answer to this question has important implications for policies aimed at reducing poverty or socioeconomic gaps in educational attainment, risky behaviors, and labor market success. If these types of policies have large spillover effects on the next generation it suggests that a concerted effort for a single generation of impoverished youth might break the cycle of poverty and reduce the need to provide similar services to future generations.

2.1 The Evidence on Head Start

Recent policy discussions of large-scale, publicly provided preschool interventions often rely on the Head Start literature as the most relevant and informative for designing and scaling up programs. The Head Start program was an early piece of President Lyndon B. Johnson's War on Poverty; operated by the White House's Office of Economic Opportunity (OEO), Head Start commenced as a summer program in 1965 (Vinovskis 2005). It was then quickly expanded to a year-round program. While Head Start today is characterized as an early childhood education program, it was designed as an anti-poverty program with significant health and community development

[6] We return to discussion of these studies in conjunction with our findings later in the paper.

components. The initial emphasis was on a variety of "preschool"-related services and supports, including nutrition, vaccinations and health care, dental services, and social development (Office of Child Development 1968, Office of Child Development 1970, Vinovskis 2005).

Head Start served a decidedly disadvantaged population during the 1960s. The median family income of participants was less than half that of all families in the U.S. and approximately 50 percent of early full-year program participants were black (Office of Child Development 1968). Between nine and 17 percent of families reported having no running water inside the home. Only five percent of mothers reported some postsecondary schooling or more, with approximately 25 percent indicating that they graduated from high school, and 65 to 70 percent of mothers with less than a high school education. Approximately 25 percent lived in female-headed households and between 65 and 70 percent of participating children's mothers were unemployed (Office of Child Development 1968). It is also notable that, among the general population, very few children participated in structured preschool outside the home before entering formal schooling, and that kindergarten availability was not universal at the time that Head Start was being introduced (Cascio 2009).[7]

While there is some debate about the pattern of short-run Head Start effects, prior quasi-experimental studies suggest Head Start has had significant long-term effects for cohorts of children who participated from the late 1960s through the 1980s.[8] Leveraging sibling comparisons and discontinuities in grant-writing assistance and program eligibility, studies have documented increased educational attainment, better health, higher earnings, and less involvement in risky behaviors (Carneiro and Ginja 2014, Deming 2009, Garces et al. 2002, Ludwig and Miller 2007), even in the presence of short-term test-score fadeout (Deming 2009). A follow-up to Deming's study finds that these effects persist later into adulthood, including impacts on participants' later-life parenting practices (Bauer and Schanzenbach 2016).

[7] In the late 1960s, only 10 percent of three- and four-year old children participated in a school setting, according to Current Population Survey school enrollment data, reaching 20 percent in 1970 driven in part by Head Start (Gibbs, Ludwig and Miller 2013, Johnson and Jackson 2018).

[8] While the Head Start Impact Study (HSIS) found initial positive effects on cognitive skill for participants in the mid-2000s, there were no persistent effects at first and third grade follow-ups (Puma et al. 2005, Puma et al. 2010, Puma et al. 2012). Re-analyses of the HSIS data suggest a more nuanced picture (Montialoux 2016). These analyses revealed that there is considerable variation in impact by center (Walters 2015), that effects are most pronounced among children who would otherwise be in parental or relative care (Kline and Walters 2016), and that Hispanic children and children with low skills at program entry experience the greatest benefit (Bitler, Hoynes and Domina 2014).

2.2 Potential Pathways for Intergenerational Transmission

Head Start was conceived of as an anti-poverty program intended to disrupt the cycle of disadvantage passed through generations (Greenberg 1990, Vinovskis 2005). The designers intended for the program to be multi-generational in its service delivery and to prepare young people for successful transitions to formal schooling (Vinovskis 2005). The program's comprehensive view of "school readiness" led to a broad emphasis on a variety of supports for children's healthy development, including medical and dental screenings, mental health, nutrition, and parental engagement, and on connecting children and families with other available services (Office of Child Development 1968, Greenberg 1990, Zigler and Valentine 1979). This bundle of interventions likely affected the long-run outcomes of Head Start-participating children directly, while also indirectly influencing the quality of schooling, peer groups, and subsequent environments with which those children interacted after Head Start (Johnson and Jackson 2018).

When thinking about the intergenerational spillovers of those first-generation improvements, existing literature documents the importance of a mother's human capital and health in influencing her children's prenatal health, birth outcomes, and early childhood health (Almond and Currie 2011, Currie and Moretti 2003, East, Miller, Page and Wherry 2019, Miller and Wherry 2019). There is also a growing body of evidence about the earliest years of life, when children's brains exhibit the greatest developmental plasticity, serving as a critical period in life-cycle skill development (Heckman 2007, Heckman and Mosso 2014, Knudsen, Heckman, Cameron and Shonkoff 2006). While Head Start exposure occurs at age four or five in the first generation, the spillovers potentially affect the second generation prenatally and in the earliest years of childhood. These early life conditions then have cascading effects on children's subsequent development, and may interact in complementary ways with other investments and inputs that parents then make (Cunha and Heckman 2007, Heckman and Masterov 2007). If the intervention affects multiple inputs to human capital production in the second generation, and those inputs are complementary, one could plausibly observe improvements in long-term outcomes for the children of those directly exposed that are larger than for participants themselves.

2.3 Head Start's Beginnings

To build on the existing evidence of Head Start effectiveness and extend our understanding of the intergenerational effects of anti-poverty efforts, we capitalize on plausibly random exposure to the Head Start program over geography and birth cohorts during the program's introduction (Figure 1). The Head Start program was rolled out quickly as a featured and politically popular component of the War on Poverty (Zigler and Valentine 1979). Officially announced at the White House in February 1965, the program was operational that summer rendering strategic funding behavior or coordinated grantee response more challenging and less likely. Historical accounts of OEO programs describe the "great administrative confusion" (Levine 1970) and "wild sort of grant-making operation" (Gillette 1996). Grant funds were distributed directly to local grantees as a means to circumvent governors, state legislatures, and agencies that may have prevented the funds from reaching disadvantaged black children (Gibbs, Ludwig and Miller 2011, Vinovskis 2005). Historical accounts of the development and implementation of Head Start describe early implementation as happening "with explosive speed" (White and Phillips 2001), in OEO Director Sargent Shriver's "characteristic cyclonic manner" (Greenberg 1990) with emphasis on a rapid, large-scale launch (Vinovskis 2005). Government documents and historical records also reveal that the administration did not anticipate the volume of interest it received from local communities, resulting in significant expansion of the rollout, with a concerted emphasis on broad coverage across the country (Office of Child Development 1968, Zigler and Styfco 2010, Vinovskis 2005).[9]

In the early years of the program, approximately half of U.S. counties received Head Start funding, with the majority of counties ever operating Head Start opening their first Head Start program between 1965 and 1970 (Bailey, Sun and Timpe 2018, Johnson and Jackson 2018). As a result of the uncoordinated, local distribution of funding, programs became available in different counties at different times, over a relatively short window. In Appendix Figures A1 and A2, we plot the bivariate relationship between 1960 county characteristics and Head Start adoption through 1970. These plots illustrate the substantial variation in year of adoption among counties with similar baseline characteristics. Consistent with the family-income based targeting of the program

[9] The Johnson administration noted that 49 percent of all applications were from rural areas and that, of the 300 poorest counties in the country, 261 would operate Head Start programs in that early rollout (Johnson 1965).

and early efforts to ensure that the most disadvantaged communities were not excluded from the process, higher poverty counties tended to adopt the program earlier (Greenberg 1990). However, once we condition on the 1960 poverty level (% Family Income < 3K), other county characteristics are largely unrelated to year of adoption (Appendix Table A1).[10,11] Altogether, the baseline county characteristics explain very little of the variation in year of adoption, consistent with the historical accounts of the early rollout period and supporting the validity of our rollout approach.[12]

Our strategy is similar to the novel approach employed in looking at the impact of other War on Poverty and poverty reduction programs, including the Food Stamp Program (Almond, Hoynes and Schanzenbach 2011, Hoynes, Schanzenbach and Almond 2016), the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) (Hoynes, Page and Stevens 2011), and Community Health Centers (Bailey and Goodman-Bacon 2015). Authors have also leveraged geographic variation in program expansions to assess impact of early childhood interventions in Germany and Denmark (Rossin-Slater and Wst 2019, Cornelissen, Dustmann, Raute and Schnberg 2018). Three concurrent papers use the introduction of Head Start over geography and time to explore impact on first-generation outcomes; all three papers document long-term impact for those exposed to the Head Start program (Johnson and Jackson 2018, Thompson 2018, Bailey et al. 2018), particularly when coupled with subsequent schooling investment (Johnson and Jackson 2018). Notably, we focus on a more constrained window of program rollout than work on other poverty reduction programs, and other papers leveraging geographic variation in Head Start. Due to overlap with the NLSY79 birth cohorts, we are capitalizing on the plausibly exogenous introduction of Head Start across counties in the narrow timeframe of 1966 to 1970.

We supplement our rollout strategy with a second difference-in-differences strategy that takes advantage of variation in the provision of grant-writing assistance across counties. Because of concerns about very disadvantaged communities being unable to respond to the call for grant applications, OEO sent government interns from across agencies out to poor counties to distribute

[10] An F-test of the non-poverty controls fails to reject the null that they are equal to zero (p-value = 0.783).

[11] This is consistent with Bailey and Duquette's (2014) account of how War on Poverty programs rolled out in the earliest years of implementation.

[12] Of course, if early adopters are on a different trend than later adopters than these trends would confound our approach. As we discuss below, to address these concerns, we explore the robustness of our results to specifications that interact county baseline covariates with trends in year of birth.

applications, meet with community leaders, mobilize local resources, and assist in preparation of applications (Greenberg 1990). This variation was used previously by Ludwig and Miller (2007), who capitalize on a discontinuity in the provision of this Head Start grant-writing assistance at the 300th poorest county to estimate program impacts on first-generation outcomes.[13] As they note, this approach is challenging when using survey data as many counties are not represented at all, and there are limited individual observations around the discontinuity. Because of these issues, we use the variation in a difference-in-differences framework, taking advantage of the clustered sampling approach of the NLSY79 to examine the difference in outcomes of the children of mothers born too early to participate in Head Start with the children of those born later, coupled with how that difference varies depending on whether a mother's birth county was poor enough to receive grant-writing assistance.

3 Data

To explore Head Start's intergenerational effects we rely on rich, longitudinal survey data that connect mothers and their children. The NLSY79 is a nationally representative sample of adolescents who were 14 to 22 years old when they were first surveyed in 1979. The survey follows 12,686 young men and women, with annual interviews through 1994 and biennial interviews continuing since then. In addition to rich data on labor market participation and transitions, the data provide extensive information on education and training, health, mobility, and family formation. The data facilitate analysis on a representative sample of young men and women living in the United States in 1979. The timing is particularly advantageous for the purposes of this study as individuals born during the early 1960s are differentially exposed to Head Start via its introduction and rollout. Our analytic sample is restricted to NLSY79 respondents (mothers) born between 1960 and 1964.[14]

[13] Ludwig and Miller (2007) summarize historical accounts that Head Start administrators did not want funding only dispersed to cities and the relatively more advantaged and well-resourced communities that both heard about the program and were able to respond. They describe the efforts to send interns out into the field in the several weeks prior to the application deadline.

[14] This is due to measurement and missingness issues with grandmother (mothers of NLSY79 respondents) education levels for pre-1960 cohorts. While maternal education levels for 1960 to 1964 cohorts correspond very closely with other data sources (Current Population Survey) both in levels and trends, there is a significant positive jump in education levels for those born prior to 1960, opposite the trend observed in the CPS. These issues are problematic as we use grandmother education levels, a proxy for Head Start eligibility, to restrict our sample to mothers likely to

Beginning in 1986, a separate, related survey of all children born to NLSY79 female respondents has been collected, the CNLSY. In addition to all the mother's information from the NLSY79, the child survey includes direct information for each child collected from either the mother or child depending on age. The survey gathers data on children's schooling and training, labor market experiences, health, and engagement in risky behaviors. The CNLSY allows us to explore intergenerational effects of the mother's Head Start exposure.

In addition to the NLSY surveys, we use county-by-year data from the Community Action Programs (CAP) and Federal Outlays System (FOS) files available from the National Archives and Records Administration for Head Start availability in fiscal years 1966–1968 (see Appendix C for details). We aggregate data on Head Start grant recipients to the county level, and code a birth cohort and county pair as "exposed" if it received per four-year old Head Start funding above the tenth percentile ($22 per four-year old in the county).[15] We do not otherwise leverage data on appropriated dollar amount because of concerns about the accuracy of the recorded funding amounts in the early years of the Head Start program as well as the endogeneity of funding levels.

We focus on the second generation's long-term outcomes because these outcomes are most important in assessing whether the intergenerational transmission of program effects disrupts the cycle of poverty and because these outcomes capture the myriad ways in which a mother's Head Start access may affect her children's outcomes. These pathways are likely cumulative across childhood. From changes in parenting practices and greater likelihood of enrolling one's child in early childhood programming to heightened expectations and spillovers from a mother's own increased human capital and income, we expect the channels of impact on the second generation to accumulate over the childhood years. There are two positive outcomes: high school completion (including GED receipt) and college going (attending college for any period of time). Because of findings in prior literature, we also consider two negative outcomes with important implications for children and teens' life chances: teen pregnancy and interaction with the criminal justice system (as measured by any arrests, convictions, or probations). These outcomes are important in capturing the

be eligible for Head Start.

[15] In other words, we classify county-birth cohorts with very low funding levels as unexposed. Our results are similar using $0 or the 5th or 15th percentile of funding as the cutoff.

second generation's private returns, but also have implications for measuring the broader societal benefits of the program. To address multiple inference concerns and reduce measurement error, we follow the prior literature in constructing a summary index of our outcome measures (Kling, Liebman, and Katz 2007; Deming 2009). We normalize each outcome to have a mean zero and standard deviation one, adjust outcome signs so that a more positive outcome is better (i.e., we flip the sign on teen parenthood and crime), and take the simple average across these outcomes.

The top panel of Table 1 contains summary statistics for these outcomes. The first column provides these measures for the full sample. Columns (2) and (3) contain similar measures for the samples underlying our geographic rollout strategy. Participation in Head Start is largely restricted to children in poor households. Unfortunately, the NLSY79 lacks a measure of household income for respondents prior to their inclusion in the survey (in adolescence). Thus, we use a proxy for family resources, maternal education, to restrict our sample to individuals who are likely to have been affected. Consistent with this approach, levels of maternal educational attainment among early Head Start participants were very low. We focus our analyses on the children of NLSY79 mothers whose mothers (i.e., the grandmothers of our population of interest) did not finish high school (column (2)).[16] We refer to this as our high impact sample because close to 70 percent of participants' mothers had less than a high school degree, implying participation rates of close to 60 percent in counties with Head Start availability.[17] As seen from the summary statistics, children in our high impact sample are negatively selected relative to the full NLSY sample. They have higher rates of teen parenthood (22% vs. 19%) and interaction with the criminal justice system (31% vs. 29%), and lower rates of high school completion (77% vs. 81%) and college going (52% vs. 58%). In column (3), we restrict the sample to NLSY79 mothers whose mother completed at most a high school degree. We call this our low impact sample as we estimate participation rates of at most 30 percent in counties with Head Start availability. As might be expected, summary statistics for children in this group have somewhat better outcomes.

We present analogous statistics for maternal outcomes (i.e., NLSY79 respondents) in the bot-

[16] Roughly 65 to 70 percent of the mothers of early participants reported less than a high school education, approximately 25 percent indicated that they graduated from high school, and only five percent of mothers reported some postsecondary schooling or more.

[17] See the note to Table A4 for details of these calculations.

tom panel of Table 1. While rates of teen parenthood are substantially higher and rates of college going are substantially lower, the pattern of statistics across samples is similar.[18]

4 Estimation and Empirical Results

To circumvent issues associated with individual selection into Head Start, we rely on variation within counties over time in the availability of the Head Start program. This variation is generated by the rollout of the Head Start program in the late 1960s. We identify the effect of program availability within a mother's county of birth on the adult outcomes of her children. Between 65 and 70 percent of the mothers of early Head Start participants did not complete high school (Barnow and Cain 1977, Office of Child Development 1968), so we focus our analyses on this "high impact" population. Our basic specification is:

$$y_{ict} = \beta_0 + \beta_1 X_i + \beta_2 HSavail_{ct} + \gamma_c + \lambda_t + \varepsilon_{ict} \tag{1}$$

where y_{ict} is an adult outcome for a child; X_i includes controls for the child's sex and age as well as the mother's birth order and race; and γ_c and λ_t are county of birth and birth year fixed effects. $HSavail_{ct}$ indicates whether Head Start was available for a mother in a particular birth cohort t and birth county c. $HSavail_{ct}$ is set to one for a mother when there is a non-trivial level of Head Start funding in that mother's birth county four or five years after her year of birth.[19] Standard errors are clustered at the county of birth level.

Table A4 illustrates that our measure of Head Start availability predicts both self-reported Head Start participation and state-level participation rates derived from administrative Head Start enrollment data. The top panel presents estimates of the effect of Head Start availability on self-reported participation. When a program is available in a county four or five years after a mother's year of birth, the mother in our high impact sample is 10 percentage points more likely *to report* having participated in Head Start as a child. The Head Start participation question is

[18] Similar measures of interaction with the criminal justice system are unavailable for this sample.

[19] Explicitly, $HSavail_{ct}$ is set to one when the level of Head Start funding within a county exceeds the 10th percentile of observed funding per four year old (roughly $22). The results are similar using a cutoff of $0 or the 5th or 15th percentile of funding (Appendix Tables A2 and A3).

asked retrospectively in 1994, when sample members were 30 to 34 years old. In other words, the participation measure reflects individuals' recollection of whether they participated in a program nearly 30 years earlier, when they were four or five years old. A variety of evidence from the psychology literature indicates that retrospective reports of early childhood are extremely unreliable (see Appendix B for further discussion). Given the self-reported and retrospective nature of the Head Start participation variable, we expect there is considerable misreporting.

As has been established in the literature, measurement error in a binary variable is necessarily non-classical and will thus result in a downwardly biased estimate of the relationship between Head Start availability and participation (Kane, Rouse and Staiger 1999, Aigner 1973). Under reasonable assumptions on the extent of misreporting, the true relationship between Head Start availability and Head Start participation may be four to eight times the observed estimate in our high impact sample, or 40 to 80 percentage points (see Appendix B for the details of these calculations). This larger first stage is consistent with that indicated by Johnson and Jackson (2018), who suggest a first stage among children in poor families of as high as 86 percentage points during this period.

That said, even with full confidence in the true first stage in our sample, we do not think it is reasonable to convert our estimates to treatment on the treated (TOT) effects because there are likely important spillover effects of program availability from participants to non-participants within the same cohort. This approach is consistent with prior studies of Head Start using similar designs, which all focus on reduced-form effects of Head Start availability (or grant-writing assistance) for similar reasons.

Nevertheless, the positive relationship between program availability and self-reported participation supports our research design. The middle panel contains estimates using state-level variation in participation rates in 1966.[20] The estimates suggest close to a 30 percentage point increase in the likelihood of participation for individuals in our high impact sample. However, this approach faces serious limitations and we mention it mainly as support for our empirical strategy; measures of availability do predict participation, even in the aggregate. The table also contains implied participation rates, assuming that all participation occurred in counties with Head Start availability.

[20] State-level enrollment numbers are not available in other years during this time period. The note to Table A4 contains the details of these calculations.

These estimates provide a rough upper bound for the effect of Head Start availability on participation in our high impact sample of nearly 60 percentage points, similar to the estimated relationship between Head Start availability and participation after correcting for misclassification error.[21]

Our baseline results are contained in Table 2. The first row contains our main estimates from our high impact sample, second-generation individuals whose grandmothers did not complete high school. We observe a large positive effect (0.32 sd) of a mother's Head Start availability on our index of adult outcomes for her children that is statistically significant at the 1 percent level. This effect is driven by reductions in teen parenthood (8 percentage points) and criminal behavior (15 percentage points) and increases in high school graduation (14 percentage points) and college enrollment (17 percentage points). The second row presents estimates from our low impact sample. As expected given their lower levels of Head Start participation and disadvantage, the effects are smaller for individuals in this group: a 0.15 sd increase on our index of adult outcomes.[22]

While our baseline inference relies on standard errors clustered at the county of birth level, we have also explored the robustness of our p-values to a variety of other assumptions, including clustering on birth county by birth cohort, clustering on birth state, and the wild cluster bootstrap (Appendix Tables A5 and A6). Finally, we consider "p-values" generated by the generally conservative approach of randomization inference.[23] Under this procedure (essentially a large set of placebo assignments), we randomly reassign the pattern of Head Start introduction timing to counties of birth and estimate our basic specification. We do this 1,000 times. The distribution of these estimates is contained in Appendix Figures A3 and A4. As displayed in the figures, the estimates we observe are quite unlikely under a random assignment of the availability of Head Start to a county.[24] P-values presented are the two-tailed statistics calculated as the share of coefficient estimates obtained under random assignment of Head Start timing that are larger in absolute magnitude than the estimate produced using the true timing assignment. Our randomization inference

[21] The upper bound in our sample may be somewhat higher than 60 percentage points given the overrepresentation of poor and black families in the sample, who had higher rates of participation.

[22] The estimates are consistent with the relative participation rates in the high-impact (60 percent) versus low-impact (30 percent) samples.

[23] See Abadie, Diamond, and Hainmueller (2010) for a discussion of this procedure and examples of its implementation.

[24] The random reassignment is at the level of county of birth.

"p-values" are similar to those obtained using other more standard approaches.

Consistent with the general approach in the literature, we focus on our index of outcomes as it maximizes statistical power and limits concerns about multiple comparisons. We are most interested in the question of the extent to which positive effects spillover to the second generation. While the second-generation effects appear to be driven by reductions in teen parenthood and crime and increases in educational attainment, we do not take a strong stand on the magnitude of the contribution of different components of our index given a relatively high level of uncertainty in our point estimates. That said, we have generated adjusted p-values for the components of our index following Romano and Wolf (2005). The adjusted p-values for teen parenthood, crime, high school, and some college are 0.02, 0.00, 0.00, and 0.00 for our primary (high impact) sample. The adjusted p-values for the low impact sample for teen parenthood, crime, high school, and some college are 0.12, 0.12, 0.06, and 0.12.

4.1 Threats to Internal Validity

To interpret these estimates as the causal effect of Head Start availability, it must be the case that the availability of a Head Start program is, conditional on county and year of birth fixed effects, unrelated to other factors that would affect the outcomes of children born to women who did and did not have the program available. For example, one concern would be that the type of woman who became a mother or the type of woman included in the sample (due to non-response or our sample restrictions) was affected by the availability of a Head Start program in early childhood. To check for this we examine how Head Start availability predicts maternal background characteristics that are unlikely to have been affected by Head Start directly (columns (1)-(5) of Table A7). We do this exercise separately for the full sample and the two restricted samples we use for our rollout analyses. There is little relationship between maternal characteristics (race, maternal birth order, 1978 household poverty status) and Head Start availability. Similarly, there is no evidence that the education levels of the grandmother, which we use to focus our sample, are affected by Head Start availability. In column (6) and (7), we present analogous estimates focused on second generation characteristics unlikely to be affected by Head Start: the age and gender of the child. While there is no relationship with child age, Head Start availability is correlated with child gender in our high impact sample.[25] Given the general balance of observables across samples and characteristics, we are not particularly concerned by this difference, but as a further check we explore the relationship between Head Start availability and a predicted index based on all of the characteristics in the table.[26] We would be concerned if there were a positive and significant coefficient. Instead, we see insignificant negative point estimates, suggesting that, if anything, the children of those who have Head Start available are likely to have worse outcomes based on exogenous observable characteristics. Finally, we explore how our treatment estimates are affected by the exclusion of covariates. As expected from the balance of observable characteristics, our estimates are robust to the exclusion of covariates, supporting the argument that the availability of Head Start is conditionally exogenous.[27]

A second concern is that of endogenous program adoption or pre-existing positive trends in outcomes in counties that adopted a Head Start program.[28] The presence of meaningful pre-trends might reflect general improvements in early childhood conditions or the existence of other programs that were correlated with Head Start availability. While the relatively tight window of analysis limits concerns related to pre-existing trends, we also probe the robustness of our estimates to the inclusion of differential trends by birth county (Table 3). Allowing differential birth cohort trends interacted with baseline (1960) county characteristics, or even birth county-specific trends, does little to change our point estimates. The estimates are similarly robust to the inclusion of more specific county-cohort controls for spending on War on Poverty programs and state by birth

[25] There is also no significant effect of Head Start exposure on a woman's decision to have children, the number of children born, and whether a woman has a child who responds to the survey after age 20.

[26] To implement this approach, we first regress the second-generation outcome index on all of the maternal and child characteristics in Table A2. Nearly all of the covariates are significant predictors at the 1% level and the p-value from an F-test in this regression is 0.000, indicating that these covariates are predictive of our index of outcomes (results not shown in table). We then regress this predicted index on Head Start availability and county and year of birth fixed effects.

[27] Concerns about differential selection of families into the sample are further limited by the inclusion of family fixed effects into the specification. While this approach has a number of limitations, the resulting point estimates (in column (8)) are smaller than, but statistically indistinguishable from, the other point estimates in Table 3. The attenuation may be a result of sibling spillovers, although we hesitate to draw conclusions given the very large confidence intervals and different sample that results from restricting to families with at least two sisters in the original NLSY79 sample.

[28] Our results are robust to restricting our sample to counties that received Head Start by the end of our sample period (column (9) of Table 3), so it would need to be the case that these positive trends occurred just prior to adoption and not just that adopting counties were on a positive trend relative to non-adopters for this issue to threaten our strategy.

cohort fixed effects, which flexibly control for changes over time within states that could affect maternal outcomes.[29]

Figure 2 addresses these concerns graphically, demonstrating the relationship between Head Start availability within a county and the index of adult outcomes for the second generation. The x-axis presents the number of years between a mother's year of birth and the first year of Head Start availability in a county. Those individuals with a non-negative value are considered treated in that Head Start was available in their county of birth for their birth cohort. As observed in the figure, the estimates are flat and close to zero prior to the availability of Head Start and then positive after a program becomes available, consistent with our estimates representing a causal effect of Head Start availability.

We further address endogeneity concerns related to the availability of a Head Start program using a placebo exercise. In Table 4, we explore the effect of Head Start availability on the children of a group of individuals who are largely ineligible for the program. Specifically, we run our basic specification on the children of mothers whose mothers obtained at least a high school degree. Only a small fraction of women in this group were eligible for or participated in Head Start.[30] If something other than Head Start availability is driving our main results, we might expect to see similar effects show up for the children of women in this group. Table 4 illustrates that the point estimates for this group are small, frequently opposite-signed, and statistically indistinguishable from zero across all outcomes.[31]

4.2 Grant-Writing Assistance Strategy

While the preceding approach and accompanying exercises strongly support a causal interpretation of our main estimates generated using the rollout strategy, we leverage a second source of variation to further validate our findings. Specifically, we use variation in Head Start exposure gen-

[29]The War on Poverty program variables control for spending on Medicaid, Community Action Program (CAP) administrative grants, cash assistance, CAP health programs, and Community Health Centers.

[30]We estimate that participation rates in this group were at most 1/5 of those in our high impact sample. Appendix Table A8 further illustrates that our measure of Head Start availability has a small and non-significant relationship with self-reported Head Start participation in this subsample.

[31]We similarly find non-significant point estimates in the sample with grandmothers who obtained more than a high school degree, but the sample sizes are too small and the estimates too imprecise for this exercise to provide much information.

erated by the provision of Head Start grant-writing assistance to the poorest 300 counties, adapting the approach of Ludwig and Miller (2007). The relatively small sample size and number of birth counties in the NLSY79 makes implementation of their regression discontinuity approach infeasible. However, unlike the National Education Longitudinal Study (NELS) sample used by Ludwig and Miller, in which individuals turned four years old in the late 1970s, the NLSY79 sample contains a number of cohorts of individuals who turned four before the creation of Head Start. This feature of our data supports a difference-in-differences approach in which we compare the difference in outcomes of children of women who turned four before the creation of Head Start (prior to 1965) with the children of women who turned four later across counties that did and did not receive grant-writing assistance. Our basic specification is:

$$y_{ict} = \beta_0 + \beta_1 X_i + \beta_2 GrantAssistance_c * Post_t + \gamma_c + \lambda_t + \varepsilon_{ict} \tag{2}$$

where y_{ict} is an adult outcome for a child; X_i includes controls for the child's sex and age as well as the mother's birth order and race; and γ_c and λ_t are county of birth and birth year fixed effects. $GrantAssistance_c$ indicates whether a mother's birth county c was one of the 300 poorest in 1960 and $Post_t$ is an indicator for whether a mother's birth cohort t was after 1961. We are primarily interested in β_2, which indicates the reduced-form effect of grant-writing assistance.

Table A9 illustrates that the availability of grant-writing assistance results in an increase in self-reported retrospective Head Start participation of 20 to 30 percentage points (42 to 62 percent), depending on the bandwidth of 1960 county poverty rates used to restrict the sample. These estimates are somewhat larger than, but statistically indistinguishable from, those reported in Ludwig and Miller (2007) using a regression discontinuity strategy and self-reported retrospective Head Start participation (at age 13) in the NELS. In both cases the self-reported and retrospective nature of the Head Start participation variable suggests that these are underestimates of the true first stages, although the extent of the downward bias is likely much larger in the NLSY79 data given that the question is asked of participants in their 30s.[32] Following our earlier discussion

[32] A natural question then is why our estimated effect on self-reported participation is, if anything, larger than that in Ludwig and Miller (2007), suggesting a much larger first stage in our sample. The differences could be explained by the greater proportion of poor individuals in our sample, the different set of birth cohorts used (the NELS participants

on the extent of misreporting, the true relationship between grant-writing assistance and Head Start participation is likely multiple times the observed estimate, suggesting a first stage that approaches one.[33] Furthermore, given the extremely high share of preschool-aged children treated in these counties, we think spillover effects are quite likely. With these concerns in mind, we do not interpret the estimates in Table A9 as a first stage. Instead, as in Ludwig and Miller (2007), we focus our attention on the intent-to-treat (ITT) parameter, the effect of grant-writing assistance.

Turning to our main outcomes of interest, Table 5 indicates large intergenerational effects of Head Start grant-writing assistance on our index of second-generation outcomes. This effect is driven by reductions in teen parenthood and criminal behavior and increases in high school completion and college enrollment. Interestingly, the reduced-form estimates generated by this approach are similar in magnitude to those estimated using our rollout strategy; we reconcile the results further in the next section.

The estimates are unaffected by the exclusion of covariates, indicating that these effects are not driven by differential changes in the composition of mothers included in the NLSY79 sample (Table A10). Another potential threat to internal validity in this design would arise if the 300 poorest counties had been targeted for the provision of other forms of social spending. As noted by Ludwig and Miller (2007), this "seems unlikely since the decision to focus Head Start grant-writing assistance in the 300 poorest counties seems to have been made arbitrarily within the Head Start office rather than as some part of larger OEO-wide policy." However, as we are unable to rely on the discontinuity in funding, differential changes in social spending may still be a concern if the poorest 300 counties received substantially more support after 1965 relative to those counties on the other side of the threshold.[34] Figure III of Ludwig and Miller (2007) suggests that this was not the case, with at most very small differences in average social spending across counties with

were born 10 to 15 years after our sample by which time the effect of grant-writing assistance in the early Head Start years was likely attenuated), or the assignment of treatment county based on county of residence at age 13 versus at birth.

[33] The scaling factors here are somewhat more complicated to infer for two reasons. First, we have no administrative participation data restricted to the set of counties in our analyses from which to generate a data-driven recall error estimate. Second, we know that the earlier cohorts' participation is correctly reported as having not participated given that Head Start did not exist when these individuals were four or five.

[34] Note that all of the counties included in these regressions have very high poverty rates due to our bandwidth restriction.

poverty rates between 40 and 80 percent in 1972. As before, we can also control for county-by-year levels of War on Poverty spending directly, which has no effect on our point estimates. Finally, we can include controls for 1960 county characteristics (including the poverty rate) interacted with a trend in birth cohort. If, for example, the concern is that poorer counties received additional resources over this period, then the interaction of poverty rate and time trend should absorb this potentially problematic variation. The point estimate is hardly affected by these controls, further indicating that the estimates are reflecting the effect of grant-writing assistance.[35]

4.3 Discussion of Effect Sizes

While our analyses focus on the reduced-form effects of Head Start availability and grant-writing assistance, in this subsection we provide discussion of the implied magnitudes of treatment effects, comparing across strategies within our paper as well as reconciling our results with existing work. We begin with a discussion of the likely first stage in Head Start participation across our samples and approaches as well as the implied treatment on the treated in the absence of spillovers. We then situate our results in the small literature that explores the intergenerational effects of a schooling or early childhood intervention. Finally, as further validation of our approach, we compare the first-generation Head Start effects estimated using our strategies with previous estimates in the literature.

4.3.1 Comparing Effect Sizes across Strategies

Interestingly, the reduced-form effects estimated using our rollout and grant-writing assistance strategy are quite similar. This need not be case if Head Start availability and grant-writing assistance generate substantially different increases in Head Start participation, if the effect of Head Start participation was greater in poorer areas, or if the extent of spillovers was larger in poorer areas. While we cannot isolate these factors, we can provide some discussion, particularly with respect to the effect of Head Start availability versus grant-writing assistance on participation in Head Start. As discussed previously, we put little stock in "first-stages" estimated using the

[35]The estimates are also robust to a variety of choices of 1960 poverty rate bandwidths that we use to restrict the set of birth counties included in the analyses (Table A11).

self-reported and retrospective Head Start participation variable because the participation question is asked retrospectively in 1994, when sample members were 30 to 34 years old (Appendix Table A4). The measure thus reflects individuals' recollection of whether they participated in a program nearly 30 years earlier, when they were four or five years old. A variety of evidence indicates that retrospective reports of early childhood are extremely unreliable, even for meaningful events such as parental divorce or separation, household incarceration, parental chronic unemployment, or residence changes (see Appendix B for further details). Given the self-reported and retrospective nature of the Head Start participation variable, we expect there is considerable misreporting.

As discussed previously, this recall error will result in a downwardly biased estimate of the relationship between Head Start availability (or grant-writing assistance) and participation. In Appendix B, we walk the reader through two approaches for estimating the extent of the recall bias in our rollout estimates. The first approach uses the available administrative and self-reported participation data to generate estimates of the misclassification rates, while the second draws upon the existing literature on retrospective recall. Combined, they suggest that the true relationship between Head Start availability and Head Start participation is four to eight times the observed estimate in our high impact sample, or 40 to 80 percentage points.

Our suggested first stage estimates are broadly consistent with the other papers using rollout variation in Head Start exposure in their identification strategies. For example, Johnson and Jackson (2018) estimate participation rates of around 86 percent for early cohorts in their sample (similar to the birth cohorts used in our analyses) and use a 75 percentage-point increase among "poor children" due to rollout as their "ballpark" estimate for scaling for their full period.[36] Their sample of poor families has relatively similar rates of eligibility as our high-impact sample. While we hesitate to similarly restrict our sample due to the endogeneity of potential income measures in the NLSY79, first- and second-generation estimates using similar specifications as those employed by Johnson and Jackson are consistent with our main analyses (Appendix Tables A12 and A13).[37]

[36] They define poor children as those whose parents were in the bottom quartile of the income distribution, with 80% of these individuals below the poverty line.

[37] We follow Johnson and Jackson as closely as possible, restricting our sample based on the poverty level of the NLSY79 respondent's family in 1978 (rather than by the bottom quartile of the income distribution) due to available information. We are hesitant to use this sample restriction more generally due to concerns about the endogeneity of 1978 poverty level (which is measured after potential exposure to Head Start).

The estimates in Appendix Table A9 suggest a somewhat larger first stage using the grant-writing assistance strategy. That said, it is more difficult to use the data to infer the extent of misclassification within this sample. The literature on misclassification error suggests scaling the participation estimates by as much as four to eight times, suggesting a true first stage that approaches one. However, the literature driven misclassification rates are almost certainly overestimates for the grant-writing assistance strategy because the participation of those born too early to have attended Head Start is set automatically (and accurately) to zero. An upper bound is provided by the fraction of children in counties that received grant-writing assistance who were eligible to participate in Head Start, roughly 80 percent.[38]

Regardless of specification or sample choice, if we assume the entirety of our estimated effects of Head Start access was generated by participation (and there were no spillover effects), the implied TOT effects are large, suggesting increases in second-generation high school completion and college enrollment of at least 20 percentage points, with 90-percent confidence intervals including implied TOT estimates of closer to 10 percentage points for the rollout approach.

That said, we do not think it is reasonable to convert our estimates to TOT effects as there are likely important spillover effects of program availability; indeed, it seems likely that improving the trajectories of a significant share of a group would result in improvements for the group as a whole that are substantially larger than what we might expect to see if an individual was treated in isolation (as in a small scale experiment, for example). Concretely, teen parenthood (always) and criminal behavior (frequently) involve more than one party, so it is easy to see how these spillovers would occur. Spillovers could also emerge if subsequent schooling quality or productivity increased because a share of each school-entry cohort was more prepared. Given the concerns raised here and those raised previously by other authors pursuing similar approaches, we focus on the estimated effect of Head Start availability.

[38]This upper bound is likely an underestimate in our sample because the NLSY oversamples poor and black populations, who have higher rates of Head Start participation.

4.3.2 Comparing Intergenerational Effect Sizes with Other Estimates

Using the rollout approach, the effects on the second generation are at least as large as effects on the first generation; we cannot reject the equivalence of effect sizes for comparable measures (Appendix Table A14). While the level effects are generally larger in the second generation, the percent changes are similar for most education measures.[39]

We continue to find larger estimated effects in the second generation using our grant-writing assistance approach (Appendix Table A15). That said, our point estimates of the effect of grant-writing assistance on first-generation educational outcomes measured among 40 and 50 year-olds are generally one-half to one-third the size that Ludwig and Miller estimate among 25 year-olds (despite a likely larger first-stage in our data). This underscores the difficulty of comparing effect sizes across generations, particularly when outcomes are measured at different ages. Indeed, the effects reported by Ludwig and Miller for the first generation in the NELS are quite similar to the effects we estimate in the second generation.

While there are few benchmarks for comparison, this high intergenerational correlation in effects is consistent with some recent findings (Rossin-Slater and Wst 2019, Akresh et al. 2018, Mazumder et al. 2019, Heckman and Karapakula 2019).[40] Estimates of the intergenerational effects of new school construction in Indonesia indicate that the effect on maternal educational attainment in the first generation precedes second-generation increases in test scores and educational attainment (Akresh et al. 2018, Mazumder et al. 2019). As in our study, for most outcomes the effect on a mother's children's educational attainment is larger than the effect on her own attainment. Rossin-Slater and Wust (2019) explore the intergenerational impact of, and interaction between, a Danish preschool program and a nurse home-visiting program in infancy. They also find educational attainment effects in the first generation that persist in the second generation.

Perhaps most closely related to our study is a very recent long-term follow-up of the experimental Perry Preschool Project that explores effects among the children of participants (Heckman

[39]We do not include the estimate on highest grade completed in the second generation in our main tables as many of these individuals have not finished their education; however, the point estimates are 0.67 (se 0.23) for the high impact sample and 0.28 (se 0.18) for the low impact sample, statistically indistinguishable from the effect on the mothers.

[40]Indeed, as in our study, the magnitudes of comparable first and second generation effects are statistically indistinguishable in all of these studies.

and Karapakula 2019). Here too, the authors estimate large effects of program participation on a participant's children. While the reported outcomes are not identical across generations, among comparable outcomes the effect sizes are quite similar for participants and their children. For example, a previous study estimates an increase in the likelihood of high school graduation of around 39 percentage points for female participants and the intergenerational study finds the likelihood of graduating without a suspension increases 34 percentage points for the participants' children (Nores, Belfield, Barnett and Schweinhart 2005, Heckman and Karapakula 2019). Similarly, the studies estimate that the program increases the likelihood of employment by 27 percentage points among female participants and 26 percentage points for the participants' children (Nores et al. 2005, Heckman and Karapakula 2019).

Across varied contexts, time periods, and types of interventions, recent evidence on intergenerational effects suggests that program impact can transfer to the children of those directly exposed, and in substantial ways. Moreover, mothers' human capital has emerged as a potentially critical channel through which intergenerational effects are realized.

4.3.3 Reconciling our First-Generation Results with Existing Estimates of the Effects of Early Childhood Programs

As in evaluations of Perry Preschool, our estimated effects and the implied intergenerational relationship in effect sizes are large. While we believe our approach captures causal estimates and, importantly, the direction of Head Start effects, the confidence intervals around the estimates are relatively wide. Reconciling our first-generation estimates with published and forthcoming estimates of Head Start and similar early childhood education programs may help with interpretation of the magnitude of Head Start effects. The top panel of Appendix Table A16 provides reduced-form effects of Head Start availability for similar approaches and populations. Across different approaches and samples, the reduced-form effect of Head Start availability on comparable outcomes is similar, with generally larger effects reported by Johnson and Jackson (2018) and smaller effects reported by Thompson (2018). It is important to note that the Head Start participation rate in Thompson's sample is roughly half of that in our high impact sample due to our different sample restrictions. In nearly all cases, the 95-percent confidence interval of the implied TOT overlaps.

Similarly, the reduced-form effects of grant-writing assistance reported by Ludwig and Miller (2007) are statistically indistinguishable from the estimates we report for comparable outcomes using a similar approach. They report a 12 percentage-point increase in high school completion and 15 percentage-point increase in college enrollment among NELS respondents, somewhat larger than the six and five percentage-point increases we estimate in the first generation, although again the confidence intervals overlap.

While we think it is important to use caution in interpreting TOT estimates implied by our study, our implied TOT effects under the assumption of no spillovers allow us to provide some comparison across studies with different approaches.[41] We have similarly adjusted the ITT estimates reported by other authors, although it is worth again noting that nearly all of these authors are resistant to scaling their reduced-form estimates for similar reasons to those mentioned previously. With these caveats in mind, we note that our estimates tend to fall in the middle of the TOT estimates implied by other authors' reduced-form estimates and suggested first stages. That our first-generation estimates are in line with previous estimates provides further confidence in our approach.

5 Discussion and Conclusion

Research and policy discussion frequently focus on how to level the playing field for those born into poverty. We focus instead on whether such interventions truly break the cycle of bad outcomes. While there is increasing interest among researchers and policymakers in understanding the intergenerational spillovers of particular policies and interventions, there exist very few contexts in which these questions can be tested empirically; this is particularly true for early childhood interventions, an area of increasing focus and investment. The federal Head Start program provides a context in which data availability and the time horizon since first implementation facilitate such an empirical exploration, allowing us to contribute the first U.S. evidence on the intergenerational

[41] Informed by our earlier discussion of first stages and for ease of comparison, we scale our high impact estimates by 1/0.6 and our grant-writing assistance estimates by 1/0.7. The true first-stages may be somewhat different and we caution the reader in interpreting our scaled estimates.

transmission of early childhood intervention effects.

We find consistent evidence that the positive effects of Head Start during its earliest years transferred across generations in the form of improved long-term outcomes for the second generation. The pattern of results suggests decreases in teen parenthood and criminal engagement and increases in educational attainment across empirical approaches. The effects are large in magnitude, but broadly consistent with the first-generation effect sizes found in evaluations of similar early childhood programs that provided an array of services to disadvantaged youth.[42] Furthermore, because of the large scale of Head Start, the program likely provided benefits beyond the direct effect on participants. The key takeaway is that the benefits of Head Start transfer from one generation to the next in a substantial way.

Given the documented importance of mothers' human capital for their children's early health and developmental outcomes, we could expect that mother's own increased educational attainment and earnings would translate to improvements in birth outcomes, child health, and development. Importantly, early investments may interact complementarily with subsequent investments and interventions that amplify the effect of both those initial conditions and inputs throughout the life course (Cunha and Heckman 2007, Heckman and Masterov 2007, Heckman and Mosso 2014). It is also plausible that mothers who experienced improved life chances via their own exposure to Head Start make different choices about the peers with whom they and their children interact, the networks with which they engage, and the environments in which their children grow and learn. While any single decision may be incrementally important for their children's outcomes, the cumulative effect of a variety of improved inputs and conditions could be quite large in magnitude, especially as experienced prenatally, at birth, and throughout childhood and adolescence.

Indeed, the availability of Head Start, at least during the early years of the program, appears to have been quite successful at breaking the cycle of poor outcomes for disadvantaged families. Head Start access closes most of the gap in outcomes between individuals with more and less advantaged grandmothers. These results imply that cost-benefit analyses of Head Start and similar early

[42]Recall that, during the early years of the program, Head Start provided medical and dental care, engaged parents of the participants, and facilitated linkages to other social services for an extremely disadvantaged population of children and families.

childhood interventions underestimate the benefits of such programs by ignoring the transmission of positive effects across generations. This finding has important policy implications for optimal investment in these types of programs. Each disadvantaged child society helps now will lead to fewer who require assistance in the future.

References

Aigner, Dennis J, "Regression with a binary independent variable subject to errors of observation," *Journal of Econometrics*, 1973, *1* (1), 49–59.

Akresh, Richard, Daniel Halim, and Marieke Kleemans, "Long-term and intergenerational effects of education: Evidence from school construction in Indonesia," Working Paper 25265, National Bureau of Economic Research 2018.

Almond, Douglas and Janet Currie, "Killing me softly: The fetal origins hypothesis," *Journal of economic perspectives*, 2011, *25* (3), 153–72.

—— **, Hilary W. Hoynes, and Diane Whitmore Schanzenbach**, "Inside the War on Poverty: The impact of food stamps on birth outcomes," *Review of Economics and Statistics*, 2011, *93(2)*, 387–403.

Auten, Gerald, Geoffrey Gee, and Nicholas Turner, "Income inequality, mobility, and turnover at the top in the US, 1987-2010," *American Economic Review*, 2013, *103(3)*, 168–172.

Bailey, Martha J. and Andrew Goodman-Bacon, "The War on Poverty's experiment in public medicine: Community health centers and the mortality of older Americans," *American Economic Review*, 2015, *105(3)*, 1067–1104.

—— **, Shuqiao Sun, and Brenden Timpe**, "Prep school for poor kids: The long-run impacts of Head Start on human capital and economic self-sufficiency," 2018. mimeo.

Barnow, Burt S. and Glen G. Cain, "A reanalysis of the effect of Head Start on cognitive development: Methodology and empirical findings," *Journal of Human Resources*, 1977, *12(2)*, 177–197.

Bauer, Lauren and Diane Whitmore Schanzenbach, "The long-term impact of the Head Start program," Economic Analysis, Brookings Institution, The Hamilton Project, Washington, DC 2016.

Bitler, Marianne P., Hilary W. Hoynes, and Thurston Domina, "Experimental evidence on the distributional effects of Head Start," Working Paper 20434, National Bureau of Economic Research 2014.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes, "Why the apple doesn't fall far: Understanding intergenerational transmission of human capital," *American Economic Review*, 2005, *95(1)*, 437–449.

Black, Sandy and Paul Devereux, "Recent developments in intergenerational mobility," *Handbook of Labor Economics*, 2011, *4*, 1487–1541.

Campbell, Frances, Gabriella Conti, James J. Heckman, Seong Hyeok Moon, Rodrigo Pinto, Elizabeth Pungello, and Yi Pan, "Early childhood investments substantially boost adult health," *Science*, 2014, *343*, 1478–1485.

Carneiro, Pedro and Rita Ginja, "Long-Term impacts of compensatory preschool on health and behavior: Evidence from Head Start," *American Economic Journal: Economic Policy*, 2014, *6(4)*, 135–173.

Cascio, Elizabeth U, "Maternal labor supply and the introduction of kindergartens into American public schools," *Journal of Human resources*, 2009, *44* (1), 140–170.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan, "How does your kindergarten classroom affect your earnings? Evidence from Project STAR," *Quarterly Journal of Economics*, 2011, *126(4)*, 1593–1660.

___ **, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez**, "Where is the land of opportunity? The geography of intergenerational mobility in the United States," Working Paper 19843, National Bureau of Economic Research 2014.

___ **,** ___ **,** ___ **,** ___ **, and Nicholas Turner**, "Is the United States still a land of opportunity? Recent trends in intergenerational mobility," Working Paper 19844, National Bureau of Economic Research 2014.

Chevalier, Arnaud, "Parental education and child's education: A natural experiment," IZA Discussion Paper 1153, Institute for the Study of Labor 2007.

Corak, Miles, "Income inequality, equality of opportunity, and intergenerational mobility," *Journal of Economic Perspectives*, 2013, *27(3)*, 79–102.

Cornelissen, Thomas, Christian Dustmann, Anne Raute, and Uta Schnberg, "Who benefits from universal child care? Estimating marginal returns to early child care attendance," *Journal of Political Economy*, 2018, *126(6)*, 2356–2409.

Cunha, Flavio and James Heckman, "The technology of skill formation," Working Paper 12840, National Bureau of Economic Research 2007.

Currie, Janet and Enrico Moretti, "Mother's education and the intergenerational transmission of human capital: Evidence from college openings," *Quarterly Journal of Economics*, 2003, *118(4)*, 1495–1532.

Deming, David, "Early childhood intervention and life-cycle skill development: Evidence from Head Start," *American Economic Journal: Applied Economics*, 2009, *1(3)*, 111–134.

Duncan, Greg, Ariel Kalil, Susan E. Mayer, Robin Tepper, and Monique R. Payne, "The apple does not fall far from the tree," *Unequal chances: Family background and economic success*, 2005, pp. 23–79.

Duncan, Greg J., Chantelle J. Dowsett, Amy Claessens, Katherine Magnuson, Aletha C. Huston, Pamela Klebanov, Linda S. Pagani, Leon Feinstein, Mimi Engel, Jeanne Brooks-Gunn, Holly Sexton, and Kathryn Duckworth, "School readiness and later achievement," *Developmental Psychology*, 2007, *43(6)*, 1428–1446.

Dynarski, Susan, Joshua Hyman, and Diane Whitmore Schanzenbach, "Experimental evidence on the effect of childhood investments on postsecondary attainment and degree completion," *Journal of Policy Analysis and Management*, 2013, *32(4)*, 692–717.

East, Chloe N., Sarah Miller, Marianne Page, and Laura R. Wherry, "Multi-generational impacts of childhood access to the safety net: Early life exposure to Medicaid and the next generation's health," Working Paper 23810, National Bureau of Economic Research 2019.

Fryer, Roland G. and Steven D. Levitt, "Understanding the black-white test score gap in the first two years of school," *Review of Economics and Statistics*, 2004, *86(2)*, 447–464.

___ and ___ , "The black-white test score gap through third grade," *American Law and Economics Review*, 2006, *8(2)*, 249–281.

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

Gibbs, Chloe, Jens Ludwig, and Douglas L. Miller, "Does Head Start do any lasting good?," Working Paper 17542, National Bureau of Economic Research 2011.

___ , ___ , and ___ , "Head Start: Origins and impacts," in Martha J. Bailey and Sheldon Danziger, eds., *Legacies of the War on Poverty*, New York, NY: Russell Sage Foundation Press, 2013, pp. 39–65.

Greenberg, Polly, "Head Start–Part of a multi-pronged anti-poverty effort for children and their families... Before the beginning: A participant's view," *Young Children*, 1990, *45* (6), 40–52.

Heckman, James J., "The economics, technology, and neuroscience of human capability formation," *PNAS*, 2007, *104(33)*, 13250–13255.

___ and Dimitriy V. Masterov, "The productivity argument for investing in young children," *Review of Agricultural Economics*, 2007, *29(3)*, 446–493.

___ and Ganesh Karapakula, "Intergenerational and intragenerational externalities of the Perry Preschool Project," Working Paper 2019-033, Human Capital and Economic Opportunity Working Group 2019.

___ and Stefano Mosso, "The economics of human development and social mobility," *Annual Review of Economics*, 2014, *6*, 689–733.

——, **Rodrigo Pinto, and Peter Savelyev**, "Understanding the mechanisms through which an influential early childhood program boosted adult outcomes," *American Economic Review*, 2013, *103(6)*, 2052–2086.

Hertz, Tom, Tamara Jayasundera, Patrizio Piraino, Sibel Selcuk, Nicole Smith, and Alina Verashchagina, "The inheritance of educational inequality: International comparisons and fifty-year trends," *The BE Journal of Economic Analysis & Policy*, 2008, *7* (2), 10.

Hoynes, Hilary W., Diane Whitmore Schanzenbach, and Douglas Almond, "Long-run impacts of childhood access to the safety net," *American Economic Review*, 2016, *106(4)*, 903–934.

Hoynes, Hillary, Marianne Page, and Ann Huff Stevens, "Can targeted transfers improve birth outcomes? Evidence from the introduction of the WIC program," *Journal of Public Economics*, 2011, *95(7-8)*, 813–827.

Johnson, Rucker C. and C. Kirabo Jackson, "Reducing inequality through dynamic complementarity: Evidence from Head Start and public school pending," Working Paper 23489, National Bureau of Economic Research 2018.

Kane, Thomas J, Cecilia Elena Rouse, and Douglas Staiger, "Estimating returns to schooling when schooling is misreported," Technical Report, National bureau of economic research 1999.

Kline, Patrick and Christopher R. Walters, "Evaluating public programs with close substitutes: The case of Head Start," *Quarterly Journal of Economics*, 2016, *131(4).*

Knudsen, Eric I., James J. Heckman, Judy L. Cameron, and Jack P. Shonkoff, "Economic, neurobiological, and behavioral perspectives on building America's future workforce," in "Proceedings of the National Academy of Sciences of the United States of America" National Academy of Sciences Washington, DC 2006, pp. 10155–10162.

Lee, Chul-In and Gary Solon, "Trends in intergenerational income mobility," *Review of Economics and Statistics*, 2009, *91(4)*, 766–772.

Ludwig, Jens and Douglas L. Miller, "Does Head Start improve children's life chances? Evidence from a regression discontinuity design," *Quarterly Journal of Economics*, 2007, *122(1)*, 159–208.

Maurin, Eric and Sandra McNally, "Vive la revolution! Long-term education returns of 1968 to the angry students," *Journal of Labor Economics*, 2008, *26(1)*, 1–33.

Mazumder, Bhashkar, "Fortunate sons: New estimates of intergenerational mobility in the United States using social security earnings data," *Review of Economics and Statistics*, 2005, *87* (2), 235–255.

——, **Maria Rosales-Rueda, and Margaret Triyana**, "Intergenerational human capital spillovers: Indonesia's school construction and its effects on the next generation," in "AEA Papers and Proceedings," Vol. 109 2019, pp. 243–49.

Miller, Sarah and Laura R. Wherry, "The long-term effects of early life Medicaid coverage," *Journal of Human Resources*, 2019, *54(3)*, 785–824.

Montialoux, Claire, "Revisiting the impact of Head Start," Policy Brief, University of California, Berkeley, Institute for Research on Labor and Employment, Berkeley, CA 2016.

Murnane, Richard J., John B. Willett, Kristen L. Bub, and Kathleen McCartney, "Understanding trends in the black-white achievement gaps during the first years of school," in "Brookings-Wharton Papers on Urban Affairs" Brookings Institution Press Washington, DC 2006, pp. 97–135.

Nores, Milagros, Clive R. Belfield, W. Steven Barnett, and Lawrence Schweinhart, "Updating the economic impacts of the High/Scope Perry Preschool program," *Educational Evaluation and Policy Analysis*, 2005, *27* (3), 245–261.

Office of Child Development, "Project Head Start 1965-1967: A descriptive report of programs and participants," Report, Department of Health, Education, and Welfare, Washington, DC 1968.

——, "Project Head Start 1968: A descriptive report of programs and participants," Report, Department of Health, Education, and Welfare, Washington, DC 1970.

Oreopoulos, Philip, Marianne E. Page, and Ann Huff Stevens, "The intergenerational effects of compulsory schooling," *Journal of Labor Economics*, 2006, *24(4)*, 729–760.

Page, Marianne, "Fathers' education and children's human capital: Evidence from the World War II G.I. Bill," 2009. mimeo.

Page, Marianne E., "New evidence on the inter generational correlation in welfare participation," *Generational income mobility in North America and Europe*, 2004, p. 226.

Puma, Michael, Stephen Bell, Ronna Cook, Camilla Heid, and et al., "Head Start impact study: Final report," Report, U.S. Department of Health and Human Services, Administration for Children and Families, Washington, DC 2010.

——, ——, ——, ——, Michael Lopez, and et al., "Head Start impact study: First year findings," Report, U.S. Department of Health and Human Services, Administration for Children and Families, Washington, DC 2005.

——, ——, ——, ——, Pam Broene, Frank Jenkins, Andrew Mashburn, and Jason Downer, "Third grade follow-up to the Head Start impact study: Final report," Report 2012-45, U.S. Department of Health and Human Services, Administration for Children and Families, Washington, DC 2012.

Reardon, Sean F. and Ximena A. Portilla, "Recent trends in income, racial, and ethnic school readiness gaps at kindergarten entry," *AERA Open*, 2016, *2(3)*, 1–18.

Rossin-Slater, Maya and Miriam Wst, "What is the added value of preschool for poor children? Long-term and intergenerational impacts and interactions with an infant health intervention," Working Paper 22700, National Bureau of Economic Research 2019.

Sawhill, Isabel V., Scott Winship, and Kerry Searle Grannis, "Pathways to the middle class: Balancing personal and public responsibilities," Report, Brookings Institution, Washington, DC 2012.

Schweinhart, Lawrence J., Jeanne Montie, Zongping Xiang, W. Steven Barnett, Clive R. Belfield, and Milagros Nores, *Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40*, Ypsilanti, MI: High/Scope Press, 2005.

Solon, Gary, "Intergenerational mobility in the labor market," *Handbook of labor economics*, 1999, *3*, 1761–1800.

Thompson, Owen, "Head Start's long run impact: Evidence from the program's introduction," *Journal of Human Resources*, 2018, *53(4)*, 1100–1139.

Vinovskis, Maris A., *The Birth of Head Start*, Chicago, IL: University of Chicago Press, 2005.

Walters, Christopher R., "Inputs in the production of early childhood human capital: Evidence from Head Start," *American Economic Journal: Applied Economics*, 2015, *7(4)*, 76–102.

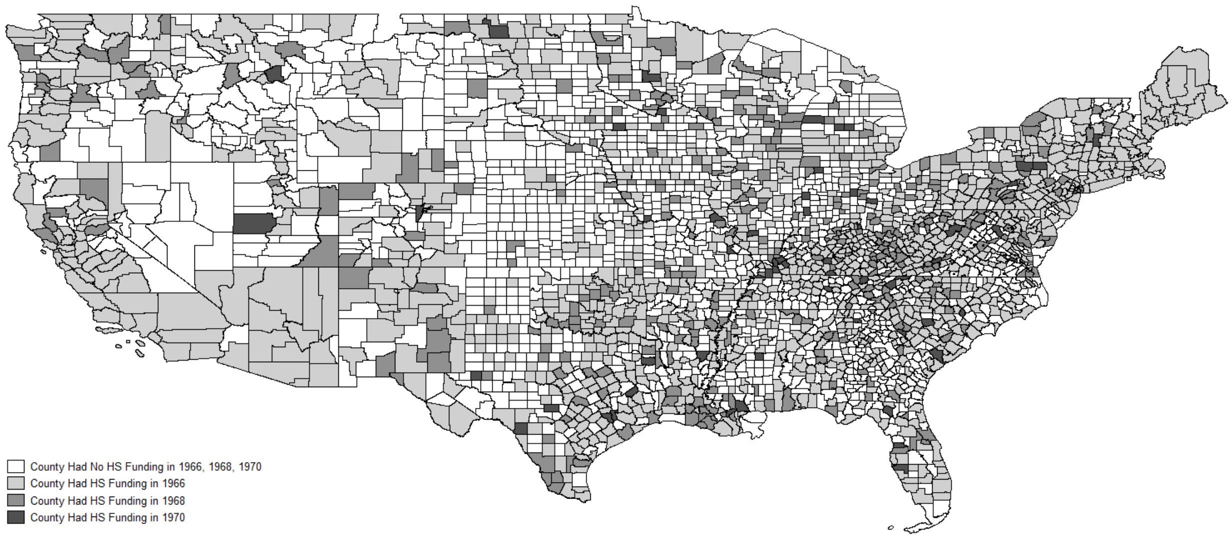
White, Sheldon H. and Deborah A. Phillips, "Designing Head Start: Roles played by developmental psychologists," in David L. Featherman and Maris A. Vinovskis, eds., *Social science and policy-making: A search for relevance in the twentieth century*, Ann Arbor, MI: University of Michigan Press, 2001, pp. 83–118.

Zigler, Edward and Jeanette Valentine, *Project Head Start: A Legacy of the War on Poverty*, New York, NY: Free Press, 1979.

___ **and Sally J. Styfco**, *The Hidden History of Head Start*, Oxford, UK: Oxford University Press, 2010.

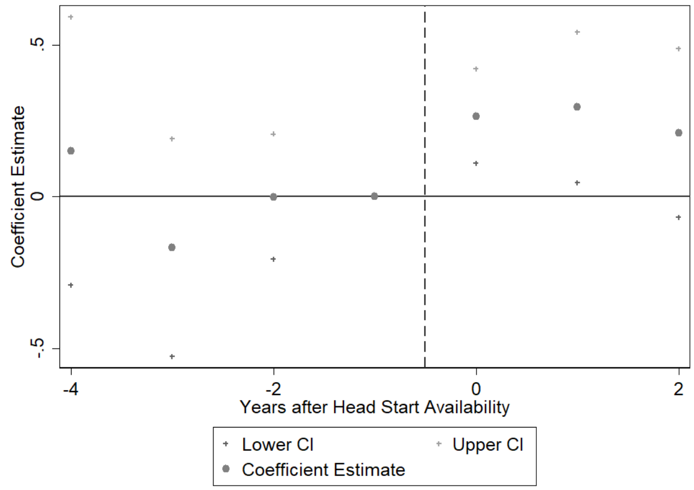
Figures

Figure 1: Early Geographic Expansions of Head Start



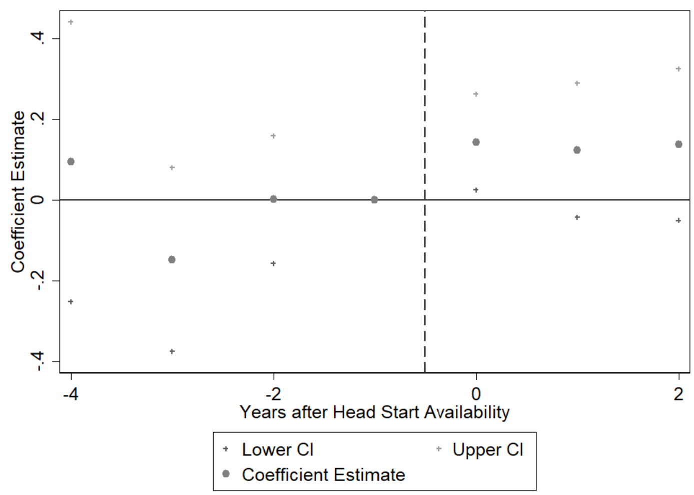
Source: National Archives and Records Administration.

Figure 2: Event Study for Second-Generation Index (High Impact)



Note: Our "high impact" sample is restricted to participants in the NLSY79 whose mothers dropped out of high school. Circles indicate coefficients on indicator variables for the difference between a mother's birth cohort and the first birth cohort to have Head Start availability within a county (non-negative years reflect cohorts with Head Start availability). The dependent variable is the index of adult outcomes in the second generation. See the notes to Tables 1 and 2 for additional information on sample restrictions and index construction. Regressions include cohort and county of birth fixed effects. Standard errors are clustered at the birth county level. Asterisks indicate 95 percent confidence intervals.

Figure 3: Event Study for Second-Generation Index (Low Impact)



Note: Our "low impact" sample, is restricted to participants in the NLSY79 whose mothers completed no education beyond high school. Circles indicate coefficients on indicator variables for the difference between a mother's birth cohort and the first birth cohort to have Head Start availability within a county (non-negative years reflect cohorts with Head Start availability). The dependent variable is the index of adult outcomes in the second generation. See the notes to Tables 1 and 2 for additional information on sample restrictions and index construction. Regressions include cohort and county of birth fixed effects. Standard errors are clustered at the birth county level. Asterisks indicate 95 percent confidence intervals.

Tables

Table 1: Descriptive Statistics of NLSY79 Women and Their Children

	(1)	(2)	(3)
		High Impact	Low Impact
Sample	(Full Sample)	(Grandmother <HS)	(Grandmother ≤HS)
Second Generation (Child) Outcomes			
Teen Parent	0.19	0.22	0.19
Crime	0.29	0.31	0.30
High School	0.81	0.77	0.81
Some College	0.58	0.52	0.56
Index	-0.0217	-0.11	-0.04
Observations	3216	1683	2729
First Generation (Mother) Outcomes			
Teen Mom	0.41	0.52	0.43
High School	0.85	0.80	0.84
Some College	0.44	0.32	0.40
Highest Grade	12.9	12.4	12.8
Income	41697	33859	39321
Observations	1638	806	1380

Note: The bottom panel presents sample means for women in the NLSY79. The top panel provides sample means for the children of these women, restricted to individuals over 20 in 2012. Each column provides sample means for a different sample, corresponding to a particular research design and set of sample restrictions. Column (1) provides sample means for NLSY79 women and their children. Columns (2) and (3) provide analogous means for the two samples used with our preferred research design, the changing availability of Head Start within counties. Each of these samples is restricted based on the education level of the mother of the NLSY79 participant (i.e., the grandmother of the children). Column (2), our "high impact" sample is restricted to participants in the NLSY79 whose mothers dropped out of high school. Column (3), our "low impact" sample, is restricted to participants in the NLSY79 whose mothers attempted no education beyond high school. All samples are restricted to mothers who were born in 1960 and after and were not part of the military sample. We largely follow Deming (2009) in our construction of crime and income measures. Crime is defined as any arrests, convictions, or probations. Income is a measure of the mother's permanent lifetime income, calculated as the deflated average of net family income for the mother. To address multiple inference concerns and reduce measurement error, we follow the prior literature in constructing a summary index of our outcome measures (Kling, Liebman, and Katz 2007; Deming 2009). We normalize each outcome to have a mean zero and standard deviation one, adjust outcome signs so that a more positive outcome is better (i.e., we flip the sign on teen parenthood and crime), and take the simple average across these outcomes.

Table 2: Geographic Variation: Reduced Form Effect of Head Start in County

	(1) Teen Parent	(2) Crime	(3) High School	(4) Some College	(5) Index
High Impact: Grandmother < High School	-0.078**	-0.150***	0.139***	0.175***	0.317***
	(0.031)	(0.043)	(0.047)	(0.056)	(0.073)
Observations	1,683	1,683	1,683	1,662	1,683
Mean	0.220	0.305	0.768	0.516	-0.113
Low Impact: Grandmother ≤ High School	-0.048*	-0.070**	0.071*	0.071	0.153**
	(0.027)	(0.034)	(0.039)	(0.047)	(0.060)
Observations	2,729	2,729	2,728	2,692	2,729
Mean	0.194	0.296	0.807	0.564	-0.0390

Note: Each column represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's county of birth. The dependent variables are second generation outcomes as indicated by the column titles. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. In addition to year of birth and county of birth fixed effects, controls include child gender, age, and age squared, as well as mother's birth order and race. * (p<0.10), **(p<0.05), ***(p<0.01).

Table 3: Geographic Variation: Reduced Form Effect of Head Start in County (robustness)

	(1) Index	(2) Index	(3) Index	(4) Index	(5) Index	(6) Index	(7) Index	(8) Index	(9) Index
Grandmother < High School	0.279***	0.317***	0.282***	0.331***	0.237**	0.194*	0.412***	0.295	0.359***
	(0.080)	(0.073)	(0.074)	(0.079)	(0.102)	(0.117)	(0.102)	(0.229)	(0.077)
Observations	1,684	1,683	1,683	1,683	1,683	1,683	779	1,683	1,290
Grandmother ≤ High School	0.139**	0.153**	0.146**	0.157**	0.153**	0.162**	0.144*	0.092	0.191***
	(0.063)	(0.060)	(0.058)	(0.064)	(0.078)	(0.075)	(0.083)	(0.160)	(0.064)
Observations	2,730	2,729	2,729	2,729	2,729	2,729	1,452	2,729	2,082
Covariates	N	Y	Y	Y	Y	Y	Y	Y	Y
Birth County Chars. (1960) x Trend	N	N	Y	N	N	N	N	N	N
WOP Measures	N	N	N	Y	N	N	N	N	N
Birth County Trends	N	N	N	N	Y	N	N	N	N
State by Year Fixed Effects	N	N	N	N	N	Y	N	N	N
Exclude South	N	N	N	N	N	N	Y	N	N
Family Fixed Effects (mother)	N	N	N	N	N	N	N	Y	N
Head Start by 1970	N	N	N	N	N	N	N	N	Y

Note: Each column represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's county of birth. The dependent variable is the index of second generation outcomes as indicated by the column titles. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. In addition to year of birth and county of birth fixed effects, controls include child's gender, age, and age squared, as well as mother's birth order and race. Column (2) contains base specification. * ($p<0.10$), **($p<0.05$), ***($p<0.01$).

Table 4: Geographic Variation: Reduced Form Effect of Head Start in County (falsification)

	(1) Teen Parent	(2) Crime	(3) High School	(4) Some College	(5) Index
HS in County	-0.015	0.005	-0.018	0.000	-0.002
	(0.041)	(0.054)	(0.047)	(0.071)	(0.078)
Observations	1,328	1,328	1,327	1,311	1,328
Mean	0.145	0.270	0.882	0.682	0.122

Note: Each column represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's county of birth. The dependent variables are second generation outcomes as indicated by the column titles. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. In addition to year of birth and county of birth fixed effects, controls include child's gender, age, and age squared, as well as mother's birth order and race. Sample restricted to mothers whose mothers had at least a high school degree. * ($p<0.10$), **($p<0.05$), ***($p<0.01$).

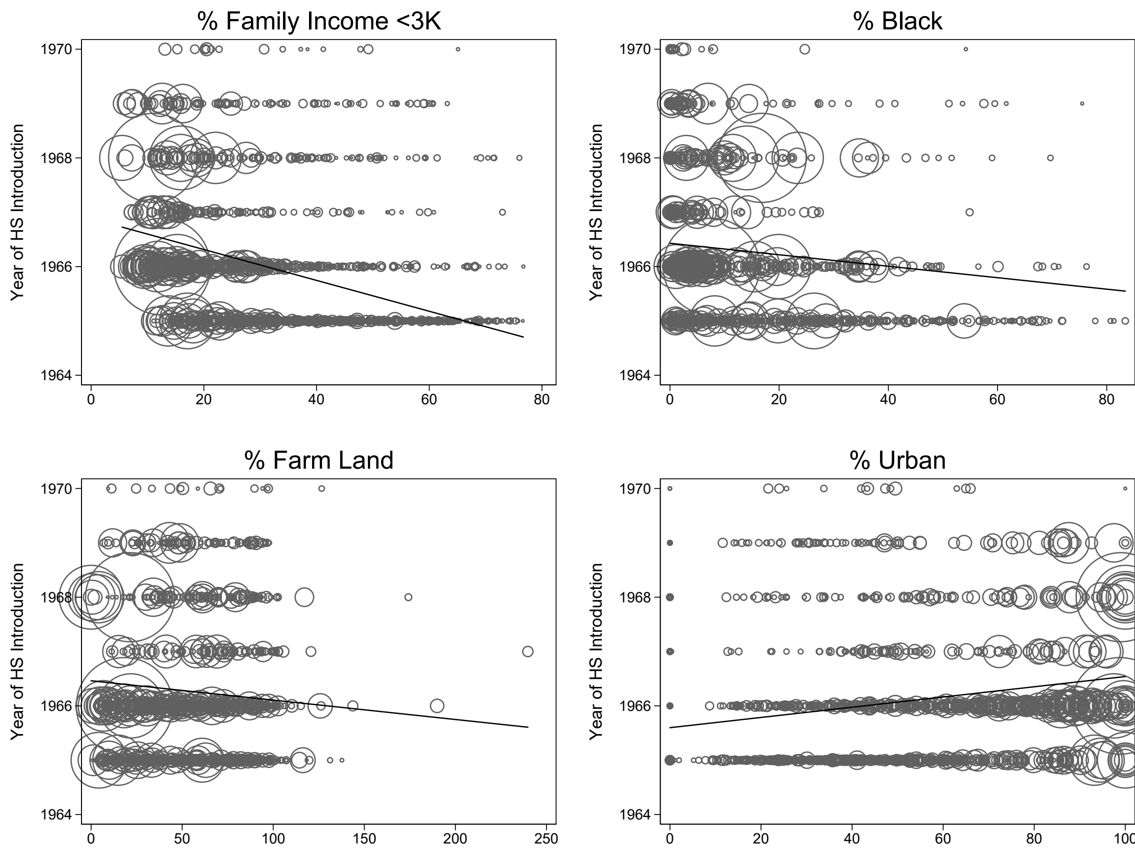
Table 5: Grant-Writing Assistance: Reduced Form Effect of Grant-Writing Assistance

	(1) Teen Parent	(2) Crime	(3) High School	(4) Some College	(5) Index
Grant*Post	-0.161***	-0.163**	0.118*	0.159*	0.345***
	(0.058)	(0.071)	(0.071)	(0.084)	(0.114)
Observations	991	991	991	983	991
Mean	0.211	0.279	0.786	0.546	-0.0654

Note: Each cell represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's county of birth. The dependent variables are second generation outcomes as indicated by the column titles. Coefficient presented for binary variable indicating county being above the 1960 poverty line for grant-writing assistance (59.18) and individual (mother) being born after 1960. In addition to year of birth and county of birth fixed effects, controls include child gender, age, and age squared, as well as mother's birth order and race. * (p<0.10), **(p<0.05), ***(p<0.01).

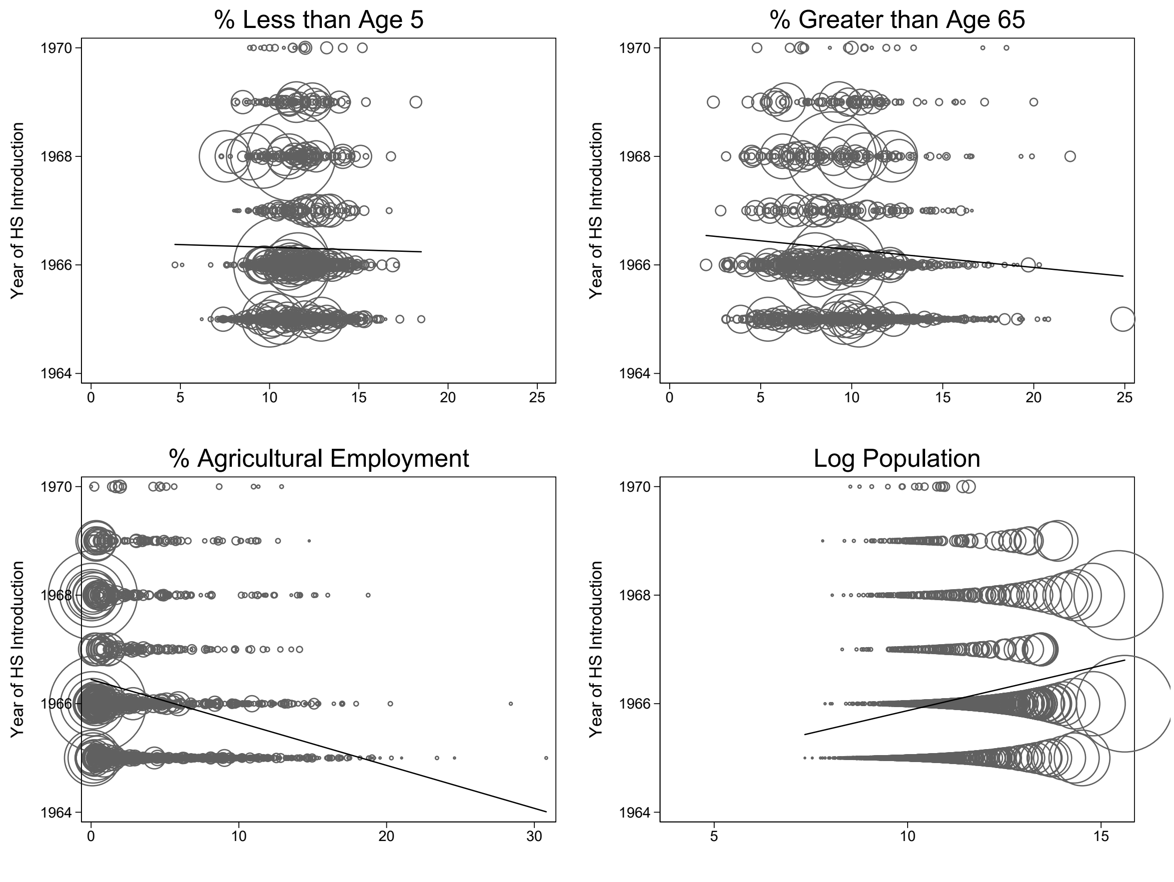
Appendix A: Supplemental Figures and Tables

Figure A1: Exploring Endogeneity of Head Start Adoption Through 1970



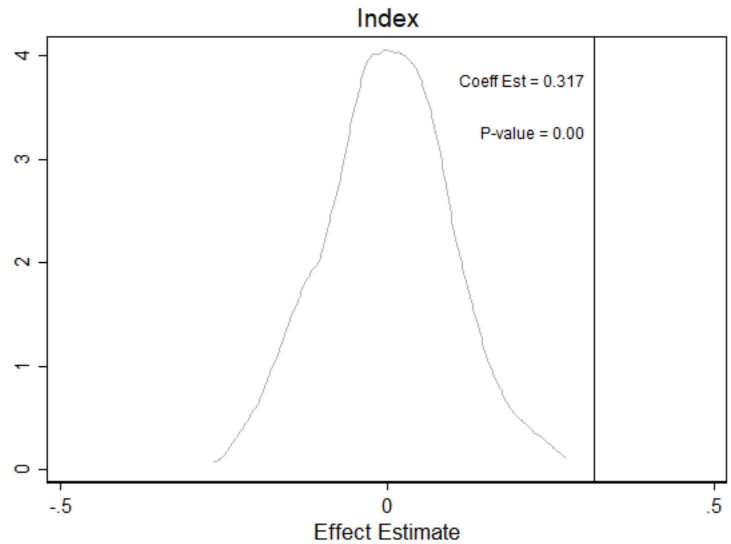
Note: Each scatter plot shows the relationship between baseline (1960) county characteristics and the year of Head Start introduction in that county. Fitted line is from a bivariate regression of year of introduction on the baseline characteristic, weighted by county population in 1960. The data are at the county-level and contain the 1,471 counties in the United States for which the relevant information was available and Head Start was adopted by 1970. Bubble size represents the population of each county in 1960.

Figure A2: Exploring Endogeneity of Head Start Adoption Timing Through 1970



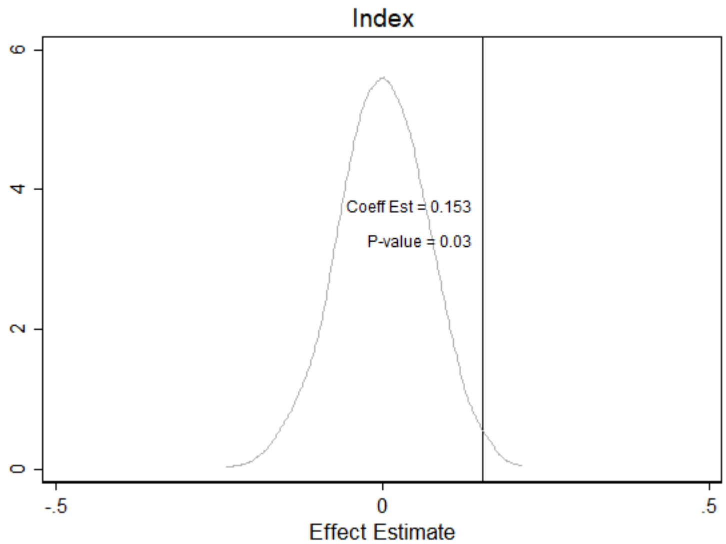
Note: Each scatter plot shows the relationship between baseline (1960) county characteristics and the year of Head Start introduction in that county. Fitted line is from a bivariate regression of year of introduction on the baseline characteristic, weighted by county population in 1960. The data are at the county-level and contain the 1,471 counties in the United States for which the relevant information was available and Head Start was adopted by 1970. Bubble size represents the population of each county in 1960.

Figure A3: High Impact Sample: Randomization Inference (index)



Note: We estimate our baseline specification for 1000 random assignments of the pattern of timing of the Head Start rollout to counties. The figure plots the smoothed distribution of the resulting coefficient estimates. The vertical line indicates the coefficient estimate using the actual timing of Head Start availability in each county. The "p-value" presented is the two-tailed statistic calculated as the share of coefficient estimates obtained under random assignment of Head Start introduction timing that are larger in absolute magnitude than the estimate using the actual timing of introduction.

Figure A4: Low Impact Sample: Randomization Inference (index)



Note: We estimate our baseline specification for 1000 random assignments of the pattern of timing of the Head Start rollout to counties. The figure plots the smoothed distribution of the resulting coefficient estimates. The vertical line indicates the coefficient estimate using the actual timing of Head Start availability in each county. The "p-value" presented is the two-tailed statistic calculated as the share of coefficient estimates obtained under random assignment of Head Start introduction timing that are larger in absolute magnitude than the estimate using the actual timing of introduction.

Table A1: Predictors of Head Start Adoption Timing

	(1) Year of Head Start Adoption
% Family Income <3K	-0.0259***
	(0.00822)
% Black	-0.00155
	(0.00774)
% Less than Age 5	-0.0611
	(0.0765)
% Greater than Age 65	-0.0411
	(0.0351)
% Agricultural Employment	0.0224
	(0.0202)
Log Population	0.0953
	(0.144)
% Farmland	0.00128
	(0.00257)
% Urban	-0.00216
	(0.00500)
Constant	1,967***
	(2.062)
Observations	1,471
R-squared	0.086
Partial F-test (conditional on % Family Income <3K)	0.567
Prob > F (conditional on % Family Income <3K)	0.783

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Estimates show the relationship between baseline (1960) county characteristics and the year of Head Start introduction in that county, restricting to counties that adopted by 1970. The column represents a single regression, weighted by the county's population in 1960, where the year of adoption is the dependent variable and the various 1960 characteristics listed in rows are the explanatory variables. The data are at the county-level and contain the 1,471 counties in the United States for which the relevant information was available and Head Start was adopted by 1970. Robust standard errors are in parentheses.

Table A2: Geographic Variation: Reduced Form Effect of Head Start in County: High Impact

	(1) Teen Parent	(2) Crime	(3) High School	(4) Some College	(5) Index
HS in County (Any Positive)	-0.070*	-0.114***	0.109	0.172**	0.270***
	(0.036)	(0.041)	(0.070)	(0.076)	(0.097)
HS in County (5th percentile)	-0.026	-0.071*	0.112**	0.115**	0.195**
	(0.030)	(0.039)	(0.050)	(0.058)	(0.076)
HS in County (10th percentile)	-0.078**	-0.150***	0.139***	0.175***	0.317***
	(0.031)	(0.043)	(0.047)	(0.056)	(0.073)
HS in County (15th percentile)	-0.067**	-0.160***	0.094*	0.082	0.232***
	(0.033)	(0.044)	(0.054)	(0.060)	(0.082)
Observations	1,683	1,683	1,683	1,662	1,683
Mean	0.220	0.305	0.768	0.516	-0.113

Note: Sample restricted to children whose maternal grandmothers had less than a high-school education. Each cell represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's county of birth. The dependent variables are second generation outcomes as indicated by the column titles. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth using various percentiles of spending per birth. In addition to year of birth and county of birth fixed effects, controls include child gender, age, and age squared, as well as mother's birth order and race. * ($p<0.10$), **($p<0.05$), ***($p<0.01$).

Table A3: Geographic Variation: Reduced Form Effect of Head Start in County: Low Impact

	(1) Teen Parent	(2) Crime	(3) High School	(4) Some College	(5) Index
HS in County (Any Positive)	-0.020	-0.036	0.063	0.056	0.108
	(0.029)	(0.045)	(0.055)	(0.063)	(0.082)
HS in County (5th percentile)	-0.022	-0.014	0.034	0.008	0.050
	(0.025)	(0.033)	(0.042)	(0.049)	(0.065)
HS in County (10th percentile)	-0.048*	-0.070**	0.071*	0.071	0.153**
	(0.027)	(0.034)	(0.039)	(0.047)	(0.060)
HS in County (15th percentile)	-0.031	-0.069**	0.059	0.031	0.112*
	(0.027)	(0.035)	(0.038)	(0.046)	(0.059)
Observations	2,729	2,729	2,728	2,692	2,692
Mean	0.194	0.296	0.807	0.564	-0.00930

Note: Sample restricted to children whose maternal grandmothers had a high-school education or less. Each cell represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's county of birth. The dependent variables are second generation outcomes as indicated by the column titles. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth using various percentiles of spending per birth. In addition to year of birth and county of birth fixed effects, controls include child gender, age, and age squared, as well as mother's birth order and race. * ($p<0.10$), **($p<0.05$), ***($p<0.01$).

Table A4: Geographic Variation: Head Start Availability and Enrollment

	Grandmother < HS	Grandmother ≤ HS
1994 Retrospective Self-reported Head Start Participation (NLSY 79)		
Head Start in County	0.100**	0.055
	(0.045)	(0.034)
Observations	805	1,374
Mean Participation | Head Start not in County	0.14	0.12
Mean Participation | Head Start in County	0.35	0.30
Fraction Enrolled in Head Start in State (OEO 66 Enrollment Counts)		
Head Start in County	0.287*	0.149*
	(0.150)	(0.077)
Observations	49	49
Implied Mean Fraction Enrolled | Head Start in County	0.58	0.30
Fraction Enrolled in Head Start Nationally (Enrollment Counts (66-69))		
Observations	4	4
Implied Mean Fraction Enrolled | Head Start in County	0.56	0.29

Note: Each cell represents a separate OLS regression. The first panel contains estimates of the effect of Head Start availability on self-reported participation in the NLSY79 (reported in 1994). The dependent variable is the mother's self-reported Head Start status. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. In addition to year of birth and county of birth fixed effects, controls include mother's birth order and race. Sample restricted to mothers whose mothers had less than a high school degree or at most a high school degree. The means are average self-reported Head Start participation levels in counties with and without Head Start availability. The second panel contains estimates from a regression of the state-level share of four-year olds participating in Head Start against the fraction of the state's four-year old population with Head Start availability. State level availability is determined by dividing the number of 4 year-olds born in counties with availability in 1966 (determined using our county-level treatment variable and natality data) by the total number of 4 year-olds born in the state. State-level participation is from OEO 1966 enrollment counts. State-level participation and availability counts are adjusted using statistics on the education levels of the mothers of participants (70% had less than a high school degree and 90% had a high school degree or less) and the mothers of children born during this time period (33% had less than a high school degree and 82% had a high school degree or less). The means are the fraction of children estimated to be enrolled in Head Start, assuming that only children born in counties with Head Start availability are enrolled. The third panel contains similar means using the national level Head Start enrollment data for 1966-1969 combined with natality data. * (p<0.10), **(p<0.05), ***(p<0.01).

Table A5: Geographic Variation: Reduced Form Effect of Head Start in County: High Impact

	(1) Index	(2) Index	(3) Index	(4) Index	(5) Index
HS in County	0.317***	0.317***	0.317***	0.317***	0.317***
SE	(0.073)	(0.073)	n/a	(0.067)	(0.076)
P-Value	0.000	0.000	0.000	0.000	0.000
95% CI	0.173-0.461	0.173-0.461	0.150-0.462	0.185-0.449	0.164-0.470
Observations	1,683	1,683	1,683	1,683	1,683
Mean	-0.113	-0.113	-0.113	-0.113	-0.113
Robust	Y	N	N	N	N
Cluster (county)	N	Y	N	N	N
Wild Cluster Bootstrap (county of birth)	N	N	Y	N	N
Cluster (county and year of birth)	N	N	N	Y	N
Cluster (state of birth)	N	N	N	N	Y

Note: Each cell represents a separate OLS regression. The dependent variable is the second-generation outcome index as indicated by the column titles. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. In addition to year of birth and county of birth fixed effects, controls include child gender, age, and age squared, as well as mother's birth order and race. * ($p<0.10$), **($p<0.05$), ***($p<0.01$).

Table A6: Geographic Variation: Reduced Form Effect of Head Start in County: Low Impact

	(1) Index	(2) Index	(3) Index	(4) Index	(5) Index
HS in County	0.153**	0.153**	0.153**	0.153***	0.153***
SE	(0.060)	(0.060)	n/a	(0.052)	(0.051)
P-Value	0.012	0.012	0.023	0.004	0.005
95% CI	0.035-0.272	0.035-0.272	0.0216-0.278	0.050-0.256	0.050-0.257
Observations	2,729	2,729	2,729	2,729	2,729
Mean	-0.009	-0.009	-0.009	-0.009	-0.009
Robust	Y	N	N	N	N
Cluster (county)	N	Y	N	N	N
Wild Cluster Bootstrap (county of birth)	N	N	Y	N	N
Cluster (county and year of birth)	N	N	N	Y	N
Cluster (state of birth)	N	N	N	N	Y

Note: Each cell represents a separate OLS regression. The dependent variables are second generation outcomes as indicated by the column titles. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. In addition to year of birth and county of birth fixed effects, controls include child gender, age, and age squared, as well as mother's birth order and race. * ($p<0.10$), **($p<0.05$), ***($p<0.01$).

Table A7: Geographic Variation Balance Checks: Head Start Availability and Observables

	(1) First born	(2) Black	(3) HH in Poverty (78)	(4) Grandmother ≤ High School	(5) Grandmother < High School	(6) Child Age (12)	(7) Male Child	(8) Predicted Index
Full Sample	0.058	0.048	0.017	0.037	0.010	0.132	0.008	-0.021
	(0.042)	(0.029)	(0.048)	(0.031)	(0.038)	(0.250)	(0.035)	(0.018)
Observations	1,638	1,647	1,558	1,549	1,549	3,216	3,217	2,866
Mean	0.504	0.348	0.330	0.896	0.538	27.15	0.497	-0.0275
Grandmother < HS	0.024	0.020	-0.011	NA	NA	0.173	0.128***	-0.028
	(0.063)	(0.045)	(0.079)			(0.348)	(0.044)	(0.024)
Observations	809	811	778	811	811	1,683	1,684	1,617
Mean	0.478	0.378	0.447	1	1	27.72	0.501	-0.122
Grandmother ≤ HS	0.062	0.042	0.014	NA	-0.003	0.281	0.038	-0.019
	(0.048)	(0.033)	(0.053)		(0.043)	(0.261)	(0.035)	(0.018)
Observations	1,380	1,385	1,321	1,385	1,385	2,729	2,730	2,611
Mean	0.504	0.354	0.347	1	0.601	27.27	0.498	-0.0535

Note: Each cell represents a separate OLS regression with robust standard errors in parentheses, clustered on county of birth. The dependent variable is indicated by the column title. To generate the predicted index we first regress our second-generation index of adult outcomes on all of the maternal and child characteristics in the table. Coefficients presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. Regressions also include controls for year of birth and county of birth fixed effects. * ($p<0.10$), **($p<0.05$), ***($p<0.01$).

Table A8: Geographic Variation: Head Start Availability and Self-reported Participation (Falsification)

	Grandmother $\geq$ High School
HS in County	-0.005
	(0.053)
Observations	735
Mean	0.15

Note: The column represents a separate OLS regression with robust standard errors in parentheses, clustered on county of birth. The dependent variable is the mother's self-reported Head Start status. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. In addition to year of birth and county of birth fixed effects, controls include mother's birth order and race. Sample restricted to mothers whose mothers had at least a high school degree. * (p<0.10), **(p<0.05), ***(p<0.01).

Table A9: Effect of Grant-Writing Assistance on Self-reported Head Start Participation

	(1)	(2)	(3)	(4)	(5)
Bandwidth	+/- 10	+/- 15	+/- 20	+/- 25	+/- 30
Grant*Post	0.276*	0.318**	0.253*	0.160	0.213
	(0.149)	(0.137)	(0.151)	(0.155)	(0.143)
Observations	232	335	438	512	735
Mean	0.484	0.481	0.481	0.481	0.481

Note: Each cell represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's county of birth. The dependent variable is first-generation self-reported Head Start participation. Coefficient presented for binary variable indicating a county being above the 1960 poverty line for grant-writing assistance (59.18) and individual (mother) being born after 1960. Columns vary the bandwidth in percent poverty around this cutoff. In addition to year of birth and county of birth fixed effects, controls include mother's birth order and race. * (p<0.10), **(p<0.05), ***(p<0.01).

Table A10: Effect of Grant-Writing Assistance on Child Index: Robustness to Specification Choice

	(1) Index	(2) Index	(3) Index	(4) Index	(5) Index
Grant*Post	0.330***	0.345***	0.324**	0.348***	0.457***
	(0.125)	(0.114)	(0.159)	(0.105)	(0.166)
Observations	992	991	991	991	736
Mean	-0.0654	-0.0654	-0.0654	-0.0654	-0.0654
Covariates	N	Y	Y	Y	Y
Birth County Chars. (1960) x Trend	N	N	Y	N	N
WOP Measures	N	N	N	Y	N
Exclude 1961	N	N	N	N	Y

Note: Sample restricted to counties with 1960 poverty rates within 20 percentage points of the grant-writing cutoff. Each cell represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's county of birth. The dependent variable is the second generation outcome index. Coefficient presented for binary variable indicating county being above the 1960 poverty line for grant-writing assistance (59.18) and individual (mother) being born after 1960. Columns vary the specification choice as indicated in the rows at the bottom of the table. In addition to year of birth and county of birth fixed effects, controls include child gender, age, and age squared, as well as mother's birth order and race. * ($p<0.10$), **($p<0.05$), ***($p<0.01$).

Table A11: Effect of Grant-Writing Assistance on Child Index: Robustness to Bandwidth Choice

	(1)	(2)	(3)	(4)	(5)
Bandwidth	+/- 10	+/- 15	+/- 20	+/- 25	+/- 30
Grant*Post	0.214	0.289**	0.345***	0.314***	0.276**
	(0.132)	(0.119)	(0.114)	(0.114)	(0.110)
Observations	538	774	991	1,146	1,597
Mean	-0.0709	-0.0500	-0.0654	-0.0478	-0.0296

Note: Each cell represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's county of birth. The dependent variable is the second generation outcome index. Coefficient presented for binary variable indicating county being above the 1960 poverty line for grant-writing assistance (59.18) and individual (mother) being born after 1960. Columns vary the bandwidth in percent poverty around this cutoff. In addition to year of birth and county of birth fixed effects, controls include child gender, age, and age squared, as well as mother's birth order and race. * (p<0.10), **(p<0.05), ***(p<0.01).

Table A12: Geographic Variation: Reduced Form Effect of Head Start in County: Poor Families

	(1)	(2)	(3)	(4)	(5)
Panel A: Mom Effects:					
	Teen Mom	High School	Some College	High Grade	Perm Inc
HS in County	0.031	0.072	0.144**	0.652**	0.114
	(0.096)	(0.082)	(0.065)	(0.258)	(0.093)
Observations	485	484	484	484	485
Mean	0.569	0.750	0.325	12.29	-0.411
Panel B: Interg Effects:					
	Teen Parent	Crime	High School	Some College	Index
HS in County	-0.131**	-0.125**	0.180***	0.123*	0.338***
	(0.051)	(0.056)	(0.063)	(0.068)	(0.103)
Observations	1,068	1,068	1,068	1,054	1,068
Mean	0.271	0.333	0.732	0.482	-0.201

Note: Sample restricted to mothers whose families were in poverty in 1978. Each cell represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's county of birth. The dependent variables are first and second generation outcomes as indicated by the column titles. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. In addition to year of birth and county of birth fixed effects, controls include child gender, age, and age squared, as well as mother's birth order and race. * (p<0.10), **(p<0.05), ***(p<0.01).

Table A13: Geographic Variation: IV Effect of Head Start Spending in County: Poor Families

	(1)	(2)	(3)	(4)	(5)
Panel A: Mom Effects:					
	Teen Mom	High School	Some College	High Grade	Perm Inc
HS in County	0.024	0.075	0.148**	0.659***	0.124
	(0.065)	(0.074)	(0.072)	(0.252)	(0.091)
Observations	482	481	481	481	482
Mean	0.569	0.750	0.325	12.29	-0.411
Panel B: Interg Effects:					
	Teen Parent	Crime	High School	Some College	Index
HS in County	-0.125**	-0.120**	0.173***	0.120**	0.325***
	(0.050)	(0.054)	(0.057)	(0.059)	(0.093)
Observations	1,061	1,061	1,061	1,047	1,061
Mean	0.271	0.333	0.732	0.482	-0.201

Note: Sample restricted to mothers who were in poverty in 1978. Each cell represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's county of birth. The dependent variables are second generation outcomes as indicated by the column titles. Coefficient presented for IV estimate of Head Start spending per child (scaled to average program size) in mother's birth county 4 or 5 years after the year of mother's birth. Spending is instrumented using Head Start availability. In addition to year of birth and county of birth fixed effects, controls include child gender, age, and age squared, as well as mother's birth order and race. * ($p<0.10$), **($p<0.05$), ***($p<0.01$).

Table A14: Geographic Variation: Reduced Form Effect of Head Start in County (Mother)

	(1) Teen Parent	(2) High School	(3) Some College	(4) Highest Grade	(5) Perm Income (std)	(6) Number of Kids
Grandmother $<$ High School	0.048	0.104**	0.091*	0.543**	0.139	-0.162
	(0.072)	(0.049)	(0.053)	(0.215)	(0.089)	(0.133)
Observations	809	806	806	806	809	809
Mean	0.515	0.795	0.319	12.35	-0.217	2.742
Grandmother $\leq$ High School	0.052	0.063*	0.065	0.499***	0.053	0.022
	(0.049)	(0.036)	(0.044)	(0.188)	(0.070)	(0.101)
Observations	1,380	1,371	1,371	1,371	1,380	1,380
Mean	0.431	0.842	0.397	12.75	-0.0749	2.624

Note: Each column represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's county of birth. The dependent variables are first-generation (mother) outcomes indicated by the column titles. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. In addition to year of birth and county of birth fixed effects, controls include mother's birth order and race. * (p<0.10), **(p<0.05), ***(p<0.01).

Table A15: Grant-Writing Assistance: Reduced Form Effect of Grant-Writing Assistance

	(1) Teen Mom	(2) High School	(3) Some College	(4) High Grade	(5) Perm Inc
Grant*Post	-0.083	0.060	0.051	0.338	0.072
	(0.116)	(0.073)	(0.072)	(0.386)	(0.072)
Observations	443	442	442	442	443
Mean	0.485	0.835	0.419	12.92	-0.197

Note: Each cell represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's county of birth. The dependent variables are first generation outcomes as indicated by the column titles. Coefficient presented for binary variable indicating county being above the 1960 poverty line for grant-writing assistance (59.18) and individual (mother) being born after 1960. In addition to year of birth and county of birth fixed effects, controls include birth order and race. * ($p<0.10$), **($p<0.05$), ***($p<0.01$).

Table A16: Head Start First-Generation Effect Size Comparison

Study	Identification	Subgroup, if applicable	High School Completion	Implied ToT	Years of Education	Implied ToT	Income/ Wages - % change	Implied ToT	Long-term Index*	Implied ToT
Johnson & Jackson (forthcoming)	geo variation spending		0.106 (0.029)	0.162	0.327 (0.084)	0.5	10%	15%	–	–
Johnson & Jackson (forthcoming)	geo variation spending IV		0.173 (0.104)	0.265	0.954 (0.513)	1.46	15%	23%	–	–
Thompson (2018)	geo variation spending		0.012 ^{ns} (0.010)	0.049	0.125 (0.051)	0.512	4-5%	16-20%	0.081 (0.023)	0.332
Barr & Gibbs (this paper) - 1st gen	geo variation availability	females	0.104 (0.049)	0.173	0.543 (0.215)	0.905	12% ^{ns}	20%	0.155 (0.084)	0.258
Carneiro & Ginja (2014)	RD in eligibility	males	0.282 (0.189)	0.379	–	–	–	–	0.194 ^{ns} (0.139)	0.261
Ludwig & Miller (2007)	RD in grant-writing assistance		0.117 ^{ns} (0.080)	0.68	0.584 (0.358)	3.395	–	–	–	–
Barr & Gibbs (this paper) - 1st gen	DD in grant-writing assistance	females	0.060 ^{ns} (0.073)	0.086	0.338 ^{ns} (0.386)	0.483	6% ^{ns}	9%	0.103 ^{ns} (0.122)	0.147
Deming (2009)	family fixed effects		0.086 (0.031)	0.086	–	–	–	–	0.228 (0.072)	0.228
Garces et al. (2002)	family fixed effects		0.037 ^{ns} (0.053)	0.037	–	–	19% ^{ns}	19%	–	–
Garces et al. (2002)	family fixed effects	whites	0.203 (0.098)	0.203	–	–	57% ^{ns}	57%	–	–
Garces et al. (2002)	family fixed effects	whites, mothers ≤ high school	0.283 (0.119)	0.283	–	–	100%	100%	–	–

Notes: When estimate is not an average treatment effect on the treated, we use the authors' estimated first stage, or participation rate, to generate an implied treatment on the treated (ToT). It is important to note that nearly all of these authors caution the reader in interpreting the scaled estimates as TOTs due to misclassification error in self-reported participation and related issues. Specifically, for Johnson and Jackson (forthcoming), we follow the procedure outlined in Section VI.A in which they discuss a comparison of their effects of Head Start availability with previous estimates of the effects of participation. For Thompson (2018), we scale estimates by 1/.244, the first-stage estimate implied by a cross-sectional regression of state-level spending on state-level availability in a single year. Estimates from Barr and Gibbs (current paper) are scaled by 1/.6 in the diff-in-diff approach leveraging geographic variation in Head Start availability and by 1/.7 in the diff-in-diff approach using a discontinuity in grant-writing assistance. For Carneiro and Ginja (2014), we scale estimates by 1/.744, which is the authors' first-stage estimate for male survey respondents in the relevant age group, corresponding to the reduced-form estimates. For Ludwig and Miller (2007), we scale the NELS estimates for the authors' preferred specification by the first-stage they estimate in an equivalent specification. The bottom panel reports estimates of the effect of Head Start participation and thus are not scaled. Standard errors in parentheses. * Long-term index contains different measures across studies. "ns" denotes a result that is not significantly different from zero at the 10% level.

.

Appendix B: Recall Error and the Relationship Between Head Start Availability and Retrospective Self-Reported Participation

Table A4 illustrates that our measure of Head Start availability predicts both self-reported Head Start participation and state-level participation rates derived from administrative Head Start enrollment data. The top panel presents estimates of the effect of Head Start availability on self-reported participation. When a program is available in a county four or five years after a mother's year of birth, the mother in our high impact sample is 10 percentage points more likely *to report* having participated in Head Start as a child. Given the self-reported and retrospective nature of the Head Start participation variable, we expect there is considerable misreporting. As has been established in the literature, measurement error in a binary variable is necessarily non-classical and will thus result in a downward biased estimate of the relationship between Head Start availability and participation. For the purposes of explanation, consider the following specification of the true relationship between Head Start participation HS and Head Start availability $HSavail$:

$$HS = \beta_1 X + \beta_2 HSavail + \varepsilon \tag{3}$$

Here, HS is an individual's true Head Start participation, X is a vector of other observables (in practice this includes year and county of birth fixed effects), and $HSavail$ is an indicator for the availability of Head Start. Instead of observing HS, we observe self-reported Head Start participation HS_{sr}, which is HS measured with error. Given that HS is a binary variable, misclassification will necessarily lead to non-classical measurement error and a downward biased estimate of β_2 (Kane et al. 1999, Aigner 1973).

The extent of this bias has been previously derived. Let $m_1 = Pr(HS_{sr} = 0|HS = 1)$ represent the fraction of individuals who self-report not participating in Head Start when they did participate. Similarly, let $m_0 = Pr(HS_{sr} = 1|HS = 0)$ represent the fraction of individuals who self-report participating in Head Start when they did not participate. Under the assumption of constant rates of misclassification (m_1 and m_0), $\beta_2^{OLS} = (1 - m_0 - m_1) * \beta_2$.

Determining the extent of the bias depends on knowledge of the misclassification rates m_1

and m_0. Because we do not have measures of these misclassification rates, we take two approaches to guide the selection of reasonable choices for m_1 and m_0. First, we use the available administrative and self-reported participation data to generate estimates of the misclassification rates. Second, we draw upon the existing literature on retrospective recall rates to inform our selection of misclassification rates.

Data-driven Estimates of Misclassification Rates

For the data driven exercise, we draw upon the following equations:

$$(\overline{HS}_{sr}|HSavail = 1) = Pr(HS_{sr} = 1|HS = 0) * Pr(HS = 0|HSavail = 1) + Pr(HS_{sr} = 1|HS = 1) * Pr(HS = 1|HSavail = 1)$$

and,

$$(\overline{HS}_{sr}|HSavail = 0) = Pr(HS_{sr} = 1|HS = 0) * Pr(HS = 0|HSavail = 0) + Pr(HS_{sr} = 1|HS = 1) * Pr(HS = 1|HSavail = 0)$$

From the data, we can calculate the left hand side of both equations. If we assume that Head Start participation occurs entirely in counties with Head Start funding, we can set $Pr(HS = 1|HSavail = 0)$ to 0 and $Pr(HS = 0|HSavail = 0)$ to 1, and use the administrative data on Head Start participation levels to provide estimates of $Pr(HS = 0|HSavail = 1)$ and $Pr(HS = 1|HSavail = 1)$. This simplifies both equations to:

$$(\overline{HS}_{sr}|HSavail = 1) = Pr(HS_{sr} = 1|HS = 0) * Pr(HS = 0|HSavail = 1) + Pr(HS_{sr} = 1|HS = 1) * Pr(HS = 1|HSavail = 1)$$

and,

$$(\overline{HS}_{sr}|HSavail = 0) = Pr(HS_{sr} = 1|HS = 0) = m_0$$

We now have two equations and two unknowns. From the data on our high impact sample, we know that $(\overline{HS}_{sr}|HSavail = 0) = m_0$ is somewhere between 0.20 and 0.24.[43] From the data, we know that $(\overline{HS}_{sr}|HSavail = 1)$ is equal to 0.35. Plugging in our population-level averages for Head Start participation in counties with Head Start availability (0.58) and solving provides bounds for m_1 of 0.54 to 0.57. Combining this with the equation for $(\overline{HS}_{sr}|HSavail = 1)$, indicates that the relationship between Head Start availability and participation is likely biased down by a factor of 3.9 to 5.3. In other words, under a reasonable set of assumptions, the true magnitude of β_1 is closer to 39 to 53 percentage points. The true estimate may be even higher if Head Start availability increases rates of misclassification.

Literature-driven Selection of Misclassification Rates

A second way to guide the selection of reasonable choices for m_1 and m_0 is to draw upon existing literature on retrospective recall rates. While numerous studies in the psychology literature indicate that recall of early life events is often quite poor, there are only a handful of studies that have managed to track individuals over time so as to compare actual or reported events during childhood or adolescence with retrospective information regarding the same events. These studies all indicate the poor reliability of adulthood retrospective accounts of childhood events across a series of outcomes including various forms of childhood abuse, parental divorce or separation, parental chronic unemployment, household legal trouble, household illness or disability, maternal depression, household substance abuse, household incarceration, residence changes, height, weight, injuries, and contacts with the criminal justice system (Henry et al. 1994, Naicker et al. 2017; Reuben et al. 2016). Even parental death, which is recalled far more accurately than these other measures, is imperfectly measured in retrospective reports.

While several studies indicate that retrospective accounts are fraught with misclassification issues, only one contains the necessary underlying data to easily construct measures of $\frac{1}{1-m_0-m_1}$ (Naicker et al. 2017). Naicker and coauthors (2017) compare the reports of adolescents at ages 11,

[43]This is simply the mean participation rate reported in the high-impact sample for counties and birth cohorts that do not have Head Start available when Head Start was available somewhere in the country for that particular birth cohort.

15, and 18 with retrospective reporting by the same individuals at age 23 across a variety of events. The mean and median measures $m_0 + m_1$ are 0.77 to 0.88, suggesting a scaling factor between 4.35 and 8.36. Using these measures would imply a "first-stage" relationship of 43.5 to 83.6 percentage points. It is also important to note that the gap here between event and recall is substantially shorter than the 30 to 35 year gap between the year of likely Head Start participation and the time the retrospective question was asked (1994) in the NLSY79. This suggests that recall may be even worse, and thus the scaling factor should be even higher, in the NLSY. Reuben et al. (2016) ask retrospective questions to a much older group of individuals, age 38, but focuses on recall of a similar set of childhood experiences to those explored by Naicker and colleagues (2017). This gap between the timing of the event and the timing of recall is much closer to that observed in the NLSY79. While the data provided do not allow for the construction of misclassification rates, Cohen's Kappa values, which are constructed in both studies, are similarly very low, indicating poor agreement between retrospective and contemporaneous reporting.

Endnotes

Henry, Bill, Terrie E. Moffitt, Avshalom Caspi, John Langley, and Phil A. Silva, "On the "remembrance of things past": A longitudinal evaluation of the retrospective method," *Psychological Assessment*, 1994, *6*, 92-101.

Naicker, Sara N., Shane A. Norris, Musawenkosi Mabaso, and Linda M. Richter, "An analysis of retrospective and repeat prospective reports of adverse childhood experiences from the South African Birth to Twenty Plus cohort" *PloS ONE*, 2017, *12(7)*, 1-19.

Reuben, Aaron, Terrie E. Moffitt, Avshalom Caspi, Daniel W. Belsky, and et al., "Lest we forget: comparing retrospective and prospective assessments of adverse childhood experiences in the prediction of adult health," *Journal of Child Psychology and Psychiatry*, 2016, *57(10)*, 1103-1112.

Table B1: Geographic Variation: Head Start Availability and Enrollment

	Grandmother < HS	Grandmother ≤ HS
1994 Retrospective Self-reported Head Start Participation (NLSY 79)		
Head Start in County	0.100**	0.055
	(0.045)	(0.034)
Observations	805	1,374
Mean Participation | Head Start not in County	0.14	0.12
Mean Participation | Head Start in County	0.35	0.30
Fraction Enrolled in Head Start in State (OEO 66 Enrollment Counts)		
Head Start in County	0.287*	0.149*
	(0.150)	(0.077)
Observations	49	49
Implied Mean Fraction Enrolled | Head Start in County	0.58	0.30
Fraction Enrolled in Head Start Nationally (Enrollment Counts (66-69))		
Observations	4	4
Implied Mean Fraction Enrolled | Head Start in County	0.56	0.29

Note: Each cell represents a separate OLS regression. The first panel contains estimates of the effect of Head Start availability on self-reported participation in the NLSY79 (reported in 1994). The dependent variable is the mother's self-reported Head Start status. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. In addition to year of birth and county of birth fixed effects, controls include mother's birth order and race. Sample restricted to mothers whose mothers had less than a high school degree or at most a high school degree. The means are average self-reported Head Start participation levels in counties with and without Head Start availability. The second panel contains estimates from a regression of the state-level share of four-year olds participating in Head Start against the fraction of the state's four-year old population with Head Start availability. State level availability is determined by dividing the number of 4 year-olds born in counties with availability in 1966 (determined using our county-level treatment variable and natality data) by the total number of 4 year-olds born in the state. State-level participation and availability counts are adjusted using statistics on the education levels of the mothers of participants (70% had less than a high school degree and 90% had a high school degree or less) and the mothers of children born during this time period (33% had less than a high school degree and 82% had a high school degree or less). The means are the fraction of children estimated to be enrolled in Head Start, assuming that only children born in counties with Head Start availability are enrolled. The third panel contains similar means using the national level Head Start enrollment data for 1966-1969 combined with natality data. * ($p<0.10$), **($p<0.05$), ***($p<0.01$).

Appendix C: Data Appendix

To generate measures of Head Start exposure in the late 1960s, we compile data from the National Archives and Records Administration (NARA) files on the Office of Economic Opportunity's Head Start grant funding (National Archives, n.d.). We employ two data sources, covering different spans of time, to construct county-level measures of Head Start activity.

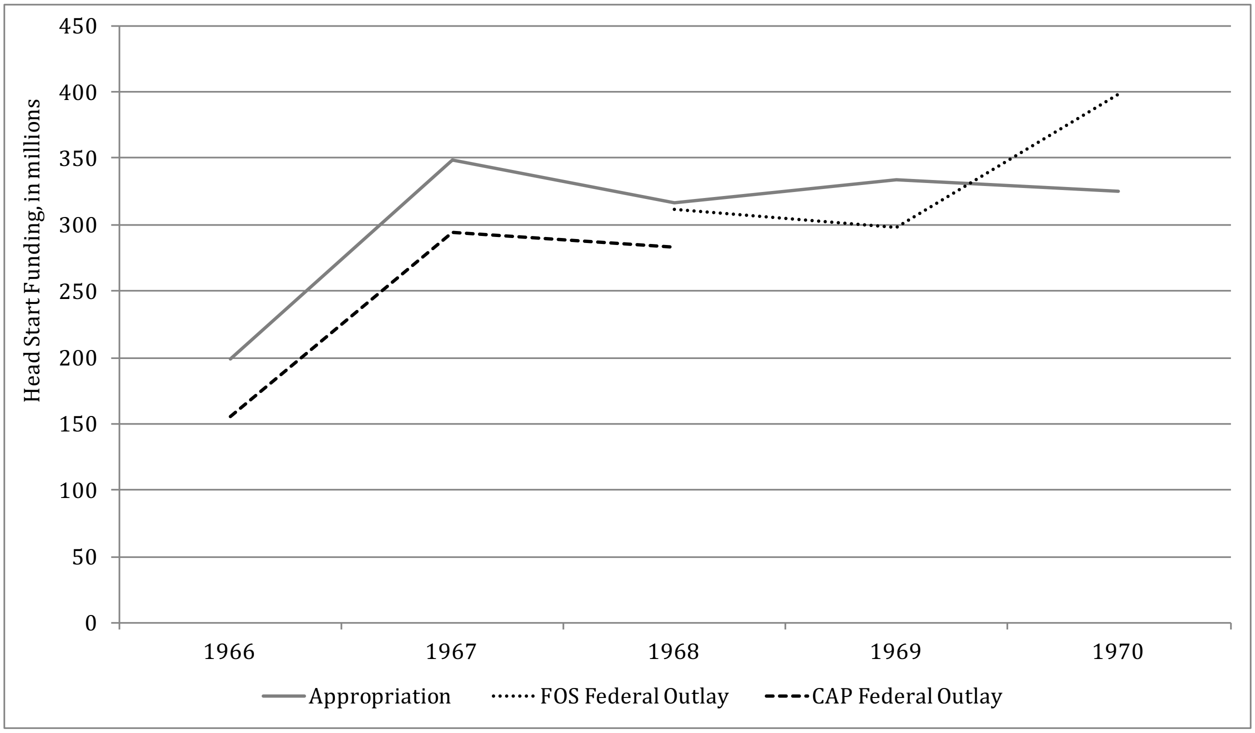
Community Action Program (CAP) Records, 1966–1968

The first data source is Records About Community Action Program (CAP) Grants and Grantees, NARA Record Group 381 (National Archives identifier 604417). These CAP files, created by the Community Services Administration, document the timeframe July 1964 through September 1981. The CAP records consist of two files, the Grantee Organization Master Files which provide information about the grantees receiving funds through CAP and the Funded Program Account Master Files which document specific grant actions. For the purposes of creating county-level funding amounts, we use the Funded Program Account Master Files which contain state and county identifiers.

We retain only records for which the grant action is Initial Application, Supplemental Funding, or Refunding at End of Program Year or for which the grant action field is blank. We exclude, therefore, grant actions Deobligation, Extension, Transfer In, and Transfer Out.

These files contain grant actions on a variety of CAP-funded poverty programs, including job training, housing services, health services, and community development. For this reason, we extract files with certain search terms in the project description field: Head Start, Headstart, child dev, preschool, pre-school, early childhood, HS child, and FYHS. We do not use the terms child care, daycare, or family care center though they appear in the project description field because these records often captured child care associated with job training or postsecondary education programs. These filters result in retaining only a subset of grant actions for each grantee number. Once we have the domain of Head Start programs, we aggregate federal funding to the county-by-FY level using state and county geographic codes. Notably, we drop nonnumeric characters from the funding amount field when they appear (always at the end of the field), assuming that these are placeholders for an input mask.

Figure C1: Crosswalk of Head Start Funding Data Sources



Notes: Funding amounts in current dollars compiled from the National Archives Records Administration, Records of the Community Services Administration. Appropriations (in current dollars) and enrollment numbers collected from the U.S. Department of Health and Human Services, Administration for Children and Families, Office of Head Start.

Because Head Start data is missing from the electronic files for FY 1965 (also documented in Bailey and Duquette's (2014) data appendix), we construct CAP records for FY 1966, 1967, and 1968. After 1968, the search terms we employ largely drop out of the project description field and no longer appear after 1972. While we undercount total Head Start grant funding relative to published federal appropriations (Office of Head Start 2015), the pattern of funding levels and changes across these three years tracks well, as displayed in Figure C1.

Federal Outlays System (FOS) Files, 1968–1980

The second data source is the Federal Outlays System (FOS) Files, also NARA Record Group 381 (National Archives identifier 599052). These records were collected from July 1967 through September 1980, also by the Community Services Administration. The files contain data on Federal Executive Branch outlays and include four files for each fiscal year: 1) a County/State Master File, 2) a City Master File, 3) a Geographic Table File, and 4) a Program Appropriations, Functions, and Agency Table File. The County/State Master Files for each year contain information on programs and outlays with state and county identifiers. We compile these records with the Program Appropriations, Functions, and Agency Table File across the years and again employ search terms to narrow to Head Start programs.

These files contain a variety of program types, so we search in the program title field for the following terms: Head Start, Headstart, child dev, early childhood, HS child, FYHS, and follow-thru program (OEO). Some terms are employed to align with Ludwig and Miller's (2007) file creation process. Terms related to preschool in the FOS files were excluded as they captured unrelated school-age programs. Notably, we included these terms in the CAP files, described previously, because of the narrower focus of those records to CAP-funded anti-poverty programs in the earliest years of the War on Poverty. Records containing the following terms in the title field (appearing in conjunction with child dev or early childhood) were dropped: handicapped, handic, child abuse prev, and child welfare. In addition, 78 observations were deleted because they were missing state identifiers. Funding outlays were then aggregated to the county-by-FY level.

For FY 1968, we have both CAP and FOS records. We retain the union of these files, i.e., if a county has positive funding in either database, they are coded as a Head Start grant-receiving county. We employ a variety of approaches to reconciling recorded grant amounts if a county appears in both CAP and FOS records. We take the maximum funding amount recorded in that year, the minimum amount, and the average when a county appears in both sources in FY 1968.