Breaking the Cycle? Intergenerational Effects of an Antipoverty Program in Early Childhood

Andrew Barr

Texas A&M University

Chloe R. 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 antipoverty programs actually disrupt the cycle of bad outcomes. We leverage the rollout of the United States's largest early-childhood program, Head Start, 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. These effects correspond to an estimated increase in discounted second-generation wages of 6%–11%, depending on specification. Exploration of earlier outcomes suggests an important role for changes in parenting behavior and potential noncognitive channels.

I. Introduction

> I believe this response reflects a realistic and a wholesome awakening in America. It shows that we are recognizing that poverty

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 AE00102 from the US Department of Health and Human Services, Office of the Assistant Secretary

Electronically published October 21, 2022

Journal of Political Economy, volume 130, number 12, December 2022.

https://doi.org/10.1086/720764

> 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 household headship, government assistance receipt, and risky behavior are high.[1] Family, school, and neighborhood contexts collectively shape children's trajectories and generate correspondence between their parents' outcomes and their own. The linkages are particularly acute for minorities, potentially contributing to the early emergence and persistence of achievement gaps.[2]

Societal investments in human capital formation may disrupt the transmission of poverty across generations by increasing educational attainment and labor market attachment and decreasing risky behavior. Early childhood, in particular, is a critical developmental period and opportunity for effective intervention. Multiple studies indicate that interventions in the preschool and early school years can have substantial effects on educational attainment, labor market success, and other measures of health and well-being.[3] Yet we know almost nothing about whether the benefits

for Planning and Analysis. We thank four anonymous reviewers for their time and efforts, Maria Rosales-Rueda and Terri Sabol for helpful discussant comments, and numerous seminar participants. We also thank Martha Bailey, Marianne Bitler, Jorge Luis García, Robert Kaestner, Jason Lindo, Jens Ludwig, Bhash Mazumder, Douglas Miller, and Nikolas Mittag for insightful conversations. Jiee Zhong, Yuxin Chen, 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 US Department of Health and Human Services. Replication files are available in a zip file. This paper was edited by James J. Heckman.

[1] Recent estimates suggest intergenerational correlations in income of 0.3–0.6 (Solon 1999; Mazumder 2005; Black and Devereux 2011), in education levels of 0.4–0.5 (Hertz et al. 2008), in female headship of 0.2 (Page 2004), and in welfare use of 0.3 (Page 2004). Duncan et al. (2005) review and provide 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 magnitude, evidence shows that cognitive test-score gaps by race and ethnicity exist at formal school entry and remain throughout the schooling years (Fryer and Levitt 2004, 2006; Murnane et al. 2006). While race/ethnicity achievement gaps have narrowed in recent decades, they remain pronounced across socioeconomic groups (Reardon and Portilla 2016).

[3] Long-term evidence from the Abecedarian Project, the Perry Preschool Project, Head Start, and the Project STAR (Student/Teacher Achievement Ratio) class-size reduction intervention suggests large positive effects on participants (Garces, Thomas, and Currie 2002; Schweinhart et al. 2005; Deming 2009; Chetty et al. 2011; Dynarski, Hyman and Schanzenbach 2013; Heckman, Pinto, and Savelyev 2013; Campbell et al. 2014).

carry over to the next generation. In other words, can needs-targeted early-childhood programs truly break the cycle of poverty?

Examining the Head Start program, we provide the first evidence of the intergenerational effects of a US-based early-childhood intervention implemented at scale. Head Start, funded and administered through the US Department of Health and Human Services, has been an integral part of the early-childhood conversation since it began in 1965. As the United States's largest preschool program, Head Start enrollment has grown from 400,000 to nearly a million participants today. Across multiple data sets and study designs, quasi-experimental evidence consistently indicates that Head Start yields long-term benefits, particularly for its early cohorts (Garces, Thomas, and Currie 2002; Ludwig and Miller 2007; Deming 2009; Carneiro and Ginja 2014).[4]

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

We find that Head Start availability significantly affects these measures and translates to an increase in a net present value (NPV) of wages summary measure of 6%–11%, depending on specification. Our estimates are robust to the inclusion 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. The legitimacy of the geographic rollout strategy is bolstered by estimates demonstrating no relationship between Head Start availability and outcomes for children unlikely to have been eligible.

Our findings are further supported by a second strategy that leverages variation in Head Start exposure generated by grant-writing assistance provided to the country's poorest 300 counties (Cort et al. 1966; Ludwig and Miller 2007). Because funding was contingent on a successful proposal, grant-writing assistance resulted in higher levels of funding and participation in these counties. Using a difference-in-differences approach,

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

[5] See app. B for details on the NARA data.

we compare the difference in outcomes between children of women born too early for Head Start exposure (before 1961) and children of women born later, across counties receiving and not receiving grant-writing assistance. We find large improvements in second-generation outcomes, consistent with our primary strategy.

To better understand how intergenerational effects develop, we investigate intermediate outcomes in the second generation's early life and adolescence. Overall, we find significant changes in both the second generation's home life (improved parenting behaviors) and early schooling (greater preschool participation). The changes lead to persistent developmental benefits and improved later-life well-being.

Though the confidence intervals around our estimates are relatively wide, we consistently find, across designs and specifications, that increased Head Start exposure for mothers leads to significant improvements in their children's outcomes. We note, however, that there is inevitably some uncertainty regarding the randomness and measurement of our identifying variation. Because the evidence is consistent across sensitivity analyses of both the measurement of Head Start availability and the conditional randomness of program rollout, we conclude that the findings are highly suggestive of intergenerational Head Start effects during this early rollout period. Our findings, in conjunction with other recent work, indicate that societal investments in early-childhood development likely can disrupt the intergenerational transmission of the effects of poverty.

II. A Path out of Poverty

Substantial evidence documents the path dependency of socioeconomic status. At each developmental stage from infancy through adulthood, children from the highest family income quintile are 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 successfully meeting benchmarks from early childhood to adolescence are more likely to attain a middle-class existence than those missing them (Sawhill, Winship, and Grannis 2012). The effects carry over to the next generation, resulting in high intergenerational correlations in income, educational attainment, and risky behavior.

Evidence implying relatively low rates of intergenerational mobility in the United States has led to considerable interest in why parents' resources, behaviors, and outcomes are strongly related to those of their children (Lee and Solon 2009; Auten, Gee, and Turner 2013; Corak 2013; Chetty et al. 2014). The data requirements necessary to answer questions about intergenerational transfer 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 antipoverty 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). Evidence at the middle and high school levels is mixed, with positive intergenerational effects of additional schooling generated by compulsory schooling changes in the United States and Great Britain in the 1960s and 1970s but no effect in Norway (Black, Devereux, and Salvanes 2005; Oreopoulos, Page, and Stevens 2006; Chevalier 2007). Evidence on the intergenerational spillovers of new school construction in Indonesia suggests that increasing a mother's educational attainment produces improvements in her children's test scores and schooling duration (Akresh, Halim, and Kleemans 2018; Mazumder, Rosales-Rueda, and Triyana 2019).

A substantial body of empirical evidence demonstrates that early-childhood programs in particular can have substantial effects on schooling attainment, labor market success, and other measures of well-being into adulthood (Schweinhart et al. 2005; Deming 2009; Chetty et al. 2011; Dynarski, Hyman, and Schanzenbach 2013; Campbell et al. 2014; Carneiro and Ginja 2014; Thompson 2018; Johnson and Jackson 2019; Bailey, Sun, and Timpe 2021; Anders, Barr, and Smith, forthcoming). Despite this growing literature, we know little about whether these effects persist into the subsequent generation. Rossin-Slater and Wüst (2020) explore the intergenerational impact of a Danish preschool and nurse home-visiting program. The researchers find educational attainment effects in the first generation that persist in the second generation. A recent follow-up of Perry Preschool Project participants documents substantial long-term effects for their children, providing compelling experimental identification of a well-documented and more intensive program's effects (García, Heckman, and Ronda 2021).

Alongside this evidence on early-childhood intervention, our study relies on a quasi-experimental approach to estimate intergenerational effects of a program at scale. We provide some of the first evidence—and to our knowledge the first evidence in a US context—on whether the effects of widely offered early-childhood programs transfer across generations. The challenge in evaluating any program implemented at scale is the level of uncertainty around treatment exogeneity and implementation specifics. We address the associated concerns and demonstrate that our evidence, coupled with other work on the long-term effects of the Head Start program and the long-term and intergenerational effects of model preschool programs, reinforces the importance of early-childhood investments for both those exposed to the programs and their children. The existence of such spillovers suggests that a concerted effort to invest in one generation of impoverished youth could break the cycle of poverty and reduce the need to provide similar services to future generations.

A. *The Evidence on Head Start*

Policy discussions of designing and scaling publicly provided preschool interventions often rely on the Head Start literature. Head Start was an early piece of President Lyndon B. Johnson's War on Poverty; operated by the White House's Office of Economic Opportunity (OEO), it commenced as a summer program in 1965 (Vinovskis 2005). It was soon expanded to a year-round program. While Head Start today is characterized as an early-childhood education program, it was designed as an antipoverty initiative, with significant health and community development components. The initial emphasis was on "preschool"-related services and supports, including nutrition, vaccinations and health care, dental services, and social development (Office of Child Development 1968, 1970; Vinovskis 2005).

Head Start served a decidedly disadvantaged population during the 1960s. Participants' median family income was less than half that of all US families, and approximately 50% of early full-year program participants were Black (Office of Child Development 1968). Between 9% and 17% of families reported having no running water inside the home. Only 5% of mothers reported some postsecondary schooling or further education, with approximately 25% indicating that they had graduated from high school and 65%–70% having less than a high school education. Approximately 25% lived in female-headed households, and between 65% and 70% of participating children's mothers were unemployed (Office of Child Development 1968). Among the general population, few children participated in structured preschool outside the home before entering formal schooling, and kindergarten availability was not universal when Head Start was introduced (Cascio 2009).[6]

While there is some debate about Head Start's short-run effects, prior quasi-experimental studies suggest that the program has had significant long-term effects for children born in the late 1960s through the 1980s.[7] 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 (Garces, Thomas, and Currie 2002; Ludwig and Miller 2007;

[6] In the late 1960s, only 10% of 3- and 4-year-old children participated in a school setting, according to Current Population Survey school enrollment data; the number reached 20% in 1970, driven in part by Head Start (Gibbs, Ludwig, and Miller 2013; Johnson and Jackson 2019).

[7] While the HSIS found initial positive effects on participants' cognitive skills in the mid-2000s, no persistent effects were observed at first- and third-grade follow-ups (Puma et al. 2005, 2010, 2012). Reanalyses of the HSIS data suggest a nuanced picture, with considerable variation in impact by center and more concentrated and persistent benefits for Hispanic children, those with low skills at program entry, and those who would otherwise be in parental or relative care (Bitler, Hoynes, and Domina 2014; Walters 2015; Kline and Walters 2016; Montialoux 2016).

Deming 2009; Carneiro and Ginja 2014), even in the presence of short-term test-score fade-out (Deming 2009). A follow-up to Deming's study finds that the effects persist into later adulthood and affect participants' own parenting practices (Bauer and Schanzenbach 2016).

B. *Potential Pathways for Intergenerational Transmission*

Head Start was conceived as an antipoverty program intended to disrupt the cycle of disadvantage passed through generations (Greenberg 1990; Vinovskis 2005). The designers intended the program to be multigenerational in its service delivery and to prepare young people for transitions to formal schooling (Vinovskis 2005). Head Start's comprehensive view of "school readiness" led to inclusion of supports for healthy child development, including medical and dental screenings, mental health services, nutrition, parent engagement and education, and referrals to other social services (Office of Child Development 1968; Zigler and Valentine 1979; Greenberg 1990). The bundled interventions likely affected the long-run outcomes of Head Start-participating children directly, while also indirectly influencing their quality of schooling, peer groups, and subsequent environments (Johnson and Jackson 2019).

When thinking about the intergenerational spillovers of those first-generation improvements, existing literature documents the importance of a mother's human capital and health for her children's prenatal health, birth outcomes, and early-childhood health (Currie and Moretti 2003; Almond and Currie 2011; East et al. 2019; Miller and Wherry 2019). Growing evidence also suggests that the earliest years of life, when children's brains exhibit developmental plasticity, serve as a critical period in skill development (Knudsen et al. 2006; Heckman 2007; Heckman and Mosso 2014). While Head Start exposure occurs when the first generation is 4 or 5 years old, spillovers potentially affect the second generation prenatally and in the earliest years of childhood. Early-life conditions then have cascading effects on development and may complement other investments and inputs that parents make in their children (Cunha and Heckman 2007; Heckman and Masterov 2007). If an intervention affects multiple inputs to second-generation human capital production and the inputs are complementary, one could observe larger long-term improvements for the children of those directly exposed than for the participants themselves.

C. *Head Start's Beginnings*

Our research capitalizes on plausibly random exposure to Head Start over geography and birth cohorts during program introduction (fig. 1). OEO rolled Head Start out quickly as a featured and politically popular component of the War on Poverty (Zigler and Valentine 1979). The White House

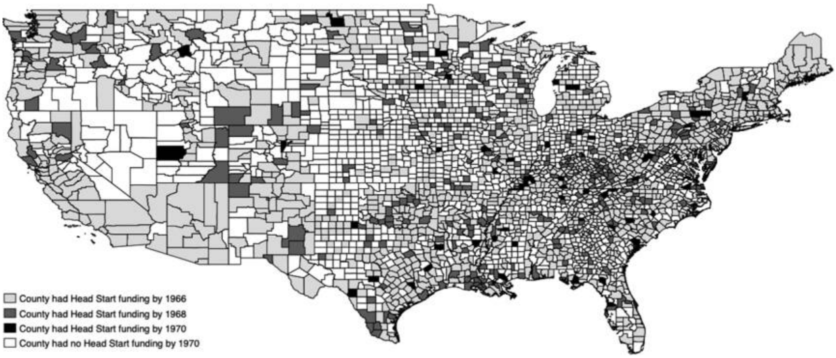

FIG. 1.—Early geographic expansions of Head Start. Sources: NARA Record Group 381, CAP grants and grantees, and FOS files.

officially announced the program in February 1965, and it was operational by summer, rendering strategic funding behavior and coordinated grantee response more challenging and less likely. Historical accounts describe "great administrative confusion" (Levine 1970) and a "wild sort of grant-making operation" (Gillette 1996, 193). OEO distributed funds directly to local grantees to circumvent governors, state legislatures, and agencies that may have prevented the funds from reaching disadvantaged Black children (Vinovskis 2005; Gibbs, Ludwig, and Miller 2011). Historical accounts describe early Head Start implementation as happening "with explosive speed" (White and Phillips 2001, 85), in OEO Director Sargent Shriver's "characteristic cyclonic manner" (Greenberg 1990, 43), and with emphasis on a rapid, large-scale launch (Vinovskis 2005). Government documents and historical records reveal that the administration did not anticipate the volume of interest from local communities, resulting in significant expansion of the rollout with broad coverage across the country (Office of Child Development 1968; Vinovskis 2005; Zigler and Styfco 2010).[8]

In the program's early years, approximately half of US counties received Head Start funding, and the majority that eventually implemented the program did so between 1965 and 1970 (Johnson and Jackson 2019; Bailey, Sun, and Timpe 2021). As a result of the funding's uncoordinated, local distribution, programs became available in different counties at different times over a relatively short window. In figures A1 and A2, we plot the bivariate relationship between 1960 county characteristics and Head Start adoption through 1970. The plots illustrate the substantial variation in adoption year among counties with similar baseline characteristics.

[8] The Johnson administration noted that 49% of all applications were from rural areas and there would be an emphasis on involving the country's 300 poorest counties, called "special-target counties," in the rollout (Office of the President 1965; Cort et al. 1966).

Higher-poverty counties tended to adopt the program earlier (Greenberg 1990). However, conditioning on 1960 poverty level (% Family Income < $3,000), other county characteristics are largely unrelated to adoption year (table A1).[9] Together, the baseline county characteristics explain little of the observed adoption-year variation, are consistent with historical accounts of early Head Start rollout, and support the validity of our rollout approach.[10]

Given the broader adoption of War on Poverty programs during this period, one might be concerned that Head Start's launch coincided with the introduction of other programs with the potential to affect child outcomes. The major programs likely to affect children were at most modestly more likely to start in counties when Head Start was initially available (fig. A3). Community health centers and Food Stamp programs were no more likely to show up at the time of initial Head Start availability. We similarly see minimal evidence of an increase in grants funding senior programs. There is evidence of a small increase in the presence of a grant related to an uncategorized "CAP (Community Action Program) Health" project. However, as Bailey, Sun, and Timpe (2021) note, funding for the program was orders of magnitude smaller than that for Head Start. Indeed, most of the programs' funding levels were substantially lower than that for Head Start. Differences in programs' age targeting, or lack thereof, further limit concerns that their availability confounds identification of the effects of Head Start availability. It is unlikely, and we see no evidence in historical accounts to suggest, that CAP Health or other programs would affect children age-eligible for Head Start but not those slightly older. That said, the availability of other antipoverty programs is an important part of the context into which Head Start emerged, and they may have magnified Head Start's effects.

We employ a strategy similar to the 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).[11] Researchers have also leveraged geographic program expansion variation to assess early-childhood intervention impacts in

[9] An F-test of the nonpoverty controls fails to reject the null that they are equal to zero (p-value = .783). The pattern is consistent with Bailey and Duquette's (2014) account of War on Poverty programs' earliest implementation years.

[10] If early adopters are on a different trend than later adopters, these trends would confound our approach. To address the concern, we explore the robustness of our results to specifications interacting county baseline covariates with birth-year trends (see table A7).

[11] These studies have found little relationship between Head Start spending and their programs of interest.

Germany and Denmark (Cornelissen et al. 2018; Rossin-Slater and Wüst 2020). Three papers, developed concurrently to ours, use the introduction of Head Start over geography and time to explore first-generation outcomes; all three document long-term impacts for those exposed to Head Start (Thompson 2018; Johnson and Jackson 2019; Bailey, Sun, and Timpe 2021), particularly when coupled with subsequent schooling investment (Johnson and Jackson 2019). We focus on a more constrained window of program rollout than other work on poverty-reduction programs and papers leveraging geographic Head Start variation. Because of overlap with the NLSY79 birth cohorts, we capitalize on Head Start's plausibly exogenous introduction across counties from 1966 to 1969. In identifying program effects, we capitalize on Head Start's narrow targeting to 4- and 5-year-old children from disadvantaged households.

We supplement our rollout strategy with a difference-in-differences strategy relying on variation in grant-writing assistance provision across counties. Because of concerns that disadvantaged communities would be unable to respond to its grant application call, OEO sent government interns to poor counties to distribute applications, meet with community leaders, mobilize local resources, and assist in application preparation (Cort et al. 1966; Greenberg 1990). Ludwig and Miller (2007) capitalize on a discontinuity in the provision of grant-writing assistance at the 300th-poorest county to estimate Head Start's impacts on first-generation outcomes. As they note, the approach is challenging when using survey data because many counties are not represented and there are limited individual observations around the discontinuity. We therefore use the variation in a difference-in-differences framework, taking advantage of the NLSY79's clustered sampling to examine the difference in outcomes between children of mothers born too early for Head Start and children of mothers born later, coupled with how the difference varies with whether a mother's birth county received grant-writing assistance.

III. Data

To explore Head Start's intergenerational effects, we rely on rich, longitudinal survey data connecting mothers and their children. The NLSY79 is a nationally representative sample of adolescents who were 14–22 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 for 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.[12]

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 the child, depending on age. The survey gathers data on children's schooling and training, labor market experiences, and engagement in risky behaviors as well as earlier-life measures including birth weight, evaluations of the home environment, and an array of cognitive and noncognitive measures.

In addition to the NLSY surveys, we use county-by-year data from the CAP and Federal Outlays System (FOS) files available from NARA for Head Start availability in fiscal years 1966–70 (see app. B for details). We aggregate data on Head Start grant recipients to the county level and code a birth cohort as "exposed" if their county of birth received per-child Head Start funding above the 10th percentile ($22 per 4-year-old in the county) when the cohort reached 4 or 5 years old.[13] While the cutoff was somewhat arbitrarily chosen to exclude low funding levels potentially associated with small pilot programs or planning grants, our results are insensitive to this choice.[14] We do not otherwise leverage data on appropriated dollar amounts 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

[12] 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 1960–64 cohorts correspond very closely to those in other data sources (Current Population Survey) in both levels and trends, there is a significant positive jump in education levels for those born before 1960, opposite the trend observed in the Current Population Survey. These issues are problematic, as we use grandmother education levels, a proxy for Head Start eligibility, to restrict our sample to mothers likely to be eligible for Head Start. A second concern is related to sample selection across cohorts induced by the sample design and the later decision to exclude certain subsamples of the NLSY79. In app. I, we provide additional estimates including the 1957–59 birth cohorts and using alternative approaches to focus the sample on groups of individuals more likely to have been affected. The effects are similar but somewhat smaller.

[13] The measure for a cohort is the average funding per child in the cohort 4 and 5 years after birth, when the cohort was turning 4 or 5 years old. The modal age of participation in the summer program was 5, while the modal age of participation in the full-year program was 4 between 1966 and 1968. Throughout this period, over 90% of participants were ages 3–5.

[14] Our results are similar using $0 or any percentile between the 1st and 30th as the cutoff. We discuss the measurement of Head Start availability and the sensitivity of the results to different assumptions on the extent of misclassification of Head Start availability in sec. IV and app. C.

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 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 parenthood and interaction with the criminal justice system (as measured by any reported arrests, convictions, or probations). These outcomes are important in capturing the 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 outcome measure. While much of the prior literature has used a summary index of the average of standardized outcomes, we focus on a more easily interpreted summary measure similar to that used in García et al. (2020).[15] We use a nonparametric approach to map the observed effects on our key outcome measures into an effect on the NPV of wages through age 50 (hereafter NPV Wages). Specifically, we match the CNLSY participants in our analytical sample to similar individuals in a synthetic cohort in which earnings are observed over a broader range of ages to construct an NPV Wages measure for each individual. The details of this approach are contained in appendix D and generally follow García et al. (2020). In appendix E, we demonstrate the robustness of the results to multiple alternative constructions of the summary measure.

Table 1 contains summary statistics for our sample. Column 1 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 before 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

[15] We include estimated effects on a standardized index in app. E for comparability with prior papers.

TABLE 1
Descriptive Statistics

	Full Sample (1)	High-Impact Sample (Grandmother < HS) (2)	Low-Impact Sample (Grandmother ≤ HS) (3)
Teen parent	.17	.21	.17
Crime	.26	.28	.27
High school	.85	.80	.84
Some college	.58	.50	.56
NPV Wages ($)	334,186	316,842	331,431
Black	.38	.44	.39
Male	.50	.50	.50
Mother's household in poverty (1978)	.33	.45	.35
Observations	3,989	1,978	3,321

Note.—Sample is restricted to CNLSY respondents aged 20 or above by 2016. 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 the full sample. 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 on the basis of the education level of the mother of the NLSY79 participant (i.e., the grandmother of the children of interest). Column 2, the high-impact sample, is restricted to participants in the NLSY79 whose mothers dropped out of high school (HS). Column 3, the 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. Crime is defined as any arrests, convictions, or probations, following Deming (2009). Income (NPV Wages) is a measure of the mother's permanent lifetime income, calculated as the deflated average of net family income for the mother. See app. D for additional details on the construction of the NPV wage summary measure.

of interest) did not finish high school (col. 2).[16] We refer to this as our high-impact sample, because close to 70% of participants' mothers had less than a high school degree, implying participation rates of close to 60% 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. Their mothers were more likely to grow up in poverty, and they have higher rates of teen parenthood (21% vs. 17%) and interaction with the criminal justice system (28% vs. 26%) and lower rates of high school completion (80% vs. 85%) and college going (50% vs. 58%). In column 3, we restrict the sample to NLSY79 mothers whose mothers completed at most a high school degree. We call this our low-impact sample, as we estimate participation rates of at most 30% in counties with Head Start availability. As might be expected, summary statistics for children in this group indicate somewhat better outcomes.

[16] Roughly 65%–70% of the mothers of early participants reported less than a high school education, approximately 25% indicated that they graduated from high school, and only 5% of mothers reported some postsecondary schooling or more.

[17] See the note to table A2 for details of these calculations.

IV. 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 program's rollout in the late 1960s. We identify the average effect of program availability within a mother's county of birth on the adult outcomes of her children. Between 65% and 70% of the mothers of early Head Start participants did not complete high school (Office of Child Development 1968; Barnow and Cain 1977), 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, respectively. The term $HSavail_{ct}$ indicates whether Head Start was available for a mother in a particular birth cohort t and birth county c. It is set to 1 for a mother when the level of Head Start funding in that mother's birth county 4 or 5 years after her year of birth exceeds the 10th percentile of the funding measure throughout the period. Standard errors are clustered at the county-of-birth level.

Appendix G 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 of table A2 presents estimates of the effect of Head Start availability on self-reported participation. When a program is available in a county 4 or 5 years after a mother's year of birth, the mother in our high-impact sample is roughly 10 percentage points more likely to report having participated in Head Start as a child. As discussed in the appendix, the considerable misreporting that is present in this variable results in a substantial downward bias on the true first-stage estimate.[18] Under our best assumptions on the extent of misreporting, the true relationship between Head Start availability and Head Start participation is over 5 times that estimated, with a true first stage in our high-impact sample of between 0.53 and 0.6. This larger first stage is consistent with that indicated by Johnson and Jackson (2019), who suggest a first stage among children in poor families of as high as 86% during this period.

[18] The Head Start participation question was asked retrospectively in 1994, when sample members were 30–34 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 4 or 5 years old. A variety of evidence from the psychology literature indicates that retrospective reports of early childhood are unreliable (see app. G for further discussion).

That said, even with full confidence in the 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 nonparticipants within the same cohort. Instead, our parameter of interest captures the reduced-form (or intent-to-treat) effect of Head Start availability among specific subpopulations with high rates of eligibility for the program. This effect measures direct effects on participants and spillover effects on nonparticipating individuals in the same cohort, as well as any indirect effects operating through contemporaneous or subsequent improvements in the environment that result from Head Start availability. Our parameter averages the effects among those with Head Start available into a single treatment effect, combining larger and smaller effects experienced by different individuals. 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.

Our baseline results are contained in table 2. Panel A 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 ($35,471, 11%) of a mother's Head Start availability on our summary measure of second-generation discounted wages (NPV Wages) that is statistically significant at the 1% level. As we discuss further below, this increase ranges from 6% to 11%, depending on specification choice. The effect on the summary measure is driven by reductions in teen parenthood (8 percentage points) and criminal engagement (13 percentage points) and increases in high school graduation (11 percentage points) and college enrollment (18 percentage points). While these effects are large in magnitude, there is significant uncertainty surrounding the point estimates. Furthermore, while the broad conclusion of important intergenerational spillovers is largely insensitive to specification choice or adjustments for measurement error in Head Start availability, discussed in more detail later in this section, the magnitude of the estimates is as much as 30%–50% smaller, depending on these choices.[19] Panel B 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 ($13,624, 4%) increase in the summary measure.[20]

In panel C, 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

[19] Adjustments for measurement error are addressed in app. C and table C1 and discussed in more detail later in this section.

[20] The estimates are consistent with the relative participation rates in the high-impact (60%) and low-impact (30%) samples.

TABLE 2

REDUCED-FORM EFFECT OF HEAD START AVAILABILITY IN COUNTY

	Teen Parent (1)	Crime (2)	High School (3)	Some College (4)	NPV Wages ($) (5)
	A. High-Impact Sample				
Grandmother < high school education	−.078**	−.125***	.109***	.182***	35,471***
	(.031)	(.040)	(.032)	(.048)	(8,012)
Sensitivity to availability measure[a]	[−.085, −.023]	[−.125, −.054]	[.055, .115]	[.099, .187]	[17,858, 35,471]
Sensitivity to alternative specifications[b]	[−.082, −.038]	[−.140, −.096]	[.075, .121]	[.097, .182]	[20,852, 43,372]
Observations	1,978	1,978	1,978	1,978	1,880
Mean	.205	.283	.809	.499	316,903
	B. Low-Impact Sample				
Grandmother ≤ high school education	−.037	−.057*	.049*	.097***	13,624*
	(.024)	(.031)	(.027)	(.036)	(7,299)
Sensitivity to availability measure[a]	[−.042, −.003]	[−.058, −.001]	[.013, .052]	[.044, .099]	[2,019, 14,005]
Sensitivity to alternative specifications[b]	[−.046, −.028]	[−.073, −.051]	[.033, .052]	[.090, .118]	[12,042, 15,164]
Observations	3,321	3,321	3,321	3,321	3,143
Mean	.173	.266	.842	.556	331,431
	C. Placebo Sample (Likely Ineligible for Head Start)				
Grandmother ≥ high school education	−.003	.008	−.006	−.018	−8,104
	(.032)	(.043)	(.031)	(.050)	(10,258)
Observations	1,769	1,769	1,769	1,769	1,669
Mean	.115	.223	.902	.682	359,753

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. The coefficient presented is for the 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 and age indicators as well as mother's birth order and race.

[a] Cells in this row provide the range of estimates produced using different measures of Head Start availability (discussed in app. C), including different cutoffs (1st–20th percentiles) for the binary Head Start availability variable and various approaches to dealing with discrepant counties presented in table C1.

[b] Cells in this row provide the range of estimates produced using the different empirical specification choices in cols. 1–6 of table A7, including controls for birth-county characteristics by birth-cohort trends, county by birth-cohort trends, birth-cohort by birth-state fixed effects, and War on Poverty measures.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

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.[21] 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. The point estimates for this group are small, frequently opposite signed, and statistically indistinguishable from zero across all outcomes.[22] While our baseline inference relies on standard errors clustered at the county-of-birth level, our conclusions are robust to a variety of other assumptions (tables A3, A4; figs. A5, A6).[23]

The use of archival grant funding data from the 1960s introduces additional uncertainty into our Head Start availability measures. While we have undertaken considerable work to validate our Head Start availability measures (see app. C), some measurement concerns remain. These measurement concerns are largely driven by discrepancies between our measures and those produced by other researchers, as well as suggestive evidence (from newspaper articles and aggregate reports) that we may be missing a modest set of counties with Head Start programs during the early years. Our conclusions are robust to a variety of sensitivity analyses designed to address these concerns (app. C). These checks include dropping mothers born in discrepant county-cohort cells, counties with any discrepancy, or counties with uncertain codes, or simulating the effect of this uncertainty (table C1). Our conclusions are similarly insensitive to exercises that provide bounds on our estimates from simulating the effect of recoding as funded additional potentially missing programs in the early years of Head Start's introduction. Finally, we demonstrate that the results are insensitive to the percentile threshold choice (figs. C3–C5). While the conclusion of significant intergenerational spillovers is largely insensitive to reasonable adjustments to measurement of Head Start availability, the magnitude of the estimates suggests effects that may be 30%–50% smaller than our main estimates.

Focusing on the high-impact sample, table 3 presents results separately by gender and race. Consistent with recent work evaluating the intergenerational effects of the Perry Preschool Project, we see larger level effects

[21] We estimate that participation rates in this group were at most one-fifth of those in our high-impact sample. Table A6 further illustrates that our measure of Head Start availability has a small and nonsignificant relationship with self-reported Head Start participation in this subsample.

[22] We similarly find nonsignificant 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.

[23] This includes clustering on birth county by birth cohort, clustering on birth state, and the wild cluster bootstrap (tables A3, A4), as well as "*p*-values" generated by the generally conservative approach of randomization inference (figs. A5, A6; see Abadie, Diamond, and Hainmueller 2010 for a discussion of this procedure and examples of its implementation). The results are also robust to adjusting the *p*-values following Romano and Wolf (2005), with adjusted *p*-values for the high-impact sample outcomes all remaining below .05.

TABLE 3
Reduced-Form Effect of Head Start Availability in County (Heterogeneity)

	Teen Parent (1)	Crime (2)	High School (3)	Some College (4)	NPV Wages ($) (5)
Males	−.031	−.229***	.137**	.183***	34,113***
	(.040)	(.060)	(.055)	(.066)	(11,522)
Observations	988	988	988	988	938
Mean	.140	.401	.751	.402	392,389
Females	−.133**	−.029	.081**	.195**	25,251**
	(.054)	(.046)	(.040)	(.077)	(11,439)
Observations	990	990	990	990	942
Mean	.270	.166	.866	.597	241,897
Black	−.114**	−.061	.114**	.230***	37,397***
	(.050)	(.045)	(.049)	(.052)	(9,510)
Observations	866	866	866	866	823
Mean	.218	.281	.789	.476	267,475
Non-Black	−.048	−.197***	.096**	.130	29,580**
	(.051)	(.066)	(.048)	(.082)	(11,704)
Observations	1,112	1,112	1,112	1,112	1,057
Mean	.195	.285	.824	.518	355,317

Note.—Sample is restricted to CNLSY respondents whose maternal grandmothers had less than a high school education (high-impact sample). 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. The coefficient presented is for the 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 and age indicators as well as mother's birth order and race.

** $p < .05$.

*** $p < .01$.

for males than for females on our summary measure of NPV Wages. However, this appears to be largely driven by the higher overall life-cycle earnings experienced by males; the point estimates for individual outcomes are generally similar, with the unsurprising exceptions of teen parenthood, where we see larger effects for females, and criminal behavior, where we see larger effects for males. In the bottom half of the table, we observe larger effects for Black children in the second generation, consistent with the higher levels of participation among Blacks, even when conditioning on maternal education.

A. *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 have the program available and children born to those who did not. For example, one concern would be that the type of woman who became a mother or the type of woman included in the sample (as a result of nonresponse 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 (cols. 1–5 of table A5). 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 level of the grandmother, which we use to focus our sample, is affected by Head Start availability. In columns 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.[24] 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.[25]

A second concern is that of endogenous program adoption or preexisting positive trends in outcomes in counties that adopted a Head Start program.[26] The presence of meaningful pretrends 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 preexisting trends, we also probe the robustness of our estimates to the inclusion of differential trends by birth county (table A7). Allowing differential birth-cohort trends interacted with baseline (1960) county characteristics, or even birth-county-specific trends, does little to change our conclusions, but it does reduce the point estimates somewhat. Findings are similarly robust to the inclusion of more specific county-cohort controls for spending on War on Poverty programs and state by year-of-birth fixed effects, which

[24] 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.

[25] 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 col. 8) are smaller than, but statistically indistinguishable from, the other point estimates in table A7. The attenuation may be a result of sibling spillovers, although we hesitate to draw conclusions, given the very large confidence intervals and the different sample that results from restricting to families with at least two sisters in the original NLSY79 sample.

[26] Our results are robust to restricting our sample to counties that received Head Start by the end of our sample period (col. 9 of table A7), so it would have to be the case that these positive trends occurred just before adoption and not just that adopting counties were on a positive trend relative to nonadopters for this issue to threaten our strategy.

flexibly control for changes over time within states that could affect maternal outcomes.[27]

Figure 2 addresses these concerns graphically, demonstrating the relationship between Head Start availability within a county and our summary NPV Wages measure 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 nonnegative 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 before 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.

The pattern of effects across subgroups provides further support for the identification assumption. We see larger effects for those with higher rates of eligibility (based on a grandmother's education or race). Finally, the lack of effects of Head Start availability among those largely ineligible for the program serves as a placebo exercise. If there are other programs or community investments occurring in these counties that could affect second-generation outcomes, we might expect to detect the effects regardless of likely eligibility for Head Start. In panel C of table 2, we see no evidence of Head Start availability effects among the children of likely-ineligible mothers.

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. Discussed at greater length in appendix F, we use variation in Head Start exposure generated by the provision of Head Start grant-writing assistance to the poorest 300 counties, adapting the approach of Ludwig and Miller (2007). Table F1 indicates large intergenerational effects of Head Start grant-writing assistance on our summary measure of second-generation outcomes; we reconcile the results further in the next section.

B. 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

[27] The War on Poverty program variables control for spending on Medicaid, CAP administrative grants, cash assistance, CAP health programs, and community health centers. The insensitivity of results to the inclusion of these controls is consistent with the (conditional on year) minimal correlation between the timing of Head Start and other War on Poverty program adoption within a county (fig. A3).

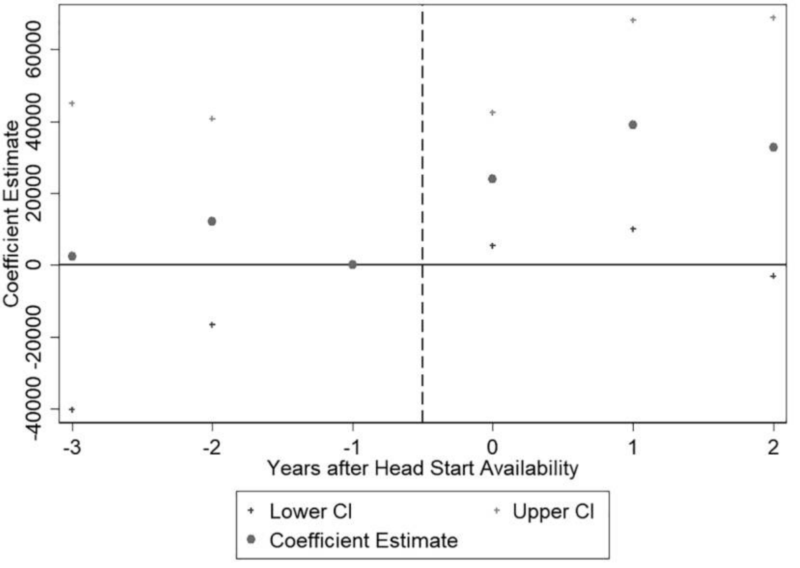

Fig. 2.—Event study for second-generation NPV Wages measure (high-impact). The high-impact sample is restricted to CNLSY respondents whose maternal grandmothers had less than a high school education. Circles indicate coefficients on indicator variables for the difference between a mother's birth cohort and the first birth cohort to have Head Start availability within a county (nonnegative years reflect cohorts with Head Start availability). The dependent variable is the summary measure of NPV wages in the second generation (see app. D for additional details on the construction of this measure). See the notes to tables 1 and 2 for additional information on sample restrictions. In addition to year-of-birth and county-of-birth fixed effects, controls include child gender and age indicators as well as mother's birth order and race. Standard errors are clustered at the birth-county level. Plus signs indicate 95% confidence intervals (CI).

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.[28] We then situate our results in the small literature that explores the intergenerational effects of a schooling or early-childhood intervention.[29]

As discussed above (and further in app. G), we put little stock in "first stages" estimated with the self-reported and retrospective Head Start

[28] A more complete discussion of the first stage is contained in app. G.

[29] We also compare the first-generation Head Start effects estimated using our strategies with previous estimates in the literature in app. H. As noted in that appendix, the confidence intervals around our first-generation impact estimates include all estimates of first-generation effects in the literature, including those of Thompson (2018), Johnson and Jackson (2019), and Bailey, Sun, and Timpe (2021). Given the imprecision of our estimates, we are limited in our ability to draw strong conclusions about the magnitude of our effects, but our estimates are larger than those estimated in Thompson (2018) and Bailey, Sun, and Timpe (2021) and smaller than those estimated in Johnson and Jackson (2019). Further discussion of these differences is contained in app. H.

participation variable because the participation question was asked retrospectively in 1994, when sample members were 30–34 years old. Our investigation in appendix G suggests a likely first stage in our high-impact sample of 0.53–0.6. Our estimated first stages are broadly consistent with those of the other papers using rollout-based variation in Head Start exposure in identifying effects. For example, Johnson and Jackson (2019) estimate participation rates of around 86% 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.[30] Their sample of poor families has rates of eligibility relatively similar to those of our high-impact sample. While we hesitate to similarly restrict our sample because of the potential endogeneity of income measures in the NLSY79, first- and second-generation estimates using specifications similar to those employed by Johnson and Jackson are consistent with our main analyses (tables A8, A9).[31]

The estimates in table F2 suggest a somewhat larger effect on reported Head Start participation, using the grant-writing assistance strategy. Assuming the same upper-bound misclassification rates among this sample would suggest a true first stage of around 0.8. This estimate is consistent with an upper bound inferred from the fraction of children in counties that received grant-writing assistance who were eligible to participate in Head Start, roughly 80%, as well as some accounts from this early period suggesting that Head Start served nearly all Head Start–aged children in their communities (Cort et al. 1966).[32] While these statistics and case studies suggest quite high participation rates, we expect that not all high-poverty counties were treated to the same extent. Furthermore, there is considerable uncertainty in the misclassification rates for participation within this group. The true first stage for this sample and strategy may be significantly lower.

Regardless of specification or sample choice, if we assume that the entirety of our estimated effects of Head Start access was generated by participation (and there were no spillover effects), the implied TOT

[30] 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.

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

[32] This upper bound is likely an underestimate in our sample because the NLSY79 oversamples poor and Black populations, who have higher rates of Head Start participation. For example, an early evaluation report indicates that Knox County, Kentucky, one of the country's poorest 300 counties in 1960, served 452 children in its first year providing Head Start and reported 450 income- and age-eligible children in the community (Cort et al. 1966). On the basis of data from the 1960 census, we estimate that there were only around 540 4-year olds in total in the county at the time.

effects from our main estimates are large, suggesting increases in second-generation NPV Wages of over 20%, with similarly large increases in high school completion and college enrollment. Allowing for a somewhat smaller first stage (as discussed in app. G) would suggest implausibly large effects. However, as noted above, while the conclusion of significant intergenerational spillovers is largely insensitive to specification choice or adjustments for measurement error in Head Start availability, the magnitude of our estimates is not. Across a range of definitions of Head Start availability and different specification choices, our estimates suggest effects that are 30%–50% smaller than our main estimates. For our NPV Wages measure, this includes estimated effects as small as 6.5%, with 95% confidence intervals as low as 2.3%. These estimates would imply TOT effects of closer to 11% and 4%, respectively.

That said, we do not think it is reasonable to interpret these scaled estimates as TOT effects, because 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. 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 substantial share of each school-entry cohort was more prepared. These subsequent schooling improvements, or responses to children's skill development, may constitute a channel through which longer-term outcomes were further boosted by the Head Start investments and were, in fact, an emphasis of the OEO's guidance to local communities in setting up their programs. As a result of these likely spillovers and the uncertainty surrounding the first stage in our sample, we focus on capturing the effect of program availability among those with a specific family disadvantage profile for whom participation was likely quite high. This approach captures both direct effects on participants and indirect effects on their likely peers. Put another way, among all relatively disadvantaged children (as proxied by maternal education levels), we estimate the average effect of providing Head Start to the 53%–60% who are the most disadvantaged.

Comparing intergenerational effect sizes with other estimates.—Using the rollout approach, the level effects on the second generation are at least as large as effects on the first generation, while the percent changes are similar across the education measures (table A11). We continue to find larger estimated effects in the second generation using our grant-writing assistance approach (table F7). 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 (2007) 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 using National Educational Longitudinal Study data 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 (Akresh, Halim, and Kleemans 2018; Heckman and Karapakula 2019; Mazumder, Rosales-Rueda, and Triyana 2019; Rossin-Slater and Wüst 2020; García, Heckman, and Ronda 2021). 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, Halim, and Kleemans 2018; Mazumder, Rosales-Rueda, and Triyana 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 Wüst (2020) 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 (García, Heckman, and Ronda 2021). This experimental study is distinct from other intergenerational inquiries, including our own, because it identifies the effect of participation in a specific, high-quality program rather than the average effect of a large-scale and heterogeneous program's availability on a cohort of individuals. While the study focuses on a different parameter, the authors similarly estimate large intergenerational transmission of effects. 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 employment of around 27 percentage points for female participants, and the intergenerational study finds an increased likelihood of 26 percentage points for the participants' children (Nores et al. 2005; García, Heckman, and Ronda 2021). While the intergenerational effects of the Perry Preschool Project differ by gender, the studies collectively document persistent effects for the participants themselves and the participants' children on important behavioral, health, and educational attainment outcomes (Nores et al. 2005; García, Heckman, and Ronda 2021).

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.

C. Mechanisms

To better understand the channels through which mothers' exposure to Head Start affects the second generation, we explore several intermediate outcomes measured in the second generation's early-childhood and adolescent years. To select the constructs, we relied on previous literature on first-generation effects of Head Start and other early-childhood programs and existing evidence on intergenerational spillovers from mothers' educational attainment (Currie and Moretti 2003; Oreopoulos, Page, and Stevens 2006; Deming 2009; Carneiro and Ginja 2014; Thompson 2018; Bailey, Sun, and Timpe 2021), as well as more recent working papers on potential channels of intergenerational transmission (García, Heckman, and Ronda 2021; Kose 2021). This investigation includes representation of important domains in child development, including early home environment and parenting, health, cognitive performance, and noncognitive skills. We report program effects for the high-impact sample in table 4.[33]

While we cannot rule out small changes, we estimate no effect of Head Start availability on birth weight in the second generation.[34] We find evidence, however, of significant changes in the second generation's childhood environment. Head Start availability in the first generation leads to a higher likelihood that one's children participate in any preschool, and specifically in Head Start.[35] We also see evidence of improvement on dimensions of the childhood home environment. Head Start exposure in the first generation leads to higher Home Observation Measurement of the Environment (HOME) scores on the comprehensive summary measure, as well as on the cognitive stimulation and emotional support subscores, suggesting improvements in parenting on multiple dimensions.[36] We do not detect any differences on the Behavior Problems Index (BPI-E), which measures externalizing problem behaviors on the basis of mothers' reports.[37]

We also explore outcomes throughout childhood and into adolescence. While the estimated effect on test scores is positive and of modest

[33] While table 4 is restricted to individuals in our main sample and therefore last observed at age 20 or older, in table A13 we present results for children observed in the CNLSY at any age.

[34] We similarly see no effect when estimating effects on incidence of low birth weight or very low birth weight.

[35] Both Head Start and preschool participation were measured contemporaneously in the CNLSY, or soon after early childhood for children born to NLSY79 respondents before 1982, so this measure does not suffer from recall-induced measurement error as in the NLSY79, in which respondents reported Head Start participation retrospectively in adulthood.

[36] The HOME Inventory is compiled from interviewer observations and mother reports and captures the nature and quality of the home environment.

[37] It may be worth noting that, unlike for the HOME Inventory or test scores (which are all standardized to mean zero and standard deviation one within age of observation within the full sample), the mean level of the BPI-E within the high-impact sample suggests only modestly worse outcomes among this more disadvantaged group.

TABLE 4
Reduced-Form Effect of Head Start Availability in County (Earlier Outcomes)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	A. Early Childhood						
	Birth Weight	Preschool	Head Start	HOME	HOME (Cog)	HOME (Emot)	BPI-E
HS in County	1.030	.133***	.112***	.235***	.198**	.203***	−.053
	(2.912)	(.044)	(.036)	(.070)	(.081)	(.062)	(.107)
Observations	1,817	1,889	1,912	10,860	10,258	9,620	8,469
Mean	114.7	.636	.297	−.276	−.263	−.191	.0965
	B. Later Childhood						
	Test Scores (8)	LD (9)	Repeat (10)	Crime (11)	Poor Health (12)	CES-D (13)	Rosenberg (14)
HS in County	.095	−.010	−.103**	−.057**	.020	.002	.194**
	(.081)	(.025)	(.043)	(.022)	(.033)	(.075)	(.077)
Observations	8,845	1,974	1,965	1,777	1,845	4,811	3,522
Mean	−.318	.0502	.308	.0972	.140	.0310	.0236

Note.—Sample is restricted to CNLSY respondents whose maternal grandmothers had less than a high school education (high-impact sample). 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. "Birth Weight" is measured in ounces, "Preschool" and "Head Start" are participation indicators, "LD" is an indicator for the reported presence of a learning disability, "Repeat" is an indicator for any grade repetition, "Crime" is an indicator for any reported engagement with the criminal justice system before age 18, and "Poor Health" is an indicator for self-reported health less than "fair" at ages 15–18. The "HOME," "BPI-E," "Test Scores," "CES-D," and "Rosenberg" measures are all standardized to have a mean of zero and a standard deviation of one within age of observation. "HOME" is the Home Observation for Measurement of the Environment total percentile score standardized; "HOME (Cog)" is the cognitive stimulation percentile score standardized; "HOME (Emot)" is the emotional support percentile score standardized; "BPI-E" is the Behavioral Problems Index (Externalizing) percentile score standardized on which higher values indicate more externalizing behavioral problems; and "Test Scores" captures pooled, standardized Peabody Picture Vocabulary Test or Peabody Individual Achievement Test scores measured at different ages. "CES-D" is the percentile score for the depression scale standardized; and "Rosenberg" is the percentile score for the measure of self-esteem standardized. The coefficient presented is for the binary variable indicating Head Start (HS) 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 and age indicators as well as mother's birth order and race.

** $p < .05$.

*** $p < .01$.

size, lack of precision limits our capacity to reject null effects. This result is largely consistent with other Head Start literature (in the first generation), which tends to find at most modest effects on cognitive measures in the later childhood years. There are no detectable effects on the presence of a learning disability, poor health in adolescence, or depression. It is not clear whether (or in what direction) Head Start availability should influence the presence or reported presence of learning disabilities or depression in the next generation, but we include these measures for comparability with prior work. In contrast, we do observe improvements in self-esteem among the second generation. We also observe reductions in grade repetition and adolescent criminal engagement that foreshadow the increased educational attainment and reduced criminal behavior we observe in young adulthood. Overall, the pattern of effects is consistent with Head Start exposure among mothers leading to significant changes in both the home and early-schooling environments of the second generation, leading to persistent developmental benefits and improved later-life well-being. Importantly, the exploration of potential mechanisms suggests many improvements in early life and childhood that likely serve as pathways to longer-term increased educational attainment and reduced risky behavior.

V. Discussion and Conclusion

Research and policy discussions frequently focus on how to level the playing field for those born into poverty. We focus more broadly on whether such interventions truly break the cycle of bad outcomes. While there is increasing interest among researchers and policy makers in understanding the intergenerational spillovers of particular policies and interventions, there exist very few contexts in which these questions can be tested empirically; this challenge is particularly acute for early-childhood interventions, an area of growing 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 large-scale US evidence on the intergenerational transmission of early-childhood intervention effects.

We find strongly suggestive 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. These effects correspond to an estimated 6%–11% increase on a summary measure of discounted wages through age 50, depending on specification choice. Given the documented importance of mothers' human capital for their children's early health and developmental outcomes, we would expect that mothers' own increased educational attainment and resources would translate to improvements in 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). Indeed, we observe direct measures of improvement in the home environment and second-generation preschool participation that precede improvements in self-esteem, reductions in grade repetition, and lower criminal engagement in adolescence. 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.

There are few settings in which researchers can overcome the challenges associated with estimating intergenerational program effects, and while the rollout of Head Start is one such context, it is not without its limitations. The key takeaway is that the benefits of Head Start appear to transfer from one generation to the next in a substantial way. That said, we recognize that there is always some uncertainty regarding the conditional randomness of a program's rollout, and perhaps additional uncertainty in our context, given reliance on historical records for measurement of our identifying variation. Because the results are robust to considerable sensitivity analyses on both margins, we conclude that the findings are highly suggestive of important intergenerational transmission. While the evidence strongly suggests important intergenerational transmission, we urge caution in interpreting the magnitude of our point estimates of Head Start effects in the second generation, given the aforementioned uncertainties as well as the imprecision of our estimates.

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. These results imply that cost-benefit analyses of Head Start and similar early-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

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *J. American Statis. Assoc.* 105 (490): 493–505.

Akresh, Richard, Daniel Halim, and Marieke Kleemans. 2018. "Long-Term and Intergenerational Effects of Education: Evidence from School Construction in Indonesia." Working Paper no. 25265 (November), NBER, Cambridge, MA.

Almond, Douglas, and Janet Currie. 2011. "Killing Me Softly: The Fetal Origins Hypothesis." *J. Econ. Perspectives* 25 (3): 153–72.

Almond, Douglas, Hilary W. Hoynes, and Diane Whitmore Schanzenbach. 2011. "Inside the War on Poverty: The Impact of Food Stamps on Birth Outcomes." *Rev. Econ. and Statis.* 93 (2): 387–403.

Anders, John, Andrew Barr, and Alexander A. Smith. Forthcoming. "The Effect of Early Childhood Education on Adult Criminality: Evidence from the 1960s through 1990s." *American Econ. J. Econ. Policy*.

Auten, Gerald, Geoffrey Gee, and Nicholas Turner. 2013. "Income Inequality, Mobility, and Turnover at the Top in the US, 1987–2010." *A.E.R.* 103 (3): 168–72.

Bailey, Martha J., and Nicholas J. Duquette. 2014. "How Johnson Fought the War on Poverty: The Economics and Politics of Funding at the Office of Economic Opportunity." *J. Econ. Hist.* 74 (2): 351–88.

Bailey, Martha J., and Andrew Goodman-Bacon. 2015. "The War on Poverty's Experiment in Public Medicine: Community Health Centers and the Mortality of Older Americans." *A.E.R.* 105 (3): 1067–104.

Bailey, Martha J., Shuqiao Sun, and Brenden Timpe. 2021. "Prep School for Poor Kids: The Long-Run Impacts of Head Start on Human Capital and Economic Self-Sufficiency." *A.E.R.* 111 (12): 3963–4001.

Barnow, Burt S., and Glen G. Cain. 1977. "A Reanalysis of the Effect of Head Start on Cognitive Development: Methodology and Empirical Findings." *J. Human Resources* 12 (2): 177–97.

Bauer, Lauren, and Diane Whitmore Schanzenbach. 2016. "The Long-Term Impact of the Head Start Program." Economic analysis, The Hamilton Project, Brookings Inst., Washington, DC.

Bitler, Marianne P., Hilary W. Hoynes, and Thurston Domina. 2014. "Experimental Evidence on the Distributional Effects of Head Start." Working Paper no. 20434 (August), NBER, Cambridge, MA.

Black, Sandra E., and Paul J. Devereux. 2011. "Recent Developments in Intergenerational Mobility." In *Handbook of Labor Economics*, vol. 4B, edited by David Card and Orley Ashenfelter, 1487–541. Amsterdam: North-Holland.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2005. "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital." *A.E.R.* 95 (1): 437–49.

Campbell, Frances, Gabriella Conti, James J. Heckman, et al. 2014. "Early Childhood Investments Substantially Boost Adult Health." *Science* 343 (6178): 1478–85.

Carneiro, Pedro, and Rita Ginja. 2014. "Long-Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start." *American Econ. J. Econ. Policy* 6 (4): 135–73.

Cascio, Elizabeth U. 2009. "Maternal Labor Supply and the Introduction of Kindergartens into American Public Schools." *J. Human Resources* 44 (1): 140–70.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yang. 2011. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR." *Q.J.E.* 126 (4): 1593–660.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner. 2014. "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility." Working Paper no. 19844 (January), NBER, Cambridge, MA.

Chevalier, Arnaud. 2007. "Parental Education and Child's Education: A Natural Experiment." IZA Discussion Paper no. 1153, Inst. Study Labor, Bonn.

Corak, Miles. 2013. "Income Inequality, Equality of Opportunity, and Intergenerational Mobility." *J. Econ. Perspectives* 27 (3): 79–102.

Cornelissen, Thomas, Christian Dustmann, Anne Raute, and Uta Schönberg. 2018. "Who Benefits from Universal Child Care? Estimating Marginal Returns to Early Child Care Attendance." *J.P.E.* 126 (6): 2356–409.
Cort, H. Russell, Jr., William D. Commins Jr., Kenneth L. Deavers, Ruth Ann O'Keefe, and James F. Ragan Jr. 1966. "Results of the Summer 1965 Project Head Start." Technical report, Office Econ. Opportunity, Washington, DC.
Cunha, Flavio, and James Heckman. 2007. "The Technology of Skill Formation." *A.E.R.* 97 (2): 31–47.
Currie, Janet, and Enrico Moretti. 2003. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings." *Q.J.E.* 118 (4): 1495–532.
Deming, David. 2009. "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start." *American Econ. J. Appl. Econ.* 1 (3): 111–34.
Duncan, Greg, Ariel Kalil, Susan E. Mayer, Robin Tepper, and Monique R. Payne. 2005. "The Apple Does Not Fall Far from the Tree." In *Unequal Chances: Family Background and Economic Success*, edited by Samuel Bowles, Herbert Gintis, and Melissa Osborne Groves, 23–79. Princeton, NJ: Princeton Univ. Press.
Dynarski, Susan, Joshua Hyman, and Diane Whitmore Schanzenbach. 2013. "Experimental Evidence on the Effect of Childhood Investments on Postsecondary Attainment and Degree Completion." *J. Policy Analysis and Management* 32 (4): 692–717.
East, Chloe N., Sarah Miller, Marianne Page, and Laura R. Wherry. 2019. "Multigenerational Impacts of Childhood Access to the Safety Net: Early Life Exposure to Medicaid and the Next Generation's Health." Working Paper no. 23810 (February), NBER, Cambridge, MA.
Fryer, Roland G., Jr., and Steven D. Levitt. 2004. "Understanding the Black-White Test Score Gap in the First Two Years of School." *Rev. Econ. and Statis.* 86 (2): 447–64.
———. 2006. "The Black-White Test Score Gap through Third Grade." *American Law and Econ. Rev.* 8 (2): 249–81.
Garces, Eliana, Duncan Thomas, and Janet Currie. 2002. "Longer-Term Effects of Head Start." *A.E.R.* 92 (4): 999–1012.
García, Jorge Luis, James J. Heckman, Duncan Ermini Leaf, and María José Prados. 2020. "Quantifying the Life-Cycle Benefits of an Influential Early-Childhood Program." *J.P.E.* 128 (7): 2502–41.
García, Jorge Luis, James J. Heckman, and Victor Ronda. 2021. "The Lasting Effects of Early Childhood Education on Promoting the Skills and Social Mobility of Disadvantaged African Americans." Working Paper no. 29057 (July), NBER, Cambridge, MA.
Gibbs, Chloe, Jens Ludwig, and Douglas L. Miller. 2011. "Does Head Start Do Any Lasting Good?" Working Paper 17452 (September), NBER, Cambridge, MA.
———. 2013. "Head Start: Origins and Impacts." In *Legacies of the War on Poverty*, edited by Martha J. Bailey and Sheldon Danziger, 39–65. New York: Russell Sage Found.
Gillette, Michael L. 1996. *Launching the War on Poverty: An Oral History.* New York: Twayne.
Greenberg, Polly. 1990. "Head Start—Part of a Multi-pronged Anti-poverty Effort for Children and Their Families . . . Before the Beginning: A Participant's View." *Young Children* 45 (6): 40–52.
Heckman, James J. 2007. "The Economics, Technology, and Neuroscience of Human Capability Formation." *Proc. Nat. Acad. Sci. USA* 104 (33): 13250–55.

Heckman, James J., and Ganesh Karapakula. 2019. "Intergenerational and Intragenerational Externalities of the Perry Preschool Project." Working Paper no. 2019-033, Human Capital and Econ. Opportunity Working Group, Dept. Econ., Univ. Chicago.
Heckman, James J., and Dimitriy V. Masterov. 2007. "The Productivity Argument for Investing in Young Children." *Rev. Agricultural Econ.* 29 (3): 446–93.
Heckman, James J., and Stefano Mosso. 2014. "The Economics of Human Development and Social Mobility." *Ann. Rev. Econ.* 6:689–733.
Heckman, James J., Rodrigo Pinto, and Peter Savelyev. 2013. "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes." *A.E.R.* 103 (6): 2052–86.
Hertz, Tom, Tamara Jayasundera, Patrizio Piraino, Sibel Selcuk, Nicole Smith, and Alina Verashchagina. 2008. "The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends." *B.E. J. Econ. Analysis and Policy* 7 (2): 10.
Hoynes, Hilary, Marianne Page, and Ann Huff Stevens. 2011. "Can Targeted Transfers Improve Birth Outcomes? Evidence from the Introduction of the WIC Program." *J. Public Econ.* 95 (7–8): 813–27.
Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond. 2016. "Long-Run Impacts of Childhood Access to the Safety Net." *A.E.R.* 106 (4): 903–34.
Johnson, Rucker C., and C. Kirabo Jackson. 2019. "Reducing Inequality through Dynamic Complementarity: Evidence from Head Start and Public School Spending." *American Econ. J. Econ. Policy* 11 (4): 310–49.
Kline, Patrick, and Christopher R. Walters. 2016. "Evaluating Public Programs with Close Substitutes: The Case of Head Start." *Q.J.E.* 131 (4): 1795–848.
Knudsen, Eric I., James J. Heckman, Judy L. Cameron, and Jack P. Shonkoff. 2006. "Economic, Neurobiological, and Behavioral Perspectives on Building America's Future Workforce." *Proc. Nat. Acad. Sci. USA* 103 (27): 10155–62.
Kose, Esra. 2021. "The Intergenerational Effects of Head Start on Infant Health." Manuscript.
Lee, Chul-In, and Gary Solon. 2009. "Trends in Intergenerational Income Mobility." *Rev. Econ. and Statis.* 91 (4): 766–72.
Levine, Robert A. 1970. *The Poor Ye Need Not Have with You: Lessons from the War on Poverty.* Cambridge, MA: MIT Press.
Ludwig, Jens, and Douglas L. Miller. 2007. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design." *Q.J.E.* 122 (1): 159–208.
Maurin, Eric, and Sandra McNally. 2008. "Vive la Révolution! Long-Term Education Returns of 1968 to the Angry Students." *J. Labor Econ.* 26 (1): 1–33.
Mazumder, Bhashkar. 2005. "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data." *Rev. Econ. and Statis.* 87 (2): 235–55.
Mazumder, Bhashkar, Maria Rosales-Rueda, and Margaret Triyana. 2019. "Intergenerational Human Capital Spillovers: Indonesia's School Construction and Its Effects on the Next Generation." *AEA Papers and Proc.* 109:243–49.
Miller, Sarah, and Laura R. Wherry. 2019. "The Long-Term Effects of Early Life Medicaid Coverage." *J. Human Resources* 54 (3): 785–824.
Montialoux, Claire. 2016. "Revisiting the Impact of Head Start." Policy brief, Inst. Res. Labor and Employment, Univ. California, Berkeley.
Murnane, Richard J., John B. Willett, Kristen L. Bub, and Kathleen McCartney. 2006. "Understanding Trends in the Black-White Achievement Gaps during

the First Years of School." In *Brookings-Wharton Papers on Urban Affairs*, edited by Gary Burtless and Janet Rothenberg Pack, 97–135. Washington, DC: Brookings Inst.

Nores, Milagros, Clive R. Belfield, W. Steven Barnett, and Lawrence Schweinhart. 2005. "Updating the Economic Impacts of the High/Scope Perry Preschool Program." *Educ. Evaluation and Policy Analysis* 27 (3): 245–61.

Office of Child Development. 1968. "Project Head Start 1965–1967: A Descriptive Report of Programs and Participants." Report no. ED 034 569, Dept. Health, Educ., and Welfare, Washington, DC.

———. 1970. "Project Head Start 1968: A Descriptive Report of Programs and Participants." Report no. ED 047 816, Dept. Health, Educ., and Welfare, Washington, DC.

Office of the President. 1965. Written statement accompanying President Lyndon B. Johnson's remarks on the launch of Project Head Start, May 18.

Oreopoulos, Philip, Marianne E. Page, and Ann Huff Stevens. 2006. "The Intergenerational Effects of Compulsory Schooling." *J. Labor Econ.* 24 (4): 729–60.

Page, Marianne E. 2004. "New Evidence on the Intergenerational Correlation in Welfare Participation." In *Generational Income Mobility in North America and Europe*, edited by Miles Corak, 226–44. Cambridge: Cambridge Univ. Press.

———. 2009. "Fathers' Education and Children's Human Capital: Evidence from the World War II G.I. Bill." Manuscript.

Puma, Michael, Stephen Bell, Ronna Cook, et al. 2012. *Third Grade Follow-Up to the Head Start Impact Study: Final Report.* Report 2012-45. Washington, DC: Dept. Health and Human Services.

Puma, Michael, Stephen Bell, Ronna Cook, and Camilla Heid. 2010. *Head Start Impact Study: Final Report.* Washington, DC: Dept. Health and Human Services.

Puma, Michael, Stephen Bell, Ronna Cook, Camilla Heid, and Michael Lopez. 2005. *Head Start Impact Study: First Year Findings.* Washington, DC: Dept. Health and Human Services.

Reardon, Sean F., and Ximena A. Portilla. 2016. "Recent Trends in Income, Racial, and Ethnic School Readiness Gaps at Kindergarten Entry." *AERA Open* 2 (3). https://doi.org/10.1177/2332858416657343.

Romano, Joseph P., and Michael Wolf. 2005. "Exact and Approximate Stepdown Methods for Multiple Hypothesis Testing." *J. American Statis. Assoc.* 100 (469): 94–108. https://doi.org/10.1198/016214504000000539.

Rossin-Slater, Maya, and Miriam Wüst. 2020. "What Is the Added Value of Preschool for Poor Children? Long-Term and Intergenerational Impacts and Interactions with an Infant Health Intervention." *American Econ. J. Appl. Econ.* 12 (3): 255–86.

Sawhill, Isabel V., Scott Winship, and Kerry Searle Grannis. 2012. "Pathways to the Middle Class: Balancing Personal and Public Responsibilities." Report, Center on Children and Families, Brookings Inst., Washington, DC.

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

Solon, Gary. 1999. "Intergenerational Mobility in the Labor Market." In *Handbook of Labor Economics*, vol. 3A, edited by Orley C. Ashenfelter and David Card, 1761–800. Amsterdam: North-Holland.

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

Vinovskis, Maris A. 2005. *The Birth of Head Start: Preschool Education Policies in the Kennedy and Johnson Administrations.* Chicago: Univ. Chicago Press.

Walters, Christopher R. 2015. “Inputs in the Production of Early Childhood Human Capital: Evidence from Head Start,” *American Econ. J. Appl. Econ.* 7 (4): 76–102.

White, Sheldon H., and Deborah A. Phillips. 2001. “Designing Head Start: Roles Played by Developmental Psychologists.” In *Social Science and Policy-Making: A Search for Relevance in the Twentieth Century*, edited by David L. Featherman and Maris A. Vinovskis, 83–118. Ann Arbor: Univ. Michigan Press, 83–118.

Zigler, Edward, and Sally J. Styfco. 2010. *The Hidden History of Head Start.* Oxford: Oxford Univ. Press.

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