

Reducing Inequality through Dynamic Complementarity: Evidence from Head Start and Public School Spending[†]

By RUCKER C. JOHNSON AND C. KIRABO JACKSON*

We compare the adult outcomes of cohorts who were differentially exposed to policy-induced changes in Head Start and K–12 spending, depending on place and year of birth. IV and sibling-difference estimates indicate that, for poor children, these policies both increased educational attainment and earnings, and reduced poverty and incarceration. The benefits of Head Start were larger when followed by access to better-funded schools, and increases in K–12 spending were more efficacious when preceded by Head Start exposure. The findings suggest dynamic complementarities, implying that early educational investments that are sustained may break the cycle of poverty. (JEL H52, H75, I21, I26, I28, I32, I38)

Children born to less advantaged households and communities typically experience lower levels of educational attainment, employment, earnings, health, and well-being as adults than children born to more advantaged ones (Chetty et al. 2014). Differences between individuals from more and less advantaged backgrounds manifest early in childhood and tend to grow as children age (Fryer and Levitt 2006, Currie and Thomas 2001, McLeod and Kaiser 2004, Heckman and Mosso 2014). Accordingly, remediating the ill effects of childhood poverty may require early investments in the skills of disadvantaged children that are followed by sustained investments over time.

This paper studies whether early childhood investments designed to promote school readiness among disadvantaged children that are followed up with increases in public school spending are particularly effective at improving their long-run outcomes. This question is one specific manifestation of the long-standing

* Johnson: Goldman School of Public Policy, University of California, Berkeley, 2607 Hearst Ave, Berkeley, CA 94720 and NBER (email: ruckerj@berkeley.edu); Jackson: School of Education and Social Policy, Northwestern University, 2120 Campus Drive, Evanston, IL 60208 and NBER (email: kirabo-jackson@northwestern.edu). John Friedman was coeditor for this article. Johnson and Jackson are equal authors on this paper. We wish to thank the PSID staff for access to the confidential restricted-use PSID geocode data. We are grateful to Edward Zigler and seminar participants at the NBER education/children's meetings, the University of Michigan, UC-Berkeley, Penn, Brown, Vanderbilt University, Columbia, University of Wisconsin/IRP Summer Workshop, and UC-Davis for helpful comments; and thank Martha Bailey and Andrew Goodman-Bacon; and Doug Miller and Jens Ludwig, for sharing data on 1960 county poverty rates and data from the National Archives & Records Administration. This research was supported by the National Institutes of Health and the Russell Sage Foundation (to Johnson), and the National Science Foundation award number 1324778 (to Jackson). The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

[†] Go to <https://doi.org/10.1257/pol.20180510> to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

hypothesis in economics that, because skills beget skills, children who benefit from early human capital investments may benefit more from later investments (Cunha and Heckman 2007). Testing this hypothesis is difficult, since it requires exogenous variation in multiple investments over time, a high bar that many previous papers have struggled to clear. Some earlier studies have examined whether the effect of human capital intervention varies by pre-intervention skill level (e.g., García and Gallegos 2017, Lubotsky and Kaestner 2016, Aizer and Cunha 2012). Because these studies do not use exogenous variation in prior skills, they do not speak directly to whether early and late human capital policies exhibit dynamic complementarity. Addressing this critique, other studies have examined whether the benefits of human capital investments vary among those who were exposed to non-investment skill shocks such as hurricanes, or rainfall during gestation (e.g., Adhvaryu et al. forthcoming). Because these studies do not rely on human capital investments per se for variation in initial skills, they do not examine whether human capital investments made at different stages of the life course exhibit dynamic complementarity.

To test for dynamic complementarity, we examine the interaction between two exogenous and independent human capital investment “shocks.” The first exogenous shock to human capital investment is the rollout of Head Start, the largest early childhood intervention program in the United States, which increased access to early childhood education and pediatric care for low-income children. The second exogenous shock to human capital investment is the implementation of court-ordered school finance reforms (SFRs), which reduced differences in public K–12 school spending between affluent and poor neighborhoods within states, and increased (on average) the level of per pupil spending at public K–12 schools.¹ While dynamic complementarities may not exist between *any* two human capital investments, they are most likely to exist between two education-related interventions in which one (a school-readiness program) is designed to help children benefit from the other (public K–12 schools). Our setting is particularly well suited for studying dynamic complementarities.²

To isolate the effects of these two major policies, we exploit temporal and geographic variation in exposure to these policy-induced investment “shocks” and analyze the life trajectories of individuals born between 1950 and 1976, and followed through 2015 using the Panel Study of Income Dynamics (PSID). These data allow us to study potential complementarities on an array of adult outcomes including educational attainment, earnings, poverty, and incarceration. While

¹ See Card and Payne (2002); Murray, Evans, and Schwab (1998); Hoxby (2001); and Jackson, Johnson, and Persico (2014) for a more complete discussion of the effects of SFRs on public school spending.

² Investments are “broadly defined as actions specifically taken to promote learning” in Heckman and Mosso (2014). As such, interactions between education-related interventions are likely what the theory is about. In the only other paper to examine two education interventions, Gilraine (2016) finds that the benefits of accountability due to No Child Left Behind in later grades are larger among students’ exposure to accountability in earlier grades. Looking at a health and an education intervention, Rossin-Slater and Wüst (2017) finds that the effects of access to preschool was *smaller* among those who had access to home visits during infancy. In work with a health intervention, Bhalotra and Venkataramani (2015) finds that the benefits of antibiotic treatment for blacks decline in measures of the severity of institutionalized segregation. Also Malamud, Pop-Eleches and Urquiola (2016) examines whether the benefits of attending a better school vary by parental access to abortion near the time of conception.

test scores have been the traditional focus of evaluations of Head Start and K–12 spending, the effects of interventions on long-run outcomes may go undetected by test scores.³

Identifying the interaction effects between two human capital investments requires that one credibly identify the effects of each investment individually on the same individuals and that each of the human capital investments is independent of the other (Almond and Mazumder 2013). To identify the causal effect of early childhood investments, we exploit variation in the timing of the rollout of Head Start across counties. We compare the adult outcomes of individuals who were from the same childhood county but were exposed to different levels of Head Start spending, because some were four years old when Head Start spending levels were low (or nonexistent) while others were four years old when Head Start spending levels were higher. To identify the causal effects of public K–12 school spending, we exploit geographic variation in the timing of court-ordered SFRs. Following Jackson, Johnson, and Persico (2016)—henceforth, JJP—we predict the spending change that each district would experience after the passage of a court-mandated SFR based on the type of reform and the characteristics of the district before reforms. Using instrumental variables models, we examine whether SFR-exposed cohorts (young enough to have been in school during or after a SFR) have better outcomes relative to SFR-unexposed cohorts (those who were too old to be affected by a SFR) in districts predicted to experience larger reform-induced spending increases. We present empirical tests to validate our models and to support a causal interpretation of the patterns presented.

To explore the relationship between early- and later-childhood human capital investments, we combine both identification strategies to estimate the effects of the interaction between the two. Some districts experienced increases in school spending due to a SFR when Head Start was available in the county, while other districts experienced similar K–12 spending increases when Head Start was not available. This fact allows one to test if the effects of K–12 spending increases due to SFRs are higher with greater public pre-K investments than without them. Similarly, Head Start was rolled out in different counties both before and after the local school districts experienced increases in K–12 spending due to SFRs. This fact allows one to test if the effects of Head Start spending are larger in areas that have higher levels of K–12 spending due to the passage of a court-ordered SFR.

For the interaction effect to be identified, it requires that individuals that are exposed to both a SFR and Head Start are not somehow different from those that are exposed to only a SFR or only exposed to Head Start. One can only be confident that this condition is satisfied if both policy changes are independent of each other. We argue that because SFRs occurred at the state level (affecting all public schools in a state at the same time), while Head Start (a federal program) was introduced in certain counties within states at different times, these two policies are largely independent of each other. More formally, we show that the raw correlation between

³E.g., Heckman, Pinto, and Savelyev 2013; Jackson 2018; Chetty et al. 2011; Ludwig and Miller 2007.

the two policy instruments is only 0.15; conditional on controls, there is no association between Head Start spending and SFR-induced changes in K–12 spending; and using partial *F*-statistics, there is sufficient policy variation in Head Start spending and K–12 spending for the effect of each to be identified and for the interaction between the two to be identified.

Our results show that both Head Start and SFR-based K–12 spending increases have large, positive long-run effects, and we also find strong and robust evidence of dynamic complementarity. For children from low-income families, on average, increases in Head Start spending increased educational attainment and adult earnings and reduced the likelihood of both poverty and incarceration in adulthood. We find no effect of Head Start spending on the outcomes of non-poor children. Increases in public school K–12 spending improved this same array of outcomes in adulthood. Among poor children exposed to a 10 percent reduction in K–12 spending, exposure to a typical Head Start center has small statistically insignificant effects on educational attainment, wages, incarceration, and adult poverty. However, among poor children exposed to a 10 percent increase in K–12 spending, exposure to a typical Head Start center leads to 0.59 additional years of education, being 14.8 percentage points more likely to graduate high school, 17 percent higher wages, being 4.7 percentage points less likely to be incarcerated, and being 12 percentage points less likely to be poor as an adult.

The fact that the long-run benefits of Head Start spending depend on the subsequent level of K–12 spending may help explain why some studies find positive effects of Head Start and others do not.⁴ Looking at the marginal effects of K–12 spending, for low-income children, increasing public K–12 spending by 10 percent has small effects on educational attainment, adult wages, and incarceration when not preceded by Head Start. However, among low-income children exposed to Head Start, that same 10 percent increase in K–12 per pupil spending increases educational attainment by 0.4 years, increases earnings by 20.6 percent, and reduces the likelihood of incarceration by 8 percentage points. The positive interaction effects between Head Start and K–12 spending are robust across several models (including sibling comparisons) and are only present among poor children (who were eligible for Head Start). The effect of K–12 spending was unrelated to the level of Head Start spending among non-poor children, for whom increasing K–12 spending by 10 percent increased years of education by 0.2 and earnings by 11.7 percent.

The paper contributes to the literature in three ways. First, we provide direct evidence on the long-run benefits of both Head Start and K–12 spending. Second, we present broad, robust evidence of complementarities between early and later human capital investments for low-income children. The complementarities imply that one could increase *both* equity and efficiency by redistributing spending from well-funded K–12 schools toward Head Start programs targeted at poor children. Generally, our results are the first to show that early and sustained complementary investments in the skills of low-income children can be a cost-effective strategy

⁴For positive effects, see Deming (2009); Ludwig and Miller (2007); Garcés, Thomas, and Currie (2002); Carneiro and Ginja (2014). For mixed effects, see Zigler, Gilliam, and Barnett (2011) and Lipsey, Farran, and Hofer (2015).

to break the cycle of poverty. Third, the use of quasi-experimental methods that involve two different, yet complementary, identification strategies yields a similar pattern of results and bolsters confidence in the overall set of findings.

The rest of the paper is organized as follows. Section I outlines our theoretical framework. Section II describes the Head Start program and court-ordered school finance reforms. Section III presents the data used. Section IV describes the empirical strategy. Section V presents the results. Section VI presents conclusions and a summary discussion.

I. Theoretical Framework

Research in developmental neuroscience highlights the importance of the preschool years in establishing the building blocks of subsequent human capital formation and the interconnectedness of cognitive, noncognitive, and health formation (Shonkoff and Phillips 2000). Informed by this research, Cunha and Heckman (2007) theorizes that skill development is an interactive, multistage process in which the marginal effect of investments today is higher among those with a greater stock of previously acquired skills.⁵ We refer to this characteristic of skills production as dynamic complementary *in skill development*. When this condition holds, “skills produced at one stage raise the productivity of investment at subsequent stages” (Cunha and Heckman 2007). We refer to such synergies between *investments* as dynamic complementary *in human capital investments*.⁶ If Head Start increases skills and therefore improves school readiness, Head Start may facilitate better learning in the K–12 system. If so, insofar as increased spending improves school quality, spending on Head Start and public K–12 schools would exhibit dynamic complementarity. This is what we seek to test in this paper.

Note that complementarity is not a given. Compensatory interventions or interventions designed to bring *all* children up to some basic standard of

⁵ This is what is identified by researchers who examine the effect of interventions for individuals with differing incoming levels of skills (e.g., García and Gallegos 2017, Lubotsky and Kaestner 2016, and Aizer and Cunha 2012).

⁶ Following the notation from Heckman (2007), the technology of skills production is dynamic. Skills acquired when a child is t years old is (a) below:

$$(a) \quad \theta_{t+1} = f_t(h_t, \theta_t, I_t),$$

where $t = 1, 2, \dots, T$, θ_t is a vector of skills at time t , parental capabilities are connoted by h_t , and investments during time t are connoted by I_t . Investments in time t (I_t) are construed broadly to include parental investments, schooling inputs (i.e., peers, teachers, etc.), and neighborhood and community inputs. For analytical convenience, f_t is assumed to be strictly increasing in I_t . Dynamic complementarity in human capital investments arises when the stocks of capabilities acquired by period $t - 1$ (θ_t) make investments in period t (I_t) more productive, i.e.,

$$(b) \quad (\partial^2 \theta_{t+1}) / (\partial \theta_t \partial I_t) > 0.$$

Consider that $\theta_t = f_{t-1}(h_{t-1}, \theta_{t-1}, I_{t-1})$. Because $\partial f_t / \partial I_t > 0$, if (b) holds, then (c) below must also hold:

$$(c) \quad (\partial^2 \theta_{t+1}) / (\partial I_{t-1} \partial I_t) > 0.$$

In words, dynamic complementarity in skill development implies that there is dynamic complementarity in human capital investments. However, if early investments increase the efficacy of later investments through mechanisms other than increasing skills, the converse may not hold. We show that this is not the case in our setting.

skill, may, by design, have smaller benefits for more highly skilled children. Also, note that human capital investments may exhibit dynamic complementarity for reasons other than dynamic complementary in skill development.⁷ To apply these insights to our setting, we outline two ways through which Head Start and K–12 spending may interact. The first is a direct channel that operates through dynamic complementarity in skill development. The second channel is indirect and may operate through spillovers to other students and adjustments by actors in the schooling system (Malamud, Pop-Eleches, and Urquiola 2016).

The direct channel operates through what we call “alignment.” The alignment mechanism is predicated on the idea that the sequence of when skills are taught matters (Knudsen et al. 2006, Newport 1990, and Pinker 1994) and the fact that K–12 systems target students with a specific incoming skill level. Students above the target skill level may benefit less from the K–12 system (the K–12 system may spend valuable instructional time teaching skills they have already mastered), and students below this target incoming skill level may benefit less from the K–12 system (the instruction may require skills they do not possess). Given that poor children, on average, are less likely to be school-ready at kindergarten entry (Fryer and Levitt 2004, Magnuson and Waldfogel 2005), Head Start spending, by increasing their skills, may bring them closer to the target such that they benefit more from subsequent investments experienced in the K–12 education system. Furthermore, access to pediatric care (provided to Head Start participants) may promote this skill development (Levine and Schanzenbach 2009, Cohodes et al. 2015).

Through alignment, Head Start spending increases may not improve outcomes to the same degree in all contexts. In fact, in poorly funded schools that may align instruction to a low-target skill level, Head Start participation could reduce alignment with the target level by increasing students’ incoming skills above the target. In such a scenario, relative to their peers who did not attend preschool, any advantage in skill created by Head Start will diminish over time as children who attended Head Start receive redundant instruction, and their peers who lack access to preschool catch up in elementary school grades. That is, there may be fadeout and lower long-run Head Start effects for program participants who attend poorly funded K–12 schools. In sum, through this channel, on average, the effects of Head Start spending on poor children may be larger in well-funded K–12 districts and could be negligible in poorly funded public school districts.

The first indirect channel is through “spillovers.” Research has found that higher shares of low-performing peers or disruptive peers may have deleterious impacts on students (see Sacerdote 2014). By increasing the human capital of poor children, increases in Head Start spending may affect the subsequent peer composition of the K–12 classrooms for *all* children in the county. This could make it easier for the K–12 school system to translate resources into better outcomes.⁸ *The second indirect channel* is through “adjustments.” The first is an “alignment adjustment.” If teachers in the K–12 system alter the alignment of their instruction toward an

⁷Formally, if equation (b) holds it implies that (c) will also hold. However, the converse is not always true.

⁸Neidell and Waldfogel (2010) provide evidences of this channel by documenting spillover effects from preschool between Head Start and non-Head Start children on math and reading achievement.

incoming higher ability student (in light of a lower share of low-achieving students due to Head Start), the quality of K–12 instruction could be affected for all students. Importantly, because these adjustments can move some students closer to the target and others farther from it, the “alignment adjustment” effect could be positive or negative for any given student. There could also be “budget allocation adjustments” that can affect students in different classrooms. For example, lower shares of students requiring remediation or special services (due to Head Start) may allow schools to allocate resources to other productive inputs, which may affect all students in the school.⁹

This is not an exhaustive list of all possible adjustment effects. However, the key takeaways are that policy complementarities reflect both the direct effect due to the technology of skill formation and also some spillover and adjustment effects, and adjustment and spillover effects could lead the interaction between the two interventions to be either positive or negative such that the overall interaction effect is ambiguous in sign. We present empirical evidence to shed light on what mechanisms are most likely at play in our setting.

II. Background and Overview of Head Start and School Finance Reforms

A. Background on Head Start

Head Start was established in 1964 as part of President Lyndon B. Johnson’s “War on Poverty,” and is a national, federally funded, early childhood program with the aim of improving the human capital of poor children. The Head Start curriculum aims to enhance literacy, numeracy, reasoning, problem-solving, and decision-making skills. Head Start includes educational efforts for both parents and children to enhance nutrition in the home and provides nutritious meals for the children. Participating children receive development screenings, and programs connect families with medical, dental, and mental health services. Head Start also provides first-time parents with parenting strategies (Zigler, Gilliam, and Barnett 2011). Head Start currently operates more than 19,200 centers and serves more than 900,000 children. Current Head Start expenditures average about \$8,700 per enrolled child (in 2015 dollars). This level of per pupil spending is much lower than those at model preschool programs such as Perry Preschool or Abecedarian (Blau and Currie 2006).

Because we seek to explore the effects of Head Start spending on longer run adult outcomes (among those who are adults today), we study the effects of Head Start at the inception of the program (1965 through 1980). Head Start was initially launched as an eight-week, summer-only program in 1965 and then became a primarily part-day, nine-month program in 1966. Head Start is mainly funded federally.¹⁰ To open a new Head Start center, local organizations

⁹ Head Start also teaches parenting skills; thus, another possible indirect channel is changes in parental quality. We test for this using within-family variation. We find no indication that siblings of those exposed to Head Start have improved outcomes (online Appendix I). This runs counter to the parental quality mechanism.

¹⁰ Head Start funds were allocated to states proportionately based upon each state’s relative number of children living in families with income below the poverty line and the relative number of public assistance recipients in each

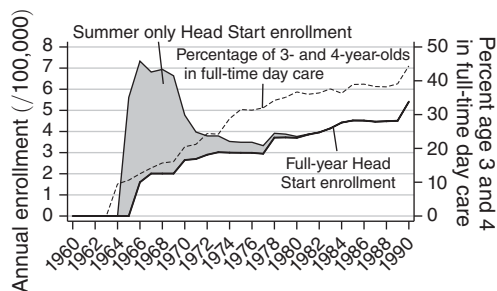
(typically nonprofit organizations, for-profit agencies, or school systems) apply to the federal government for grant funds. Grantees provide at least 20 percent of the funding. After approval, Head Start grants are awarded directly to applying organizations subject to three-year grant cycles. Each grantee must comply with student-to-teacher ratio guidelines and other standards outlined in the Head Start Act. During the first 15 years of the program, the average student-to-teacher ratio in a Head Start classroom was roughly 17:1 (Zigler and Styfco 2010). During this early era of the program, the majority of Head Start children were enrolled in part-day centers (as opposed to full-day programs, which are six or more hours per day such as Abecedarian), and often part-year (GAO 1981).

Head Start was targeted at preschool age children (three through five) and most Head Start enrollees were four years old at enrollment. At each center, at least 90 percent of enrollees had to be from families whose income was below the federal poverty line, and at least 10 percent of children had to have a disability. Figure 1, panel A, plots the raw national Head Start enrollments between 1960 and 1990. Between 1965 and 1970, most of the enrollment in Head Start was in summer-only programs. However, from 1972 and after that, most enrollment was in full-year Head Start. As such, the early rollout of Head Start represented *both* increases in Head Start participation *and* enhancements in the Head Start programs themselves. Another notable pattern is the decline in Head Start enrollments between 1969 and 1972. During this period, full-year Head Start programs enrollment was increasing at the same time that summer-only program enrollment was declining (somewhat more rapidly). To relate these enrollments to participation rates at the individual child level, for each kindergarten entry cohort we computed the cumulative likelihood across all age-eligible years that an income-eligible child would enroll in Head Start.¹¹ Figure 1, panel B, depicts our estimated likelihood of Head Start enrollment (across all age-eligible years) by kindergarten entry cohort. The likelihood of Head Start enrollment among poor children reached 86 percent for income-eligible cohorts entering kindergarten in 1969, fell in the early 1970s, and stabilized around 63 percent by 1990. This is similar to the Garces, Thomas, and Currie (2002) estimate of two-thirds. Our participation rates of between 63 and 85 percent are important to keep in mind as we interpret the magnitudes of our intent-to-treat estimates (in Section V). Figure 1 (panel A) also plots the share of three- and four-year-olds enrolled in full-time daycare over time (as reported in the Current Population Survey). This figure highlights that Head Start rollout coincides with a period in which most children were not in formal, full-time preschool, and also coincides with a general increase in the proportion of children ages 3 to 4 enrolled in full-time preschool. In the context of the estimated effects of Head Start during this rollout period, the

state. Head Start in collaboration with the Medicaid Early Pediatric Screening, Diagnosis, and Treatment Program (EPSDT) provided comprehensive prevention and treatment services to preschool children.

¹¹ The ratio of enrolled students to the income-eligible age-eligible population in a given year *is not* the same as a specific cohort's participation rate by kindergarten entry. See online Appendix M for a more detailed discussion of this. To avoid double-counting individuals who enrolled in both the summer program and the full-year programs, we assume that 40 percent of full-year enrollees were previously in a summer program.

Panel A. National Head Start enrollment and day care enrollment



Panel B. Probability of participating in any Head Start by kindergarten entry year (income-eligible children only)

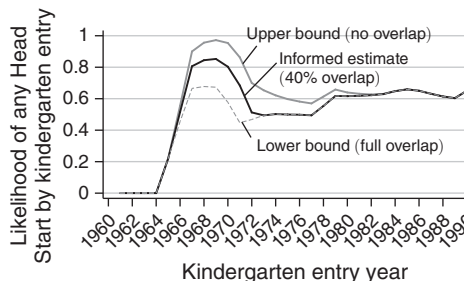


FIGURE 1. NATIONAL ENROLLMENT IN HEAD START OVER TIME

Notes: National counts of three-, four-, and five-year-olds are derived from Integrated Public Use Microdata Series (IPUMS) decennial censuses from 1960–1990. Counts for non-census years are completed using linear interpolation. The percentage of children ages three and four who are enrolled in full-year day care are as reported in the Current Population Survey (<http://www.census.gov/hhes/school/data/cps/historical/>). Head Start enrollment figures are from the Head Start fact sheet (<https://eclkc.ohs.acf.hhs.gov/hslc/data/factsheets/2015-hs-program-factsheet.html>). In panel B, the participation rate is the cumulative probability of enrolling in Head Start across all age-eligible years prior to kindergarten entry (note: this is not the same as the fraction of eligible enrollees in a given year, but is the sum of these annual probabilities across age-eligible years prior to kindergarten entry). The upper bound assumes that each enrollee in full years and summer only programs is unique (no overlap). The lower bound assumes that all full-year enrollees, where possible, were also in a summer program (full overlap). The informed estimates assume that, where possible, 40 percent of the full-year enrollees were previously summer enrollees.

counterfactual option in the early years is primarily home care, as opposed to some other full-time pre-K program.

We use Head Start spending as a way to measure both the presence of the program and also the quality, size, and extent of the program. While Head Start spending per enrollee may seem like a natural proxy for quality, such a measure fails to capture changes in spending that work through expansions in access.¹² Because the target population for Head Start is poor preschoolers and most enrollees are four years old, our measure of Head Start spending is federal Head Start spending per *poor* four-year-old in the county. Between 1965 and 1980, the average county with a Head Start center spent about \$4,000 per poor child and about \$5,300 per enrollee (in year 2000 dollars). There is considerable variation in timing of the establishment of Head Start centers. However, in most counties, the *first* Head Start center was established between 1965 and 1970.¹³ The geographic variation in the timing of the rollout of Head Start is central to our empirical strategy to isolate exogenous variation in Head Start spending across birth cohorts within a county.

¹² See online Appendix B for an illustrated discussion of this.

¹³ Online Appendix Figure A2 presents each county in the United States color-coded by the year of its first Head Start center.

B. Background on School Finance Reforms

The other major human capital interventions we study are the increases in public K–12 school spending caused by court-ordered school finance reforms (SFRs). In most states, before the 1970s, local property taxes accounted for most resources spent on K–12 schooling (Howell and Miller 1997). Because the local property tax base is typically higher in areas with higher home values, and there are high levels of residential segregation by socioeconomic status, heavy reliance on local financing contributed to affluent districts' ability to spend more per student. In response to large within-state differences in per pupil spending across wealthy/high-income and poor districts, state supreme courts overturned school finance systems in 28 states between 1971 and 2010. Because of these court decisions, many states implemented legislative reforms that led to important changes in public education funding.¹⁴ Most of these court-ordered SFRs changed the parameters of spending formulas to reduce inequality in school spending and weaken the relationship between per pupil school spending and the wealth and income level of the district.

The effect of a SFR on school spending depends on the type of school funding formula introduced by the reform and how the funding formula introduced interacts with the specific characteristics of a district. We follow JJP and categorize reforms into four types. *Foundation plans* guarantee a base level of per pupil spending and increase per pupil spending for the lowest-spending districts. *Spending-limit* plans prohibit per pupil spending levels above some predetermined amount. *Reward-for-effort* plans match locally raised funds for education with additional state funds (often with higher match rates for lower income areas). *Equalization plans* typically tax all districts and redistribute funds to lower wealth and lower income districts. These reform/formula types are not mutually exclusive.

In existing work, Card and Payne (2002), JJP, and Hoxby (2001) find that court-ordered SFRs that lead to the implementation of different funding formulas have different effects on district spending by pre-reform income and spending levels.¹⁵ In particular, JJP finds that reforms that lead to "reward-for-effort" formulas tended to increase per pupil K–12 spending in all districts; spending limits led to pronounced spending reductions in high-spending districts; foundation plans led to the largest spending increases in low-income districts; and equalization plans were more equalizing by pre-reform spending levels than by pre-reform income levels. These systematic patterns allow us to predict how much K–12 school spending increases in each district as a function of the reform type introduced (by the state) and the pre-reform characteristics of the district. Because these relationships are unrelated to decisions made by individual districts or demographic shifts that may affect public school spending levels, we can use this prediction to isolate the causal relationship between reform-induced K–12 spending increases and students' longer run outcomes.

¹⁴ See Jackson, Johnson, and Persico (2016) for a full discussion of SFRs.

¹⁵ To illustrate how the introduction of different formula types affected districts by pre-reform income and spending levels, we replicate the analysis in Jackson, Johnson, and Persico (2016). This is in online Appendix C.

III. Data

We compiled data on annual Head Start spending at the county level, and public K–12 school spending at the school district level. The Head Start spending data come from the National Archives Record Administration, Inter-university Consortium for Political and Social Research, and Surveillance, Epidemiology, and End Results population data. These are combined to form a county-level panel of Head Start spending per poor four-year-old in the county between 1965 and 1980.¹⁶ Public K–12 education funding data come from several sources that are combined to form a panel of per pupil spending for US school districts in 1967 and annually from 1970 through 2000 and are linked to a database of SFRs from JJP.¹⁷ To avoid confounding nominal with real changes in spending, we convert both Head Start and K–12 school spending across all years to 2000 dollars using the Consumer Price Index (CPI).

Our individual-level data on long-run outcomes come from the Panel Study of Income Dynamics (PSID 1968–2015), and our analysis sample includes individuals born between 1950 and 1976 who were followed into adulthood. These PSID cohorts straddle both the rollout of Head Start programs across the country and the implementation of the early waves of court-ordered SFRs. We include all information on PSID individuals between 1968 and 2015.¹⁸ We linked persons in the PSID using their census blocks during childhood to school spending data, SFR data, and Head Start spending data. We then match the earliest available childhood residential address to the school district boundaries that prevailed in 1969 to avoid complications arising from endogenously changing district boundaries over time. We detail the algorithm in Appendix D. Among potentially treated cohorts, 97 percent of the earliest address information is from *before* the policies we study were enacted so that bias due to residential sorting in response to the policies is negligible. We verify this empirically. We also merge in county-level characteristics from the 1960 census, and information on the timing of other key policy changes during childhood (e.g., school desegregation, hospital desegregation, Title I, roll out of other “War on Poverty” initiatives, and expansion of safety net programs—described in Section IV) from multiple data sources.¹⁹

We define low-income children as those whose average parental income (between ages 12 and 17) fell in the bottom quartile.²⁰ Among cohorts born between 1963–1976 for whom parental income at age four is observed, roughly 80 percent of those whom we classify as low income were below the federal poverty line at age four, and 93 percent of those who were below the poverty

¹⁶ Further details on the PSID data are in online Appendix D.

¹⁷ Details on how these databases were compiled and the coverage of districts in these data are in online Appendix E.

¹⁸ We include both the Survey Research Center component and the Survey of Economic Opportunity component, commonly known as the “poverty sample,” of the PSID sample. The PSID maintains high wave-to-wave response rates of 95–98 percent. We perform a supplementary analysis of sample attrition in the PSID, and find no evidence of selective attrition among our study sample (online Appendix Table D1).

¹⁹ See online Appendices D and E for a discussion of these data sources.

²⁰ Because the earliest year in which parental income is available is 1967 due to when the PSID data collection started, we cannot observe family income at age four for those born before 1963. However, we can observe average family income during adolescence (ages 12 through 17) for all individuals in our analytic sample, which serves as a good permanent income measure. We use this to form our group of likely Head Start eligible individuals.

threshold at age four are classified as low income by our definition. The analytic sample includes 15,232 individuals from 4,990 childhood families, 1,427 school districts, 1,120 counties, across all 50 states. From this point forward, we refer to children who are low income as “poor” children, and those not from low-income families (as defined above) as “non-poor” children. We examine a broad range of adult outcomes. These include educational outcomes—whether graduated from high school, years of completed education; labor market and economic status outcomes (in real 2000 dollars)—log wages, family income, annual incidence of poverty in adulthood²¹ (ages 20–50); and criminal involvement and incarceration outcomes—whether ever incarcerated (jail or prison) and the annual incidence of incarceration in adulthood. Table 1 contains descriptive statistics for various childhood measures and adult outcomes in our analytic sample.

IV. Empirical Strategy

A. Identifying the Effects of Head Start Spending

Our measure of Head Start spending is total federal Head Start spending in a county per poor four-year-old (in 2000 CPI-adjusted real dollars). We take advantage of the staggered introduction across geographic areas of Head Start programs during the program’s rollout. Before the rollout of Head Start to an area, there is no Head Start spending. After the introduction of Head Start in a county, spending levels subsequently increase. Figure 2 shows an event-study plot of Head Start spending per poor four-year-old before and after rollout in areas that had high and low Head Start spending in 1980 (the end of the sample period under study). Note that year “zero” is the year of the establishment of the first Head Start center in a county.

Once the first center is established, spending per poor four-year-old increases more rapidly in the high- than the low-spending counties (panel A of Figure 2). Almost all counties experienced a transitory increase in Head Start spending due to the ubiquitous introduction of summer-only programs, which falls over time. However, high-spending counties expanded enrollment (and spending) in full-year programs that was sustained over time, while the low-spending counties did not and reverted to near zero Head Start spending within four years. If higher levels of Head Start spending improve outcomes, one should observe that the post-rollout cohorts should have better outcomes than the pre-rollout cohorts, and improvements between pre- and post-rollout cohorts should be larger in counties with larger sustained increases in Head Start spending. Figure 2 reveals exactly this pattern for years of educational attainment (measured in adulthood) among poor children. Areas with small (panel B) and large increases in Head Start spending (panel C) were on similar trajectories among cohorts who were older than four years old when the first Head Start center was established (i.e., years 5 through

²¹ Based on the family income-to-needs ratio and federal poverty thresholds by family structure and household size.

TABLE 1—SUMMARY STATISTICS OF THE ANALYTIC DATASET

	All (observations = 15,232)	Poor child (observations = 6,373)	Non-poor child (observations = 8,859)
<i>Adult outcomes</i>			
High school graduate	0.85	0.71	0.89
Years of education	13.29	12.29	13.61
ln(wages), at age 30	2.49	2.24	2.56
Adult family income, at age 30	\$48,655	\$35,372	\$52,448
In poverty, at age 30	0.08	0.18	0.05
Ever incarcerated	0.05	0.08	0.04
Age (range: 20–50)	30.8	30.3	31.0
Year born (range: 1950–1976)	1962	1962	1962
Female	0.44	0.43	0.44
White	0.87	0.66	0.93
<i>Childhood school variables</i>			
Any Head Start center in county, age 4	0.33	0.33	0.34
Post-rollout: Head Start spending per poor 4-year-old, age 4	\$4,103	\$4,204	\$4,072
Child attended Head Start ^a	0.04	0.19	0.02
Child attended any preschool program	0.23	0.31	0.23
School district per pupil spending (average, ages 5–17)	\$4,366	\$4,031	\$4,470
Any court-ordered school finance reform, ages 5–17	0.13	0.11	0.14
Conditional on any: number of exposure years to school finance reform	7.37	6.90	7.50
1960 district poverty rate (percent)	21.52	28.25	19.35
<i>Childhood family variables</i>			
Income (average, ages 12–17)	\$54,488	\$22,520	\$65,130
Income-to-needs ratio (average, ages 12–17)	3.05	1.31	3.62
Mother's years of education	11.84	10.61	12.24
Father's years of education	11.82	10.04	12.36
Born into two-parent family	0.90	0.74	0.95
Low birth weight (<5.5 pounds)	0.07	0.07	0.07

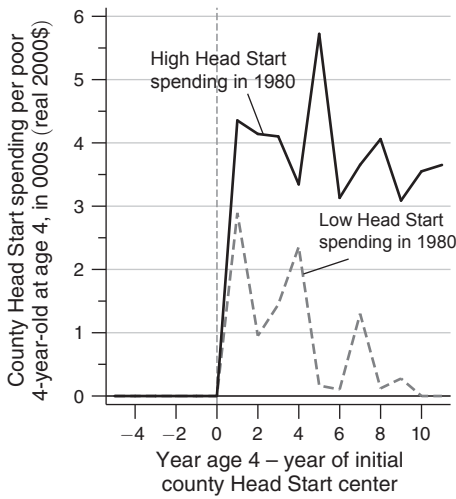
Notes: All descriptive statistics are sample weighted to produce nationally representative estimates of means. Dollars are CPI-U deflated in real 2000 dollars. “Poor child” is defined here as children whose parents were in the bottom quartile of the income distribution (approximately 80 percent of whom were below the poverty line). Analysis sample includes 15,232 individuals (218,594 person-year observations ages 20–50), from 4,990 childhood families, 1,427 school districts, 1,120 childhood counties and all 50 states.

^a Child-specific pre-K attendance and Head Start program participation info collected retrospectively in 1995 survey IW.

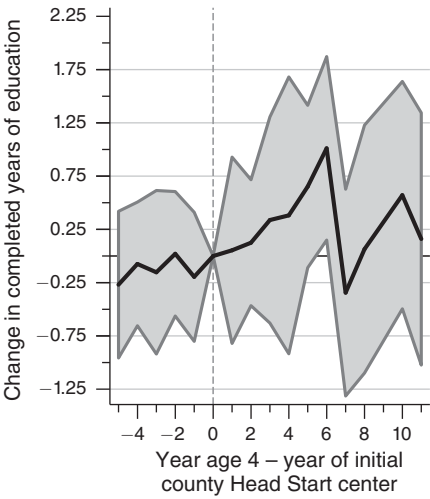
year 0). However, the post-rollout cohorts have much better outcomes in Head Start high-spending counties than in low-spending counties.

Our preferred difference-in-difference (DiD) strategy uses this variation in timing and dosage. That is, we compare the differences in long-run outcomes across birth cohorts from the same childhood county that experienced larger increases in Head Start spending at age four to the differences in outcomes across the same birth cohorts within other childhood counties that experienced small (or no) increases in

Panel A. Head Start spending:
By Head Start spending in 1980



Panel B. Educational attainment:
Low Head Start spending



Panel C. Educational attainment:
High Head Start spending

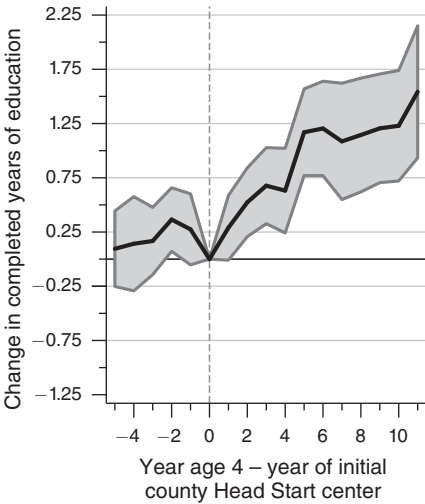


FIGURE 2. EVOLUTION ON HEAD START SPENDING AND EDUCATIONAL ATTAINMENT AT ROLLOUT (POOR CHILDREN)

Notes: Analysis sample includes PSID individuals born 1950–1976 who have been followed into adulthood. “High Head Start spending” is defined here as counties in the top quartile of Head Start spending among all US counties after rollout; “Low Head Start spending” is defined here as bottom quartile of Head Start spending among all US counties after rollout or no spending. Results are based on event study models of educational attainment on children’s exposure to county Head Start spending per poor four-year-old at age four as a function of the timing of the rollout of the program in the county. The figures present the event-study plots for both high- and low-spending counties (in 1980). The shaded gray region in the event study plots for years of education depict the 90 percent confidence interval for each event-year. The models include childhood county fixed effects, race $\times$ census division-specific birth-year trends; controls for 1960 county characteristics (poverty rate, percent black, education, percent urban, population size) each interacted with linear cohort trends; and controls for childhood family characteristics (parental income and education, mother’s marital status at birth, birth weight, gender).

Head Start spending at age four. These DiD-type comparisons are implemented in a regression framework by estimating (1) by Ordinary Least Squares (OLS):

$$(1) \quad Y_{icb} = \beta^{DiD} \cdot HS_{cb}^{age4} + \gamma \cdot C_{icb} + \theta_c + \tau_b + \varepsilon_{icb}.$$

In (1), Y_{icb} is the outcome of individual i , from childhood county c , in birth cohort b . The variable of interest (HS_{cb}^{age4}) is Head Start spending per poor four-year-old in county c (in year 2000 dollars), when birth cohort b was age four. To rely only on within-county variation in Head Start spending across cohorts, (1) includes childhood county fixed effects (θ_c); and to account for cohort effects, we include birth-year fixed effects (τ_b). We also include an extensive set of childhood-family and individual characteristics, and county-level coincident policy changes as control variables (C_{icb}) that we detail in Section IVC. The idiosyncratic error term is ε_{icb} .

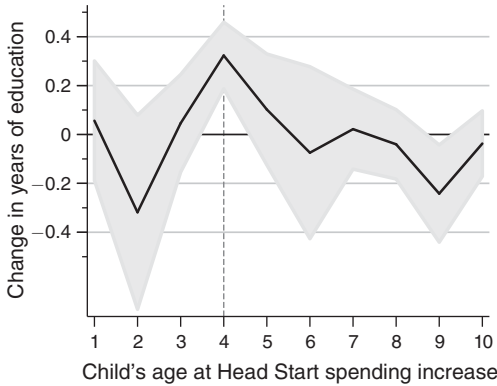
There are two identifying assumptions. First, counties that experienced larger or smaller increases in Head Start spending over time were not already on a trajectory of improving or deteriorating outcomes over time. Second, counties that saw larger or smaller increases in Head Start spending did not also undergo other unobserved changes that would also affect outcomes. Figure 2 suggests that the first condition is satisfied. To show that the second assumption is likely satisfied, we examine whether areas that had higher levels of Head Start spending may have also introduced policies and programs that may have improved child outcomes.

To test this, we estimated the marginal effect of Head Start spending levels that prevailed when individuals were different ages, conditional on the level of Head Start spending when they were four (shown in panel A of Figure 3 and online Appendix Table H8). Higher levels of Head Start spending at age four are associated with improved adult outcomes, while the spending levels at ineligible ages (age one through three or five through ten) are not.²² If areas with high levels of Head Start spending also implemented other policies that would promote better outcomes, then Head Start spending at age five, three, or seven would systematically be associated with better outcomes. This is clearly not the case. As another check on this variation, we regress each person's years of education and wage on our rich set of individual, family, and neighborhood characteristics and other social safety net programs. The fitted values from these regressions are effect-size weighted indices of childhood family and community socioeconomic factors (online Appendix Table H3). Conditional on school-district and birth-year fixed effects only, there is no association between Head Start spending and these predicted outcomes. Taken together, this is compelling evidence that our variation is valid. However, to assuage any lingering worries, we also implement a second strategy.

Because local areas with high versus low levels of Head Start spending may differ in ways that could confound our comparisons, our second identification strategy relies only on the variation in the availability of any local Head Start center at age four. To do this, we instrument for Head Start spending per poor four-year-old in county c (HS_{cb}^{age4}), with an indicator variable of whether a Head Start center

²² Online Appendix Figure H1 is an analogous figure for adult wages.

Panel A. Effects of \$1,000 increase in Head Start spending by child age on educational attainment



Panel B. Effects of county Head Start center access by child age on educational attainment

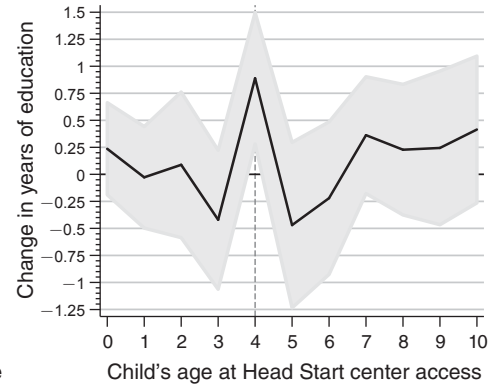


FIGURE 3. EFFECTS OF HEAD START SPENDING AND ACCESS: BY AGE OF SPENDING AND ACCESS (POOR CHILDREN)

Notes: These figures present the marginal effects of Head Start spending in an individual's childhood county at different ages, conditional on the level of Head Start spending in the childhood county at age four (when such spending should have an effect). The shaded gray region in the event study plots depict the 90 percent confidence interval for each rollout age estimate. The sample is poor children only. Models include the full set of controls as in Tables 2 and 3. The coefficients on the non-eligible years 1–3 and 5–10 are all conditional on spending at age four. The coefficient for spending at age four is based on a model with no other ages included. For the regression estimates underlying this model for years of education attained, see online Appendix Table H6.

existed in one's childhood county at four years old ($Exposed_HS_{cb}^{age4}$). Formally, we estimate the following system of equations by two-stage-least-squares (2SLS):

$$(2) \quad \widehat{HS}_{cb}^{age4} = \pi_{hs,1} \cdot Exposed_HS_{cb}^{age4} + \pi_{hs,1} \cdot C_{icb} + \theta_c + \tau_b;$$

$$(3) \quad Y_{icb} = \beta^{2SLS} \cdot \widehat{HS}_{cb}^{age4} + \gamma \cdot C_{icb} + \theta_c + \tau_b + \varepsilon_{icb}.$$

The identifying assumptions in this 2SLS model are weaker than those for the DiD model. This model is identified if counties that establish a Head Start center were not already on a trajectory of improving or deteriorating outcomes over time, and counties that established a Head Start center did not also undergo other unobserved changes that would also affect outcomes. Panel B of Figure 3 (and online Appendix Table H9) shows the effect of rollout (as opposed to spending) by age on years of education for poor children.²³ Reassuringly, there is an effect of having access to Head Start at age four but no effect of having access to Head Start for any other age (conditional on access at age four). Figures 2 and 3 suggest that the identifying conditions are satisfied. Furthermore, in Section VB, we present further evidence to support a causal interpretation of our estimates.

²³ Online Appendix Figure H1 is an analogous figure for adult wages.

B. Identifying the Effects of K–12 School Spending

Our measure of K–12 public school spending during childhood, ppe_{idb}^{5-17} , is the natural log of average public K–12 school spending per pupil (in real 2000 dollars) during school-age years (ages 5–17) in an individual’s childhood school district.²⁴ We refer to this as K–12 spending. Individuals who turned 17 years old during the year of the passage of a court-ordered SFR in their state should have completed secondary school by the time reforms were enacted. Such cohorts (and older cohorts) are “SFR *unexposed*.” Individuals who turned 16 years old or younger during the year of the passage of the first court-ordered SFR in their state would likely have attended primary or secondary school when reforms were implemented. Such cohorts are “SFR *exposed*.” One can estimate the SFR exposure effect on outcomes for individuals from a particular district by comparing the change in outcomes between SFR-exposed and SFR-unexposed birth cohorts from that district. Some districts experienced larger spending increases due to a court-ordered SFR than others. We exploit this fact and test for a causal effect of per pupil spending during childhood by testing whether the difference in outcomes between SFR-exposed and SFR-unexposed cohorts from the same school district (i.e., the SFR exposure effect) tends to be larger for those districts that experienced larger reform-induced K–12 spending increases (i.e., a SFR dose-response effect). Our identifying assumption is that the spending changes caused by the reforms within districts were unrelated to other district-level changes that could have affected adult outcomes directly.

Following JJP, we quantify the relationship between K–12 spending and adult outcomes by using only the variation above in school spending associated with the passage of a court-mandated SFR. Specifically, using the PSID, we estimate equation (4) by 2SLS. All common variables are defined as in (1):

$$(4) \quad Y_{idcb} = \beta \cdot \widehat{ppe}_{idb}^{5-17} + \gamma \cdot C_{idcb} + \theta_d + \tau_b + \varepsilon_{idcb}.$$

To rely only on variation across birth cohorts within districts, we include school district fixed effects (θ_d); to account for time trends and cohort effects, we include birth-year fixed effects (τ_b); and to account for life-cycle effects, we include flexible controls for age (cubic). Our endogenous regressor is ppe_{idb}^{5-17} , and $\widehat{ppe}_{idb}^{5-17}$ are fitted values from a first stage.

The excluded instruments in the first stage are measures of exposure to a SFR interacted with measures of dosage (to account for the fact that some districts have larger reform-induced spending increases than others). Our exposure measure, $SFRExp_{idb}$, is the number of years individual i in birth cohort c from childhood district d is expected to have been in school after the passage of the first court-ordered SFR in their home state. This exposure measure varies at the state birth-cohort level and goes from 0 (for those who were age 17 or older the year of the state’s first court-ordered SFR) to 12 (for those who were ages 5 and younger the year of the state’s court-ordered SFR). To capture variation in dosage

²⁴ We use the natural log to capture the fact that school spending likely exhibits diminishing marginal product.

conditional on exposure, in the first stage we also include the two-way interaction between $SFRExp_{idb}$ and a district-level predictor of the spending change caused by the state court-ordered SFR in that district ($\widehat{dose}_d$). More formally, the first-stage regression is as in (5) below:

$$(5) \widehat{ppe}_{idb}^{5-17} = \pi_1(SFRExp_{idb} \times \widehat{dose}_d) + \pi_2(SFRExp_{idb}) + \gamma_1 \cdot C_{idcb} + \theta_{d1} + \tau_{b,1}.$$

Following JJP, $\widehat{dose}_d$ is a predicted reform-induced spending change for each district based on reform type (implemented at the state level), pre-reform district income levels, pre-reform district spending levels and their interactions.²⁵ By construction, $\widehat{dose}_d$ is unrelated to endogenous decisions made by districts after reforms. Because we estimate $\widehat{dose}_d$ using *all* school districts while we estimate effects using the PSID sample, our approach is a two-sample-2SLS.²⁶ To assuage any concerns regarding $\widehat{dose}_d$, online Appendix Table H2 shows that the estimated point estimates obtained when using only variation in SFR exposure are almost identical (albeit less precise) than those that use both exposure and exposure times dosage.

Figure 4 shows the evolution of K–12 spending among individuals in the PSID sample from districts with high predicted dosage (i.e., $\widehat{dose}_d > 0$) and those with no predicted increases (i.e., $\widehat{dose}_d \leq 0$).²⁷ We create “event-time” indicator variables denoting the year an individual turned 17 minus the year of the first court order in the childhood state of individual i . Accordingly, negative values are cohorts who were 18 or older at the passage of a court-ordered SFR, the “0” cohort was 17 years old at the passage of a court-ordered SFR, and the “5” cohort was 12 years old at the passage of a court-ordered SFR in their state. We then estimate a regression model predicting school-age K–12 spending as a function of year fixed effects, district fixed effects, and the event-time indicators interacted with whether the district is predicted to have increased K–12 spending due to the passage of a court-ordered SFR. Because the outcome is in logs, the values represent percent changes in average school-age spending relative to the cohort from the same district that was 17 the year of the first court-ordered SFR. As shown in JJP, unexposed cohorts in districts with lower and higher predicted dosage were on similar pre-reform trajectories; however, exposed cohorts in high dose states experienced much larger increases in per pupil spending after a SFR.

²⁵ To form $\widehat{dose}_d$, we use the full universe of school districts and regress per pupil spending on indicators for years of SFR exposure, interacted with reform type, interacted with pre-reform spending levels in 1972; and indicators for years of SFR exposure, interacted with reform type, interacted with pre-reform median income levels in 1963, and region-specific year fixed effects. This regression models how per pupil spending evolves in a district after the passage of a court-ordered SFR as a function of the funding formula introduced in the state, the school spending level in the district, and the economic characteristics of the district *prior* to reforms. We take the fitted values from this regression to obtain a predicted reform-induced spending change for each district (based on these exogenous variables). See online Appendix F and Jackson, Johnson, and Persico (2016) for more details.

²⁶ This approach was popularized by Angrist and Krueger (1992) and has been used in several other settings (e.g., Björklund and Jäntti 1997, Currie and Yelowitz 2000, Dee and Evans 2003).

²⁷ Roughly two-thirds of districts in reform states are predicted to experience spending increases in the first eight years due to court-ordered SFRs. As one can see from Figure 4, because K–12 spending tended to increase in states following court-ordered SFRs in general, there are small increases in K–12 spending within 12 years post-reform even in districts with predicted initial decreases. As such, we refer to all districts as having high- or low-predicted increases.

Panel A. Per pupil K–12 spending: By predicted dose Panel B. Years of education: By predicted dose

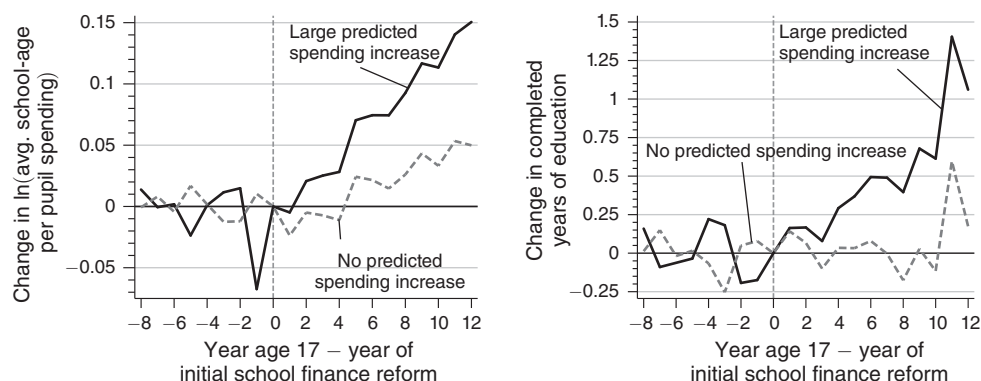


FIGURE 4. EVOLUTION OF K–12 SPENDING AND EDUCATIONAL ATTAINMENT AFTER SFR REFORM (ALL CHILDREN)

Notes: The event study figures use school district's predicted reform-induced change in spending based on the timing and type of court-ordered reform interacted with 1970 (within-state) district income and spending percentile categories—the solid black line shows estimated effects for districts with a predicted reform-induced K–12 spending increase ($dose_d > 0$), whereas the solid gray line shows the corresponding effects for districts with low predicted reform-induced K–12 spending increases or a decrease $dose_d \leq 0$. Roughly two-thirds of districts in reform states had predicted spending increases. The event study models include: school district fixed effects, race-specific year of birth fixed effects, race $\times$ census division-specific birth-year trends; controls at the county-level for the timing of school desegregation $\times$ race, hospital desegregation $\times$ race, rollout of "War on Poverty" and related safety net programs (community health centers, food stamps, Medicaid, AFDC, UI, Title-I (average during childhood years)), timing of state-funded kindergarten introduction and timing of tax limit policies; controls for 1960 county characteristics (poverty rate, percent black, education, percent urban, population size) each interacted with linear cohort trends; and controls for childhood family characteristics (parental income/education, mother's marital status at birth, birth weight, gender).

This shows that the timing of the initial court-ordered SFR in the state interacted with the predicted reform-induced spending increase for the district (based on state reform type interacted with pre-reform district characteristics) isolates exogenous variation in school spending.

If our identification strategy is valid and K–12 spending affects outcomes, outcome differences across exposed and unexposed cohorts should follow similar patterns to those of K–12 spending. Panel B of Figure 4 shows this for years of educational attainment. Areas that had small (gray line) and large (black line) reform-induced increases in K–12 spending were on similar trajectories among the unexposed cohorts (years –8 through year 0). However, the post-SFR cohorts (years 0 through 12) experienced much larger increases in years of education in the high-predicted K–12 spending increase districts than in the low-predicted K–12 spending increase districts. This figure depicts graphically the variation that undergirds our identification strategy.

The key identifying assumptions are that districts that experienced spending increases due to a SFR were not on different trajectories before reforms, and there were no coincident district-level policies or changes that confound our analysis. Figure 4 shows that this first condition is likely satisfied. We also test the second condition. If other coincident policies were driving the results (that were not targeted to school-age children), increased school spending might improve outcomes of

those who were in the same district but not of school-going age. To test this, we instrument for the K–12 spending levels that prevailed in an individual’s childhood district when they were between the ages of 18 to 22 (i.e., nonschool-going age). We find no effect on adult outcomes (online Appendix Table H1). Also, we find that conditional on school-district and birth-year fixed effects, there is no association between instrumented K–12 spending and predicted outcomes (online Appendix Table H3)—further evidence that our identifying variation is valid. While these tests are not dispositive, they support a causal interpretation of the main findings. To assuage any lingering concerns, we present additional tests in Section V.

C. Testing for Dynamic Complementarity

To test whether the marginal effect of increased Head Start spending varies by the level of K–12 spending and vice versa, we estimate the effects of public pre-K and K–12 spending on adult outcomes with the inclusion of the interaction between Head Start spending at age 4 (HS_{icb}^{age4}) and the natural log of public K–12 spending between the ages of 5 and 17 (ppe_{idb}^{5-17}). All models are estimated separately for poor and non-poor children, as we do not expect to find significant effects of Head Start spending nor evidence of dynamic complementarity among non-poor children (at least through direct channels as they are not income-eligible for Head Start). We define $INT_{idb} = (HS_{icb}^{age4} \times ppe_{idb}^{5-17})$. We estimate two different models in our analysis.

The DiD-by-2SLS Model.—In the first model, we use the within-county, across-cohort DiD variation in Head Start spending (HS_{icb}^{age4}). Because a school district may be a smaller unit of observation than a county, all models include district fixed effects (which subsumes county effects). We instrument for ppe_{idb}^{5-17} , with $(SFRExp_{idb})$ and $(SFRExp_{idb} \times \widehat{dose}_d)$. We instrument for INT_{idb} with $(HS_{icb}^{age4} \times SFRExp_{idb} \times \widehat{dose}_d)$ and $(HS_{icb}^{age4} \times SFRExp_{idb})$.²⁸ The resulting model is (6), where $\widehat{ppe}_{idb}^{5-17}$ and $\widehat{INT}_{idb}$ are fitted values from first-stage regressions:²⁹

$$(6) \quad Y_{icb} = \beta_{HS} \cdot HS_{icb}^{age4} + \beta_{K-12} \cdot \widehat{ppe}_{idb}^{5-17} + \beta_{int} \cdot (\widehat{INT}_{idb}) \\ + \gamma \cdot C_{icb} + \theta_d + \tau_b + \varepsilon_{idb}.$$

The 2SLS-by-2SLS model.—In the second model, we instrument for *all* spending variables. Now we instrument for Head Start spending (HS_{icb}^{age4}) using exposure to any Head Start center at age 4 ($Exposed_HS_{cb}^{age4}$). We instrument for ppe_{idb}^{5-17} ,

²⁸ While intuition would lead one to expect us to use all the two-way interactions between HS_{icb}^{age4} , $\widehat{dose}_d$, and $SFRExp_{idb}$, we do not use $(HS_{icb}^{age4} \times \widehat{dose}_d)$ as an excluded instrument because $\widehat{dose}_d$ cannot affect outcomes unless it is interacted with SFR exposure. This would simply introduce noise and weaken the first stage.

²⁹ Where $\hat{X}_1 = \widehat{ppe}_{idb}^{5-17}$ and $\hat{X}_2 = \widehat{INT}_{idb}$, and $w \in \{1, 2\}$,

$$\hat{X}_w = \pi_{w1}(SFRExp_{idb} \times dose_c) + \pi_{w2}(SFRExp_{idb}) + \pi_{w3}(SFRExp_{idb} \times \widehat{dose}_d) \cdot HS_{icb}^{age4} \\ + \pi_{w4}(SFRExp_{idb}) \cdot HS_{icb}^{age4} + \gamma_w C_{icb} + \theta_{wd} + \tau_{wb}.$$

with $(SFRExp_{idb})$ and $(SFRExp_{idb} \times \widehat{dose_d})$. Now we instrument for INT_{idb} with $(Exposed_HS_{cb}^{age4} \times SFRExp_{idb} \times \widehat{dose_d})$ and $(Exposed_HS_{cb}^{age4} \times SFRExp_{idb})$. The resulting model is as in (7), where $\widehat{ppe}_{idb}^{5-17}$, $\widehat{INT}_{idb}$, and $\widehat{HS}_{cb}^{age4}$ are all fitted values from first-stage regressions:³⁰

$$(7) \quad Y_{icb} = \beta_{HS} \cdot \widehat{HS}_{cb}^{age4} + \beta_{K-12} \cdot \widehat{ppe}_{idb}^{5-17} + \beta_{int} \cdot (\widehat{INT}_{idb}) \\ + \gamma \cdot C_{icb} + \theta_d + \tau_b + \varepsilon_{idb}.$$

The interaction effect between pre-K and K–12 spending can be identified because among counties that faced similar increases in Head Start spending (or had any Head Start center), some were located in school districts that experienced larger (or smaller) increases in K–12 spending due to the passage of a court-ordered reform; and among cohorts from districts that faced similar increases in K–12 spending due to the passage of a court-ordered reform, some grew up in counties that had higher (or no) levels of Head Start spending when those cohorts were age four.

To further reduce the possibility of confounding effects, vector C_{idb} includes a variety of individual, childhood family, and childhood county controls. These include parental education and occupational status, parental income, mother's marital status at birth, birth weight, child health insurance coverage, gender; and the adult economic and incarceration outcomes include flexible controls for age (cubic). The vector C_{idb} also includes birth-year fixed effects by region and race, birth-cohort linear trends interacted with various 1960 characteristics of the childhood county (poverty rate, percent black, average education, percent urban, and population size). Also, to avoid confounding our effects with that of other policies that overlap our study period, C_{idb} includes controls for childhood county-by-birth year measures of school desegregation, hospital desegregation, community health centers, state funding for kindergarten, Title I school funding, imposition of tax limit policies, average childhood spending on food stamps, Aid to Families with Dependent Children, Medicaid, and unemployment insurance (Johnson 2019; Chay, Guryan, and Mazumder 2009; and Hoynes, Schanzenbach, and Almond 2016). Standard errors are clustered at the state level.

To provide visual evidence of complementarity, Figure 5 plots the estimated changes in years of educational attainment for cohorts before and after a court-ordered SFR for districts with high predicted spending increases (i.e., $\widehat{dose_d} > 0$) and those with no predicted increases (i.e., $\widehat{dose_d} \leq 0$), separately for children with and without a local Head Start center at age four. This is the variation used in the 2SLS-by-2SLS models. Panel A shows that SFR-treated and untreated cohorts experienced similarly small changes in educational attainment in districts that had small increases in K–12 spending and were

³⁰ Where $\hat{X}_1 = \widehat{ppe}_{idb}^{5-17}$, $\hat{X}_2 = \widehat{INT}_{idb}$, $\hat{X}_3 = \widehat{HS}_{cb}^{age4}$, and $g \in \{1, 2, 3\}$,

$$\hat{X}_g = \pi_{g1}(SFRExp_{idb} \times dose_c) + \pi_{g2}(SFRExp_{idb}) + \pi_{g3}(SFRExp_{idb} \times dose_c \times Exposed_HS_{cb}^{age4}) \\ + \pi_{g4}(SFRExp_{idb} \times Exposed_HS_{cb}^{age4}) + \gamma_g C_{idb} + \theta_{gd} + \tau_{gb}.$$

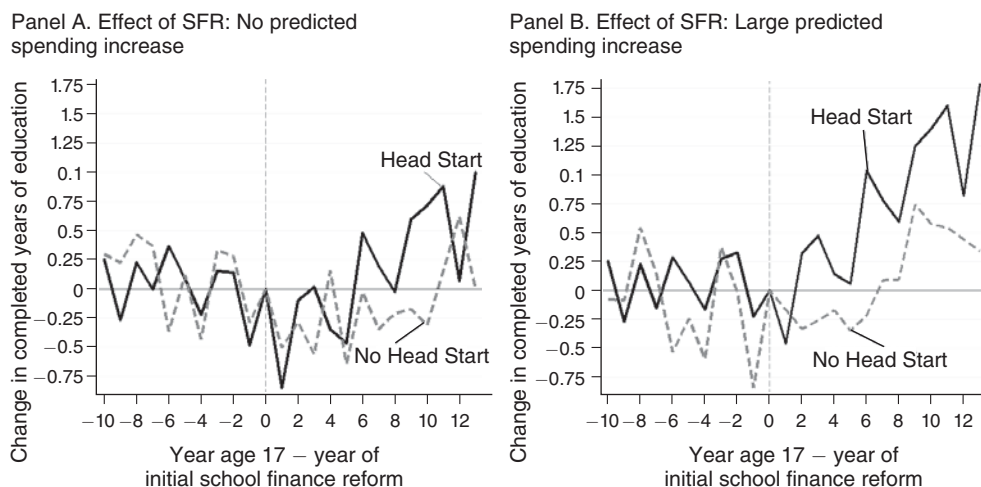


FIGURE 5. EFFECT OF K-12 SPENDING ON YEAR OF COMPLETED EDUCATION: BY HEAD START EXPOSURE STATUS (POOR CHILDREN)

Notes: The event study figures use school district's predicted reform-induced change in spending based on the timing and type of court-ordered reform interacted with 1970 (within-state) district income and spending percentile categories; panel B shows estimated effects for districts with a predicted reform-induced K-12 spending increase ($\widehat{dose}_d > 0$), whereas panel A shows the corresponding effects for districts with low predicted reform-induced K-12 spending increases or a decrease $\widehat{dose}_d \leq 0$. Roughly two-thirds of districts in reform states had predicted spending increases. These estimated effects are presented both for children whose county had no Head Start center at age four (gray line), and those who were exposed to any county Head Start spending at age four (black line), to highlight the role of dynamic complementarity. The event study models include: school district fixed effects, race-specific year of birth fixed effects, race $\times$ census division-specific birth-year trends; controls at the county-level for the timing of school desegregation $\times$ race, hospital desegregation $\times$ race, rollout of "War on Poverty" and related safety-net programs (community health centers, food stamps, Medicaid, AFDC, UI, Title-I (average during childhood years)), timing of state-funded kindergarten introduction and timing of tax limit policies; controls for 1960 county characteristics (poverty rate, percent black, education, percent urban, population size) each interacted with linear cohort trends; and controls for childhood family characteristics (parental income/education, mother's marital status at birth, birth weight, gender).

not exposed to Head Start at age four (gray line). However, among cohorts that had Head Start at age four, school-age years of exposure to SFRs led to increases in educational attainment relative to those who were not exposed to SFRs. This pattern is consistent with Head Start making even small increases in K-12 spending effective for poor children. However, if the two policies are complementary, one should see similar patterns and greater improvements for large increases in K-12 spending. This is precisely what we document in panel B of Figure 5. In districts that experienced large increases in K-12 spending after a SFR, exposed cohorts achieve more years of education than unexposed cohorts, and the relative increase is larger among those SFR-exposed cohorts that were from counties with a Head Start center at age four. Furthermore, the benefits of Head Start spending (the difference between the gray and black line in each panel) are larger among SFR-exposed cohorts that experience larger K-12 spending increases. In sum, Figure 5 presents visual evidence that Head Start and K-12 school spending exhibit dynamic complementarity.³¹ The

³¹ An analogous figure for adult wages is in online Appendix Figure L1. Though muted, one can see the same basic patterns.

lack of any differential pre-trending in either panel illustrates that the parallel trends assumption likely holds, not just for each policy (Figures 2 through 4), but also for the *interaction* between the two policies.³²

Testing for Sufficient Variation to Identify the Interaction Effects.—Identification of our key parameter of interest is based on the interaction between the two policy instruments. For our inference to be valid, these policy instruments need to be largely independent of each other. This is necessary for two reasons. First, if there were a high correlation between the two policy instruments, a model predicting both the base effects and the interaction could be under-identified. If so, there would be a weak first stage for the interaction, conditional on the instruments for the base effects. Second, if areas that were most likely to have high levels of Head Start spending were also likely to have larger SFR induced K–12 spending increases, then areas that were exposed to high levels of both may differ from areas that were only exposed to only one, or none in unobserved ways. Because our interaction is essentially a comparison of Local Average Treatment Effects (LATEs), *if* there is treatment heterogeneity, the resulting interaction effect *could* simply reflect a difference in LATEs rather than a true interaction effect.

We show that this is not a problem in our setting. First, the correlation between Head Start spending and instrumented K–12 spending is only 0.15, and conditional on controls, there is no association between the two (online Appendix Table H6). This suggests that our treatments are largely independent so that we are not comparing different LATEs. Also, following Angrist and Pischke (2009), we compute first-stage *F*-statistics for each set of excluded instruments, conditional on the other excluded instruments. The first-stage *F*-statistic on the instruments for K–12 spending (i.e., predicted SFR dosage times years of SFR exposure) is 22.41 and 23.01 in models without and with Head Start variables included, respectively (online Appendix Table H7). The first-stage *F*-statistic on the excluded instruments for Head Start spending (i.e., the existence of a Head Start center at age 4) is 59.17 and 60.76 in models without and with the K–12 instruments included, respectively. Finally, the first-stage *F*-statistic on Head Start Exposure times SFR dosage times SFR exposure is 42.46, conditional on Head Start Exposure, SFR exposure, and SFR dosage times SFR exposure. In sum, there is sufficient variation in Head Start spending and SFR-induced changes in K–12 spending for the effect of each to be identified and for the interaction between the two to be identified.

³² If the alignment channel is at play, complementarity would be larger for cohorts that were exposed to a SFR soon after Head Start than for those exposed later. The patterns in Figure 5 are consistent with this. Specifically, the cohorts that are exposed to an SFR for more years benefit much more from Head Start than those exposed to a SFR for fewer years (e.g., 8–10 years versus 2–6 years). To present another suggestive test, we estimate our main models and interact all of our K–12 spending variables with indicators measuring the age at which the SFR was implemented in one's childhood state. The results (online Appendix K) suggest that the complementarity effects are driven by those cohorts that were exposed to a SFR before the age of 9.

V. Results

We present results from specification (6) that exploits all the within-district, across-cohort variation in Head Start spending and instruments for K–12 public school spending using the SFR instruments, and specification (7) that instruments for both Head Start spending and K–12 spending. To facilitate interpretation of the base effects of K–12 spending and Head Start spending when the interaction between the two is included, both K–12 spending and Head Start spending are centered on their respective means. Thus, the coefficient on Head Start is the marginal effect of Head Start spending at the average level of K–12 spending, and the coefficient on K–12 spending is the marginal effect of K–12 spending at the average level of Head Start spending. To organize our discussion, we first discuss the base effects of K–12 spending (in logs) and Head Start spending, present empirical evidence that these estimated base effects are unbiased, and then discuss the estimated interaction effects. We present our estimated effects on education outcomes, followed by adult economic outcomes, and finally incarceration.

A. Estimating the Base Effects of Head Start and K–12 Spending

Table 2 presents the estimates for poor children. Column 1 presents the DiD-2SLS estimates of the effects on the probability of graduating from high school. The coefficient on Head Start spending per poor four-year-old is 0.025 (p -value < 0.01). That is, increasing Head Start spending per poor four-year-old in the county by \$1,000 (roughly a 25 percent increase) increases the likelihood of graduating from high school by 2.5 percentage points for a poor child exposed to the average level of K–12 spending. Given that the average level of Head Start spending, conditional on having any Head Start program in the county, is \$4,230, this implies that, for poor children, having access to the average Head Start program increased the likelihood of graduating from high school by roughly 10 percentage points. Column 2 presents effects for the 2SLS-2SLS design that instruments for all spending variables. The 2SLS coefficient on Head Start spending per poor four-year-old is quite similar (it is 0.0408 and is statistically significant), and one cannot reject that the DiD-2SLS models and the 2SLS-2SLS models yield different results. However, in the fully instrumented model, the effect of Head Start spending is slightly *larger* and less precisely estimated. Because the DiD-2SLS estimated Head Start effects tend to be smaller, and the results are similar to the fully instrumented model, we take a conservative approach and focus discussion on the DiD-2SLS results. However, Table 2 reports all results from the 2SLS-2SLS models.

Increases in Head Start spending can affect outcomes through increases in Head Start participation, increases in the quality and scope of Head Start services, and can also indirectly affect outcomes through peer effects in the K–12 system due to having better prepared schoolmates. While existing studies have focused on the effect of *enrolling in* Head Start as participants, we estimate the effect of Head Start spending on *all* eligible children. Because there are multiple channels through which spending effects may emerge, we provide

TABLE 2—MARGINAL EFFECTS OF HEAD START SPENDING AND PUBLIC PER PUPIL SPENDING AND THEIR INTERACTION: POOR CHILDREN

	Pr(high school grad)		Years of completed education		ln(wage), ages 20–50	
	DiD-2SLS (1)	2SLS-IV (2)	DiD-2SLS (3)	2SLS-IV (4)	DiD-2SLS (5)	2SLS-IV (6)
Head Start spending _(age 4)	0.02503 (0.006942)	0.04089 (0.02453)	0.07721 (0.01992)	0.2255 (0.1212)	0.02334 (0.004503)	0.03615 (0.01956)
(SFR) instrumented ln(PPE) _(age 5–17)	1.1016 (0.3268)	1.4163 (0.3390)	4.0399 (1.6751)	4.0218 (1.7856)	2.0561 (0.4348)	1.2596 (0.2690)
Head Start spending _(age 4) × ln(PPE) _(age 5–17)	0.1012 (0.05454)	0.2273 (0.06518)	0.6460 (0.2354)	0.8345 (0.4824)	0.1698 (0.06985)	0.2561 (0.07191)
<i>Marginal effects of 10% increase in K–12 spending by Head Start access</i>						
No Head Start _(age 4)	0.0673 (0.0236)	0.0455 (0.0316)	0.1307 (0.1274)	0.0492 (0.1064)	0.1338 (0.0349)	0.0176 (0.0219)
Head Start center access _(age 4)	0.1102 (0.0327)	0.1416 (0.0339)	0.4040 (0.1675)	0.4022 (0.1786)	0.2056 (0.0435)	0.1260 (0.0269)
<i>Marginal effects of Head Start with 10% increase or decrease in K–12 spending</i>						
With 10% decrease	0.0630 (0.0481)	0.0768 (0.1169)	0.0533 (0.1393)	0.6010 (0.5937)	0.0269 (0.0284)	0.0446 (0.0921)
Average	0.1059 (0.0294)	0.1730 (0.1038)	0.3266 (0.0843)	0.9540 (0.5129)	0.0987 (0.0190)	0.1529 (0.08275)
With 10% increase	0.1487 (0.0217)	0.2691 (0.0968)	0.5999 (0.1209)	1.3070 (0.5068)	0.1706 (0.0408)	0.2613 (0.0841)
Number of person-year observations	—	—	—	—	55,706	55,706
Number of children	5,419	5,419	5,419	5,419	5,613	5,613

(continued)

a sense of how our spending effects relate to the participation effects in the extant literature.

We estimate that the rollout of Head Start increased Head Start participation for poor children by about 75 percentage points. We come to this conclusion in two ways.³³ First, using national data, for cohorts entering kindergarten after 1966, the likelihood of Head Start enrollment (full-time or part-year) among income-eligible children was 63 percent (Figure 1). Because centers can enroll 10 percent of non-poor children, the participation rate among income-eligible children *could have been* as low as 57 percent. Roughly 80 percent of poor children born after 1962 in the PSID resided in a county with a Head Start center at age four during this period (this is consistent with national figures). Assuming that only children with a Head Start center in their local area at age four will participate, this implies a Head Start participation rate of $0.57/0.8 = 0.71$ (i.e., 71 percentage points), conditional on having a Head Start center in the county at age four. We arrive at a similar estimate using retrospective survey questions from the PSID. In the 1995

³³ The PSID survey data employed in Garces, Thomas, and Currie (2002) are retrospective data collected in the 1995 wave. There are some concerns about potential measurement error and recall bias in using this retrospective survey information about Head Start participation and some missing information. See online Appendix G for further discussion.

TABLE 2—MARGINAL EFFECTS OF HEAD START SPENDING AND PUBLIC PER PUPIL SPENDING
AND THEIR INTERACTION: POOR CHILDREN (CONTINUED)

	Annual incidence of poverty, age 20–50		Pr(ever incarcerated)	
	DiD-2SLS (7)	2SLS-IV (8)	DiD-2SLS (9)	2SLS-IV (10)
Head Start spending _(age4)	−0.01808 (0.005302)	−0.02576 (0.01385)	−0.006002 (0.003494)	−0.02024 (0.01082)
(SFR) instrumented ln(PPE) _(age5–17)	−0.7923 (0.2969)	−0.7971 (0.2903)	−0.8080 (0.3397)	−1.1822 (0.4550)
Head Start spending _(age4) × ln(PPE) _(age5–17)	−0.1079 (0.04267)	−0.1852 (0.05038)	−0.05169 (0.02777)	−0.1808 (0.1076)
<i>Marginal effects of 10% increase in K–12 spending by Head Start access</i>				
No Head Start _(age4)	−0.0336 (0.0301)	−0.0014 (0.0153)	−0.0589 (0.0283)	−0.0418 (0.0193)
Head Start center access _(age4)	−0.0792 (0.0297)	−0.0797 (0.0290)	−0.0808 (0.0340)	−0.1182 (0.0455)
<i>Marginal effects of Head Start with 10% increase or decrease in K–12 spending</i>				
With 10% decrease	−0.0308 (0.0206)	−0.0306 (0.0568)	−0.0035 (0.0229)	−0.0092 (0.0487)
Average	−0.0765 (0.0224)	−0.1090 (0.0586)	−0.0254 (0.0148)	−0.0856 (0.0457)
With 10% increase	−0.1221 (0.0351)	−0.1873 (0.0674)	−0.0473 (0.0138)	−0.1621 (0.0772)
Number of person-year observations	88,124	88,124	—	—
Number of children	6,373	6,373	4,536	4,536

Notes: Robust standard errors are in parentheses (clustered at childhood state level). PSID geocode data (1968–2015), matched with childhood school and neighborhood characteristics. Sample includes all individuals born 1950–1976 whose parents were in the bottom quartile of the income distribution and who have been followed into adulthood. Models: (non-instrumented and instrumented) Head Start spending per poor four-year-old at age four in the county and instrumented ln(school district per pupil spending during ages 5–17) are centered around their respective means, to facilitate interpretation of the main effects evaluated at roughly the mean; the average SFR-induced increase in school-age spending is about 10 percent. Results are based on 2SLS-IV models that include: school district fixed effects, race-specific year of birth fixed effects, race × census division-specific birth-year trends; controls at the county-level for the timing of school desegregation × race, hospital desegregation × race; controls for 1960 county characteristics (poverty rate, percent black, education, percent urban, population size, each interacted with linear cohort trends; controls for county-level per capita government safety net expenditures average during childhood, and controls for childhood family characteristics (parental income/education, mother’s marital status at birth, birth weight, gender), and age (cubic). The first-stage model includes as predictors the school-age years of exposure to school finance reform interacted with the quartile of the respective school district’s predicted reform-induced change in school spending based on the timing and type of court-ordered reform interacted with 1970 (within-state) district income and spending percentile categories. The instrument used for Head Start spending per poor four-year-old is an indicator for whether there was any Head Start center in the county at age four (based on the program’s rollout timing variation only). There exists a significant first stage. The marginal effects related to Head Start access are based on the average county Head Start spending when there is a center (~\$4,230 (in real 2000 dollars))—i.e., marginal effects are evaluated for roughly a \$4K increase in Head Start spending (to contrast impact of having access versus no access to Head Start center).

survey wave as part of a special module on early childhood, adults were asked about whether they had ever participated in a Head Start program. These data may have a number of limitations such as recall bias (see online Appendix G). However, in these data, Head Start rollout increases Head Start participation among poor children by about 80 percentage points. We use 75 percentage points as our “ballpark” estimate of the increase in the likelihood of Head Start participation

(among poor children) due to the rollout of the average Head Start center in the county during our study period.

If all of our estimated effects of having Head Start access was due to Head Start enrollment (and there were no spillover effects to other poor children), our participation margin effect implies a treatment-on-the-treated effect of $0.1/0.75 = 0.129$, or 13.3 percentage points. This is similar to the estimated enrollment effect of Head Start in existing studies.³⁴ However, most existing studies of Head Start focus on full-year Head Start programs. If one makes the conservative assumption that there is no effect of summer-only programs or part-time programs, a back-of-the-envelope calculation yields an implied treatment-on-the-treated effect of full-year Head Start on the likelihood of high school graduation of 15.3 percentage points.³⁵ This estimate is in line with the larger of the participation margin effects in the literature. However, we cannot rule out that some modest portion of our effects are driven by improvements in the quality and scope of Head Start centers (full day versus half day, full time versus summer only, better teachers, etc.), and spillovers from Head Start participants to poor nonparticipants in the K–12 school system.

Our K–12 spending results replicate JJP. The coefficient on the log of K–12 spending during the school-age years is 1.10 (p -value < 0.01). Increasing K–12 school spending (across all 12 school-age years) by 10 percent increases the likelihood of high school graduation by about 11 percentage points for a poor child exposed to the average level of Head Start spending (Table 2, column 1). Relative to baseline, this is about a 15 percent increase. The estimates indicate that increasing Head Start spending by \$4,000 would have roughly the same effect on high school graduation as increasing K–12 spending by 10 percent across all school-age years (for poor children).³⁶

Table 2, columns 3 and 4 present a similar pattern for completed years of education for poor children. The more conservative DiD-2SLS estimates reveal that increasing Head Start spending per poor four-year old in the county by \$1,000 increases the years of educational attainment by 0.077 years (p -value < 0.01) for a poor child exposed to the average level of K–12 spending. At average Head Start spending levels, a Head Start center is estimated to increase years of education by roughly a third of a year. Increasing school-age K–12 spending by 10 percent increases the number of years of completed education by about 0.4 years for a poor

³⁴ For example, Garces, Thomas, and Currie (2002) finds that participating in Head Start increases the high school graduation rates for whites by 20 percentage points, with no statistically significant effect for blacks. Deming (2009) finds that Head Start participation increases high school graduation by 11 percentage points for blacks with a small effect for whites, and increases high school graduation by 16 percentage points for those with low maternal test scores. Weikart, Marcus, and Xie (2000) finds that the average effect is 14 percentage points.

³⁵ The average enrollment rate among eligible children was 52 percent after the initial ramp-up period (for cohorts entering kindergarten after 1966). This implies a full-year Head Start participation rate of about $0.52/0.8 = 0.65$ conditional on having a Head Start center in the county at age four. If one makes the conservative assumption that there is no effect of summer only programs or part-time programs so that *all* of our estimated intention-to-treat effect was due to *full-year* Head Start enrollment, an assumed upper-bound full-year Head Start participation margin effect implies a treatment-on-the-treated effect on the likelihood of high school graduation of $0.1/0.65 = 0.153$.

³⁶ During the sample period, a 10 percent increase in K–12 spending is roughly equal to increasing per pupil K–12 spending by \$480 each year over 12 years (about \$4,300 in present value terms assuming a 7 percent interest rate).

TABLE 3—MARGINAL EFFECTS OF HEAD START SPENDING AND PUBLIC PER PUPIL SPENDING
AND THEIR INTERACTION: NON-POOR CHILDREN

	Pr(high school grad)		Years of completed education		ln(wage), ages 20–50	
	DiD-2SLS (1)	2SLS-IV (2)	DiD-2SLS (3)	2SLS-IV (4)	DiD-2SLS (5)	2SLS-IV (6)
Head Start spending _(age4)	0.000014 (0.003432)	−0.02227 (0.01864)	0.008866 (0.01635)	−0.05426 (0.1071)	0.006901 (0.005408)	0.01932 (0.02433)
(SFR) instrumented ln(PPE) _(age5–17)	0.2386 (0.1197)	0.4671 (0.2351)	2.4192 (1.1645)	2.1565 (1.5314)	0.7351 (0.3035)	0.4155 (0.2366)
Head Start spending _(age4) × ln(PPE) _(age5–17)	0.01688 (0.02347)	0.09666 (0.08062)	0.02972 (0.1937)	0.4144 (0.3706)	0.02577 (0.03090)	−0.01603 (0.1020)
Number of person-year observations	—	—	—	—	90,771	90,771
Number of children	7,983	7,983	7,983	7,983	8,195	8,195
	Annual incidence of poverty, age 20–50		Pr(ever incarcerated)			
	DiD-2SLS (7)	2SLS-IV (8)	DiD-2SLS (9)	2SLS-IV (10)		
Head Start spending _(age4)	−0.000085 (0.001716)	−0.008452 (0.005692)	−0.001274 (0.001705)	0.000937 (0.006974)		
(SFR) instrumented ln(PPE) _(age5–17)	−0.1383 (0.06316)	−0.1868 (0.1304)	−0.09837 (0.2161)	0.1298 (0.2758)		
Head Start spending _(age4) × ln(PPE) _(age5–17)	0.005707 (0.01459)	−0.000716 (0.04094)	−0.02568 (0.02271)	−0.01128 (0.06604)		
Number of person-year observations	130,470	130,470	—	—		
Number of children	8,859	8,859	5,140	5,140		

Notes: Robust standard errors are in parentheses (clustered at childhood state level). PSID geocode data (1968–2015), matched with childhood school and neighborhood characteristics. Sample includes all individuals born 1950–1976 whose parents were NOT in the bottom quartile of the income distribution, and who have been followed into adulthood. Models: (non-instrumented and instrumented) Head Start spending per poor four-year-old at age four in the county and instrumented ln(school district per pupil spending during ages 5–17) are centered around their respective means, to facilitate interpretation of the main effects evaluated at roughly the mean; the average SFR-induced increase in school-age spending is about 10 percent. Results are based on 2SLS-IV models that include: school district fixed effects, race-specific year of birth fixed effects, race × census division-specific birth-year trends; controls at the county-level for the timing of school desegregation × race, hospital desegregation × race; controls for 1960 county characteristics (poverty rate, percent black, education, percent urban, population size, each interacted with linear cohort trends; controls for county-level per capita government safety net expenditures average during childhood, and controls for childhood family characteristics (parental income/education, mother’s marital status at birth, birth weight, gender), and age (cubic). The first-stage model includes as predictors the school-age years of exposure to school finance reform interacted with the quartile of the respective school district’s predicted reform-induced change in school spending based on the timing and type of court-ordered reform interacted with 1970 (within-state) district income and spending percentile categories. The instrument used for Head Start spending per poor four-year-old is an indicator for whether there was any Head Start center in the county at age four (based on the program’s rollout timing variation only). There exists a significant first stage. The marginal effects related to Head Start access are based on the average county Head Start spending when there is a center (~\$4,230 (in real 2000 dollars))—i.e., marginal effects are evaluated for roughly a \$4K increase in Head Start spending (to contrast impact of having access versus no access to Head Start center).

child exposed to the average level of Head Start spending. As discussed previously, the 2SLS-2SLS results in column 4 are similar.

Results for non-poor children are in Table 3. As in JJP, the estimated K–12 spending effects on the education outcomes are positive, sizable, and statistically significantly different from zero. This indicates that increases in K–12 spending improve the educational outcomes of not only the poor but also the non-poor. The DiD-2SLS point estimates indicate that increasing district K–12 spending by 10 percent increases the likelihood of graduating high school by 2.3 percentage

points, and increases years of educational attainment by about 0.24 years for a non-poor child exposed to the average level of Head Start spending. These estimated K–12 spending benefits are smaller for more affluent children than for poor children, but they are positive, statistically significant, and economically important. In contrast to the positive K–12 spending effects, for non-poor children, both the DiD-2SLS and 2SLS-2SLS model reveal that increasing Head Start spending has small, insignificant effects. For both education outcomes, one cannot reject that the effect on the non-poor is zero, and one *can* reject that the Head Start effect is the same for both poor and non-poor children.³⁷ This suggests that there are no spillover effects of Head Start spending on non-poor children and that increases in Head Start spending are not associated with other broad policies that improve the outcomes of non-poor children.

The fact that we find no effect of Head Start spending for non-poor children is important. If local areas that increased Head Start spending introduced other policies that improve outcomes of all children, one would observe positive Head Start spending effects for the non-poor children. We find no such pattern. Our result, instead, implies that neither our variation in Head Start spending nor the rollout of Head Start is associated with any policies that improved the outcomes of local children who were ineligible to participate in Head Start. This coupled with the fact that Head Start spending only influences outcomes for those who were four years old at the time shows that we only see effects for children who were both income- and age-eligible for Head Start. This serves as another falsification test, of sorts, and bolsters the credibility of the research design.

The adult economic outcomes we examine are wages and the annual incidence of poverty between the ages of 20 and 50. Our models use all available person-year observations for ages 20–50 and control for a cubic in age to avoid confounding life-cycle and birth-cohort effects. Columns 5 through 8 in Table 2 present these results for children from poor families. Looking at wages, in the DiD-2SLS models (column 5) the coefficient on the log of public K–12 school spending is 2.056 (p -value < 0.1) and that on Head Start spending per poor four-year-old is 0.023 (p -value < 0.01). That is, for children from poor families exposed to average levels of Head Start spending, increasing K–12 spending by 10 percent is associated with about 20.5 percent higher adult wages. Similarly, for these same children, at average public K–12 spending levels, increasing Head Start spending by \$4,230 per poor four-year-old (the average spending amount) is associated with 9.87 percent higher wages for poor children. The results in the 2SLS-2SLS models are similar and cannot be distinguished statistically.

Columns 5 and 6 of Table 3 present the effects on adult wages for non-poor children. Similar to the educational outcomes, there are positive effects of K–12 spending, but no effect of Head Start spending on the wages in adulthood of those from non-poor families. In the DiD-2SLS models, the coefficient on the log of K–12

³⁷ We pooled the samples and estimated a single model where we interacted all variables with poverty status and tested for equality of coefficients between poor and non-poor children for our key explanatory variables. We present the results of this test for our two main adult outcomes in the conservative DiD-2SLS models in online Appendix Table J1. Our tests reject that the estimates are the same for the two populations.

public school spending is 0.7351 (p -value < 0.05), and that on Head Start spending per poor four-year-old is 0.0069 (p -value > 0.1). That is, for children from non-poor families exposed to average levels of Head Start spending, increasing K–12 spending by 10 percent is associated with 7.35 percent higher earnings between the ages of 20 and 50, while increasing Head Start spending is associated with no difference in earnings. The 2SLS-2SLS results (column 6) tell the same basic story as the DiD-2SLS models.

The pattern of estimates for the annual incidence of poverty in adulthood in columns 7 and 8 of Tables 2 and 3 mirror those for adult wages. A family is poor if their income-to-needs ratio is below the federally determined threshold for poverty. Furthermore, while adult poverty is *related* to family income and wage, it is a measure of hardship. Among poor children, Head Start spending is associated with large, statistically significant reductions in the annual incidence of poverty in adulthood (Table 2), while Head Start has small, insignificant effects on the adult outcomes of non-poor children (Table 3). However, increases in public K–12 spending are associated with significant reductions in the likelihood of poverty in adulthood for all children, on average.

The final outcome we examine is the probability that an individual has ever been incarcerated (columns 9 and 10 of Tables 2 and 3). In the DiD-2SLS model, for poor children (Table 2), a \$1,000 increase in Head Start spending reduces the likelihood of being incarcerated by 0.6 percentage points (p -value < 0.01). This implies an average Head Start rollout effect (i.e., an increase of \$4,320) of 2.5 percentage points lower likelihood of adult incarceration (at average public K–12 spending level). If one were to ascribe all of this effect to the participation margin for full-year Head Start, it would imply a Head Start participation effect of a 5 percentage point reduction in the probability of ever being incarcerated. Effects of this magnitude are in line with the results from Garces, Thomas, and Currie (2002). Table 2, column 9 also shows that increasing K–12 per-pupil spending by 10 percent (at average Head Start spending levels) reduces the likelihood of adult incarceration by 8 percentage points (p -value < 0.05). The magnitude of this effect is in line with the estimated reductions in incarceration associated with increased schooling (Lochner and Moretti 2004) and reductions in crime associated with attending a better school (Deming 2011). Note, however, that this is the first paper to document a causal relationship between increased public school K–12 spending and reduced risks of adult incarceration. The 2SLS-2SLS models in Table 2, column 10 yield similar patterns, but with somewhat larger Head Start effects and wider confidence intervals. Looking at non-poor children (Table 3), we find no effect of either Head Start or K–12 spending on the likelihood of adult incarceration among non-poor children. We attribute this to the low levels of incarceration among non-poor children. Importantly, *as with the other outcomes*, Head Start spending has no impact on those who were not income-eligible to participate.

B. Testing for Bias Due to Unobserved Family Differences

While we have presented much evidence that our variation is exogenous to other policies that may have been implemented in a locality, we have not *yet* ruled out the

possibility that our results are driven by unobserved differences across treated and untreated families within local areas. To do this, we rely on variation *within* families and compare the outcomes of siblings who were different ages at Head Start rollout or at the time of a court-ordered SFR, but were raised in the same household with the same parents. This approach accounts for observed and unobserved shared family characteristics that predict outcomes. We achieve this by augmenting (6) and (7) to include sibling fixed effects (see online Appendix Table H4). In such models, effects are similar to those in Table 2 so that unobserved family differences cannot explain the main pattern of results.

C. Evidence of Dynamic Complementarity Effects

Before presenting the magnitudes of any complementarity effects, we first establish whether such effects exist. Specifically, in the estimation of (6) and (7), we test whether the coefficient on the interaction is positive and statistically significantly different from zero. Across all outcomes for poor children, and across all specifications, increases in Head Start spending raise the marginal effect of K–12 spending and *vice versa*—that is, all of the interaction terms are statistically significant at, *at least*, the 10 percent level in all models for all outcomes (Table 2). In contrast, there is no such relationship for children from non-poor families (Table 3). For none of the outcomes is the coefficient on the interaction term statistically significant, and the signs of the coefficients across outcomes do not go in the same direction.³⁸ That is, Head Start spending had no direct or indirect effect on the outcomes of non-poor children.

To show the impact of these interaction effects, we present the marginal effects of each intervention evaluated at different levels of the other. Specifically, using the regression estimates, we compute the marginal effect of increasing Head Start spending per poor four-year-old by \$4,230 when there is a 10 percent decrease, no increase, and a 10 percent increase in K–12 spending (conditional on the direct effect of the change in K–12 spending). Similarly, we compute the marginal effect of increasing K–12 spending by 10 percent where there is no Head Start in the county and counties with average Head Start spending (\$4,230). The estimated marginal effects for each model is presented in the lower two panels of Tables 2 and 3. As before, we focus on the DiD-2SLS models, but we present the 2SLS-2SLS models to show robustness.

Looking at high school graduation among poor children, having a Head Start center with a 10 percent decrease in K–12 spending increases high school going by a statistically insignificant 6.3 percentage points. However, having a Head Start center with a 10 percent increase in K–12 spending increases high school going by 14.87 percentage points (p -value < 0.01). The marginal effect of Head Start is more than twice as large when followed by a 10 percent increase in K–12 spending than when followed by a 10 percent decrease. Also, the marginal

³⁸ We formally test that the marginal effects of Head Start and the “HeadStart $\times$ K–12” interaction are different for poor children and non-poor children for years of education and adult wages. We do this by stacking the data and testing for equality of the coefficients. We present this test for years of education and wages in online Appendix K.

effect of Head Start when there is a 10 percent decrease in K–12 spending (though economically meaningful) cannot be distinguished from zero in a statistical sense. These patterns hold in both DiD-2SLS and 2SLS-2SLS models.

We now quantify the interaction concerning the marginal effect of K–12 spending. The DiD-2SLS results indicate that increasing K–12 spending across all school-age years by 10 percent increases the likelihood of graduating high school by 6.7 and 11 percentage points, with and without Head Start, respectively. Similar comparisons for children from non-poor families reveal that the effect of K–12 spending on the outcomes of the non-poor is similar irrespective of the level of Head Start, and Head Start has no effect on the outcomes of the non-poor irrespective of the level of K–12 spending. Because the DiD-2SLS results are similar to, but more conservative than the 2SLS-2SLS estimates, we focus on these models for the remainder of the paper.

The pattern of results for years of completed education is similar to those for high school graduation. The DiD-2SLS results are presented graphically in Figure 6 (the underlying estimates are in Tables 2 and 3). For poor children (panel A), access to the average Head Start center increases completed education by 0.0533 years with a 10 percent reduction in K–12 spending, increases education by 0.32 years with no change in K–12 spending, and increases education by 0.599 years with a 10 percent increase in K–12 spending. While the effect of Head Start with a reduction in K–12 spending cannot be distinguished from zero, the effect when coupled with a 10 percent increase in K–12 spending is statistically significant at the 1 percent level. For non-poor children (panel B), there is no effect of Head Start irrespective of the increase in K–12 spending. Looking at the effect of K–12 spending, for poor children, increasing K–12 spending by 10 percent increase the years of education by 0.13 and 0.4 years, without and with Head Start, respectively. The effect of K–12 spending is more than twice as large among poor individuals exposed to Head Start than those who are not. For children from non-poor families (who are not eligible for Head Start), increasing K–12 spending by 10 percent lead to about 0.23 more years of education irrespective of the Head Start exposure.

In sum, these patterns suggest important dynamic complementarity between early childhood education spending and public K–12 spending for the educational outcomes of poor children. In fact, due to the dynamic complementarity for poor children, the pattern of results indicate that in areas with Head Start programs, increases in K–12 spending both increased outcomes for all students and simultaneously reduced educational attainment gaps. The fact that there is no evidence of complementarity for non-poor children is important. It suggests that our main effects are not simply picking up some strange LATE for those places that happen to be exposed to both high K–12 spending levels and Head Start. If our effects were due to this, one would observe positive interaction effects for all children in such districts. Instead, we find no interaction effects for the non-poor—indicating that our diagnostic tests were likely valid and further supports that our empirical strategy credibly identifies the interaction effects.

Commensurate with the educational outcomes, there is evidence of complementarity between Head Start spending and public K–12 spending in the production of adult economic outcomes for children from poor families. Because for non-poor children

Interactive effects of Head Start and K–12 spending on educational attainment (with 90% CI)

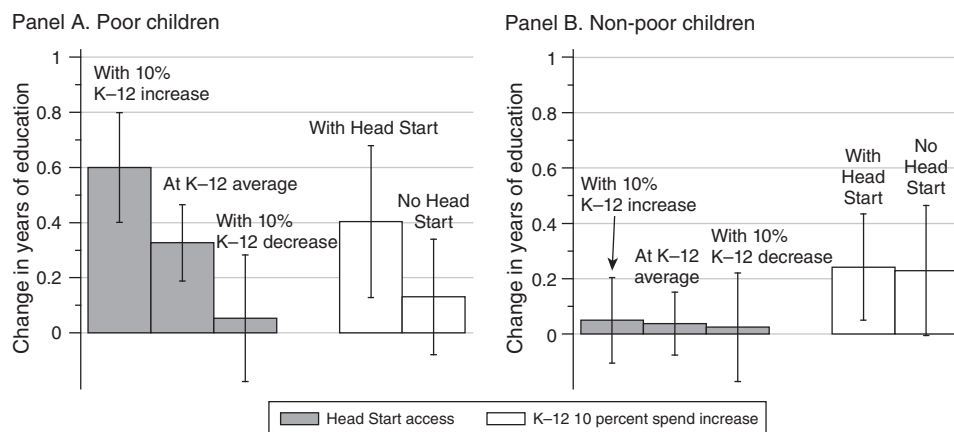


FIGURE 6. EFFECT OF HEAD START SPENDING BY K–12 SPENDING LEVELS AND VICE VERSA ON EDUCATIONAL ATTAINMENT

Notes: The reported marginal effects based upon 2SLS-difference-in-difference model results are presented in columns 3 from Tables 2 and 3. The reported marginal effects and the standard errors were computed using the delta method.

there are no interaction effects for any outcome, we focus on the results for poor children. Figure 7 presents the marginal effect on adult wages of K–12 spending by Head Start access (and vice versa). For poor children (panel A), access to Head Start (with average funding levels) increases adult wages by 2.7 percent (p -value > 0.1) when coupled with a 10 percent K–12 spending decrease, increases it by 9.8 percent when there is no change in K–12 spending (p -value < 0.01), and increases wages by 17 percent when coupled with a 10 percent increase in K–12 spending (p -value < 0.01). The dynamic complementarities are sufficiently large that the marginal effect of the same increases in Head Start spending on the adult wage is about 70 percent larger when K–12 spending increases by 10 percent than with no change. Looking at the effects of K–12 spending increases, a 10 percent increase in K–12 spending leads to 13 percent higher wages without Head Start, and 20 percent higher wages with Head Start (both effects are significant at the 1 percent level).

The effects on the annual incidence of adult poverty are consistent with those on education and wages (columns 7 and 8 of Table 2). For poor children, increasing Head Start spending from zero to average levels reduces the annual incidence of poverty in adulthood by about 3 percentage points (p -value > 0.1) when coupled with a 10 percent reduction in K–12 spending, a 7.6 percentage point reduction when coupled with no change in K–12 spending (p -value < 0.01), and reduces adult poverty by 12 percentage points when coupled with a 10 percent increase in K–12 spending (p -value < 0.01). The marginal effects of K–12 spending tell the same story. A 10 percent increase in K–12 spending leads to 3.3 and 7.96 percentage points lower adult poverty without and with Head Start, respectively. The effect of the K–12 spending increase with Head Start is significant at the 1 percent level and is more than twice as large as the effect with no Head Start.

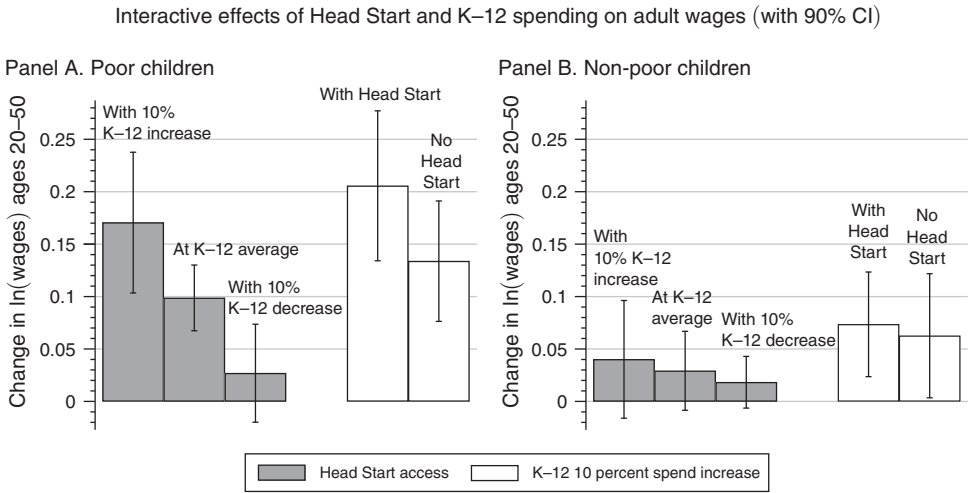


FIGURE 7. EFFECT OF HEAD START SPENDING BY K-12 SPENDING LEVELS AND VICE VERSA ON WAGES

Notes: The reported marginal effects based upon 2SLS-difference-in-difference model results are presented in columns 5 from Tables 2 and 3. The reported marginal effects and the standard errors were computed using the delta method.

As with the other adult outcomes, the reduction in the lifetime risks of incarceration associated with improvements in access to early education is larger when there are greater subsequent K-12 school investments and *vice versa*. The marginal effects are presented in columns 9 and 10 of Table 2. For poor children, increasing Head Start spending from zero to average levels has no effect on the likelihood of incarceration when coupled with a 10 percent reduction in K-12 spending. However, this same increase in Head Start exposure leads to a 2.5 percentage point reduction when coupled with no change in K-12 spending (p -value < 0.1), and a 4.73 percentage point reduction when coupled with a 10 percent increase in K-12 spending (p -value < 0.01). Looking to the effect of K-12 spending on the likelihood of being incarcerated, the marginal effects are larger with Head Start than without. A 10 percent increase in K-12 per pupil spending reduces the likelihood of being incarcerated by 5.8 percentage points with no Head Start spending (p -value < 0.05), and by eight percentage points with Head Start (p -value < 0.01).

D. Is Parenting Quality Part of the Story?

Because parent counseling was a component of Head Start, it is possible that these dynamic complementarities emerge through improvements in parenting quality. Because we have data on siblings with the same parents, we can test for improvements in parenting quality. We use only the sample of older siblings who were not themselves exposed to Head Start and test whether those with younger siblings who were exposed to Head Start have improved outcomes. If improvements in parenting quality is a part of the story, the older siblings of exposed younger

siblings should have better outcomes than the older siblings of unexposed younger siblings. However, if the Head Start effects are driven by the services provided to the children, there should be no effect. In these models (online Appendix I), we find older siblings are unaffected by Head Start exposure of the younger sibling. This suggests that parenting quality is not part of the story; our Head Start spending effects reflect real investments in the human capital of poor children; and our effects are not due to other confounding policies aimed at poor children.

E. Are the Complementarity Effects Driven by Other Coincident Policies?

Even though our estimation equations control for several coincident policies directly, one may worry that our main results are driven by some complementarity between K–12 spending and some other policy. To test for this directly, we augment our main model in equation (6) to also include interactions between food stamp spending in one's county between ages 0 to 4 with K–12 spending, and county-level spending on Medicaid between ages 0 and 4 interacted with K–12 spending. In these models, the point estimates on the interaction between Head Start spending and K–12 spending are virtually unchanged. This provides further evidence that our estimated effects are not confounded by dynamic complementarities with other policies.

VI. Benefit-Cost Considerations: Putting the Magnitudes in Perspective

It is helpful to consider how the presence of dynamic complementarity affects the optimal allocation of resources to the K–12 system versus to early childhood education (for poor children). In any given location, if average outcomes are maximized, the marginal dollar spent on Head Start will yield the same effect on outcomes as an equivalent expenditure on K–12 education. Using the estimated impacts from Tables 2 and 3, we compute the marginal impact on the average outcomes in a county of establishing a typical Head Start center that only poor children benefit from π_{HS} . We also compute the marginal impact of spending that same amount of money (in present value terms) in the K–12 system that all students attend π_{K-12} . We then compute the ratio of these marginal impacts, π_{HS}/π_{K-12} , for different poverty levels p . See online Appendix N for details of this calculation. When spending is allocated optimally, this ratio should be 1.

In Figure 8, we plot this ratio against the poverty rate, where this ratio is evaluated at the mean level of K–12 (i.e., using the empirical estimates from Table 2). We show this for adult wages (effects are similar for other outcomes). Because our empirical model is linear in Head Start spending but linear in the log of K–12 spending, the marginal effect of K–12 spending will fall relative to that for Head Start at higher levels of K–12 spending even without any complementarity. To show this relationship, on the left, we impose the condition that there is no interaction effect and then plot the resulting π_{HS}/π_{K-12} against the poverty rate where the present value is evaluated at the average K–12 spending levels, 10 percent above this average and 10 percent below this average.

As one would expect, in Figure 8, panel A, this ratio is falling with the poverty rate. This reflects the fact that K–12 spending has a larger effect on poor children, so

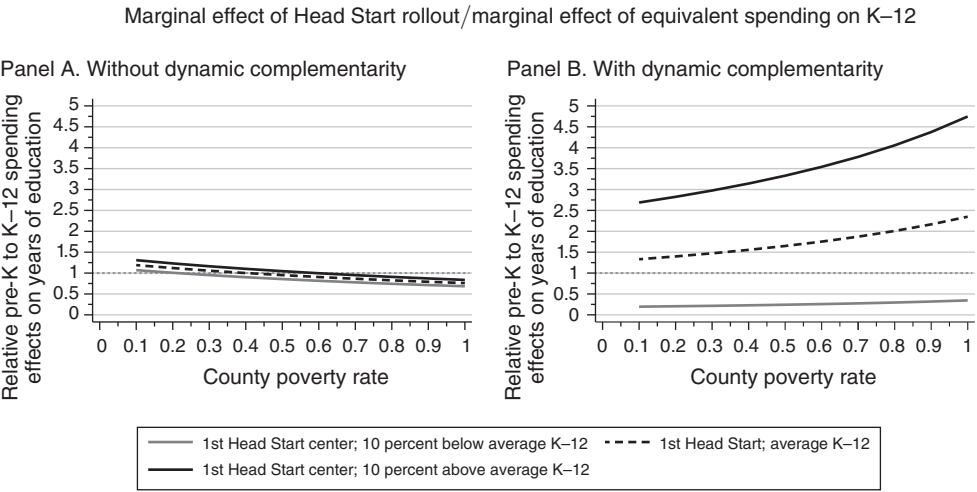


FIGURE 8. RATIO BETWEEN THE EFFECT OF HEAD START SPENDING AND K-12 SPENDING LEVELS BY POVERTY LEVEL IN THE COUNTY

Notes: The reported marginal effects are based upon 2SLS-difference-in-difference model results presented in columns 2 and 3 from Tables 2 and 3. The solid gray lines plot the ratio between the marginal effect of spending on Head Start and the effect of spending that same amount on the K-12 system (in present value terms). This ratio presented in the solid gray line is evaluated at average levels of Head Start spending and K-12 spending during the sample period. The dashed gray line presents this same ratio evaluated at \$1,000 above the average K-12 spending levels assuming no dynamic complementarity, while the solid black line presents this ratio evaluated at \$1,000 above the average K-12 spending levels using the estimated interaction effects. The difference between the solid black lines and the dashed gray lines reflect the marginal contribution of dynamic complementarity to changes in this ratio as one increases K-12 spending above the average.

that the average benefits of K-12 spending are larger in higher poverty areas. Also as expected, (even where there is dynamic complementarity) the ratio is higher when evaluated at higher levels of K-12 spending. Interestingly, with no dynamic complementarity, the relative marginal benefit of rolling out a Head Start center lies below that of K-12 spending so long as the poverty rate is above about 20 percent. With no dynamic complementarity, this is true even in areas that spend 10 percent above average in the K-12 system. To illustrate how dynamic complementarity affects these ratios, we allow for dynamic complementarity (i.e., using the estimated interaction term from Table 2) and then evaluate these same ratios (panel B). Evaluated at the average, the basic pattern is the same. However, with dynamic complementarity, the ratios are very different at K-12 spending levels 10 percent above and below the average. Where complementarities exist, in areas that spending 10 percent higher than average in the K-12 system, this ratio lies above 1 at all poverty levels, so that the marginal impact of rolling out Head Start on average wages is larger than the effects of spending that same money in the K-12 system. The flipside of this result is that in areas that spend less than 10 percent lower than the average, this ratio lies below 1 for all poverty levels. This means that in areas with low levels of K-12 spending, the marginal dollar is better spent in the K-12 system than on Head Start.

In essence, these patterns support the idea that, when such dynamic complementarities exist between early and late human capital investments, in some locations, there may be no equity-efficiency trade-off when shifting resources

toward compensatory early education programs (Cunha and Heckman 2007). More specifically, our estimates indicate that, for a district that spent \$4,500 per-pupil (about 10 percent above the average K–12 spending level), the marginal dollar spent on Head Start led to between 1.5 and 2.5 times the improvement in adult outcomes as that spent on K–12 education. Accordingly, at such spending levels, one could redistribute money from the K–12 system toward Head Start and have *both* better average outcomes and a more equitable distribution of adult outcomes. Overall, the patterns in Figure 8 suggests that during our sample period, the marginal dollar had a roughly equal effect on adult outcomes overall at levels close to the national averages that prevailed at that time. The patterns also indicate that, when resources are allocated efficiently, localities with higher levels of Head Start spending should have higher levels of K–12 spending and vice versa. Empirically, the correlation between per pupil spending and Head Start spending is roughly 0.35. This implies that, in general, localities may be taking advantage of these complementarities, but that further optimization is likely possible.

Summary and Conclusions.—This study provides new evidence on the life-cycle effects of Head Start and K–12 school spending. We explore dynamic complementarities between human capital investments made in preschool and those that subsequently occur in the K–12 system. We use children’s differential exposure to Head Start spending (at age four) and court-ordered school finance reforms (SFRs) (between the ages 5–17), depending on place and year of birth, to examine whether the marginal effect of Head Start spending on children’s adult outcomes are larger among individuals who were subsequently exposed to SFR-induced K–12 spending increases. We present extensive tests to document that the policy-induced variation in Head Start spending and K–12 public school spending we exploit is unrelated to other childhood family, community, or policy changes.

For non-poor children, SFR-induced K–12 spending increases led to significant improvements in educational and economic outcomes, while increases in Head Start spending had no effect. However, for poor children, both Head Start spending increases and SFR-induced K–12 spending increases led to significant improvements in educational outcomes, economic outcomes, and reductions in the likelihood of incarceration. Importantly, the long-run effects of increases in Head Start spending are amplified when followed by attending schools that experienced SFR-induced increases in K–12 per pupil spending. Across all the outcomes, the marginal effect of the same increase in Head Start spending was more than twice as large for students from K–12 school districts that spent at the seventy-fifth percentile of the distribution than those from K–12 school districts that spent at the twenty-fifth percentile. Similarly, the benefits of K–12 school-spending increases on adult outcomes were larger among poor children who were exposed to higher levels of Head Start spending during their preschool years. For poor children, the combined benefits of growing up in districts/counties with *both* greater Head Start spending and K–12 per pupil spending are significantly greater than the sum of the independent effects of the two investments in isolation.

There are two important caveats to our work. First, because the counterfactual childcare and pediatric care may be better today than in the late 1960s and 1970s, the marginal effect of Head Start may be smaller today than in the earlier period

that we study.³⁹ Second, public school spending levels during the period we study were lower than current levels. If school spending exhibits diminishing marginal product, the effects presented here may be larger than one would observe with similar spending increases today. These caveats do not minimize the importance of the findings or their profound implications for policy. However, they do suggest that the contemporary magnitude of the effects may be smaller than those we present here. At the same time, the returns to education have increased, so the consequences of access to high-quality human capital investments from pre-K–K-12 are large.

The cumulative nature of skill development is likely responsible for the pattern of results. Our findings highlight the importance of modeling early and later educational investments jointly and may explain some disparate results on the effects of Head Start. Indeed, our finding that the long-run effects of Head Start are larger among individuals who attended better resourced schools may provide an explanation for why Head Start may have been more successful for more socio-economically advantaged populations (Currie and Thomas 1995) and why there is a fade out of the effects of Head Start on test scores as students age (Currie and Thomas 2000). The key policy implication of our findings is that human capital investments made in, and sustained throughout, child developmental stages (preschool, elementary/middle school, adolescence) may yield greater returns than separate, isolated, short-lived reforms not sustained beyond the year in which they are implemented. The findings point to the critical role early-life investments can play in narrowing long-run gaps in well-being, and they also highlight the importance of sustained investments in the skills of disadvantaged youth.

REFERENCES

- Adhvaryu, Achyuta, Steven Bednar, Teresa Molinari, Quynh Nguyen, and Anant Nyshadham. Forthcoming. "When It Rains It Pours: The Long-Run Economic Impacts of Salt Iodization in the United States." *Review of Economics and Statistics*.
- Aizer, Anna, and Flávio Cunha. 2012. "The Production of Human Capital: Endowments, Investments and Fertility." NBER Working Paper 18429.
- Almond, Douglas, and Bhashkar Mazumder. 2013. "Fetal Origins and Parental Responses." *Annual Review of Economics* 5 (1): 37–56.
- Angrist, Joshua D., and Alan B. Krueger. 1992. "The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples." *Journal of the American Statistical Association* 87 (418): 328–36.
- Angrist, Joshua, and Steve Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.
- Bhalotra, Sonia R., and Atheendar Venkataramani. 2015. "Shadows of the Captain of the Men of Death: Early Life Health Interventions, Human Capital Investments, and Institutions." <https://ssrn.com/abstract=1940725> or <http://dx.doi.org/10.2139/ssrn.1940725>.
- Björklund, Anders, and Markus Jäntti. 1997. "Intergenerational Income Mobility in Sweden Compared to the United States." *American Economic Review* 87 (5): 1009–18.

³⁹ In the early period of Head Start, most poor children would have received home care, while today, as many as one-third of Head Start participants may have attended some other form of formal childcare (Feller et al. 2016, Kline and Walters 2016). The proportion of three- and four-year-olds in school has increased from roughly 10 percent in 1964 to almost 40 percent by 1995 (source: US Census Bureau, CPS October Supplement, 1964–2010; see Figure 1). Also, while most poor children currently receive pediatric care through Medicaid and SCHIP, during the period under study many children would only have received such care through Head Start.

- Blau, David, and Janet Currie.** 2006. "Pre-school, Day Care, and After-School Care: Who's Minding the Kids?" In *Handbook of the Economics of Education*, Vol. 2, edited by Eric Hanushek and Finis Welch, 1163–1278. Amsterdam: North-Holland.
- Card, David, and A. Abigail Payne.** 2002. "School Finance Reform, the Distribution of School Spending, and the Distribution of Student Test Scores." *Journal of Public Economics* 83 (1): 49–82.
- Carneiro, Pedro, and Rita Ginja.** 2014. "Long-Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start." *American Economic Journal: Economic Policy* 6 (4): 135–73.
- Chay, Kenneth Y., Jonathan Guryan, and Bhashkar Mazumder.** 2009. "Birth Cohort and the Black-White Achievement Gap: The Roles of Access and Health Soon after Birth." NBER Working Paper 15078.
- 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." *Quarterly Journal of Economics* 126 (4): 1593–1660.
- Chetty, Raj, and Nathaniel Hendren, Patrick Kline, and Emmanuel Saez.** 2014. "Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *Quarterly Journal of Economics* 129 (4): 1553–1623.
- Cohodes, Sarah R., Daniel S. Grossman, Samuel A. Kleiner, and Michael F. Lovenheim.** 2015. "The Effect of Child Health Insurance Access on Schooling: Evidence from Public Insurance Expansions." *Journal of Human Resources* 51 (3): 556–88.
- Cunha, Flavio, and James Heckman.** 2007. "The Technology of Skill Formation." *American Economic Review* 97 (2): 31–47.
- Currie, Janet, and Duncan Thomas.** 1995. "Does Head Start Make a Difference?" *American Economic Review* 85 (3): 341–64.
- Currie, Janet, and Duncan Thomas.** 2000. "School Quality and the Long-Term Effects of Head Start." *Journal of Human Resources* 35 (4): 755–74.
- Currie, Janet, and Duncan Thomas.** 2001. "Early Test Scores, Socioeconomic Status, School Quality and Future Outcomes." *Research in Labor Economics* 20: 103–32.
- Currie, Janet, and Aaron Yelowitz.** 2000. "Are Public Housing Projects Good for Kids?" *Journal of Public Economics* 75 (1): 99–124.
- Dee, Thomas S., and William N. Evans.** 2003. "Teen Drinking and Educational Attainment: Evidence from Two-Sample Instrumental Variables Estimates." *Journal of Labor Economics* 21 (1): 178–209.
- Deming, David J.** 2009. "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start." *American Economic Journal: Applied Economics* 1 (3): 111–34.
- Deming, David J.** 2011. "Better Schools, Less Crime?" *Quarterly Journal of Economics* 126 (4): 2063–2115.
- Feller, Avi, Todd Grindal, Luke Miratrix, and Lindsay C. Page.** 2016. "Compared to What? Variation in the Impacts of Early Childhood Education by Alternative Care Type." *Annals of Applied Statistics* 10 (3): 1245–85.
- Fryer, Roland G., Jr., and Steven D. Levitt.** 2004. "Understanding the Black-White Test Score Gap in the First Two Years of School." *Review of Economics and Statistics* 86 (2): 447–64.
- Fryer, Roland G., Jr., and Steven D. Levitt.** 2006. "The Black-White Test Score Gap through Third Grade." *American Law and Economic Review* 8 (2): 249–81.
- Garces, Eliana, Duncan Thomas, and Janet Currie.** 2002. "Longer-Term Effects of Head Start." *American Economic Review* 92 (4): 999–1012.
- García, Jorge Luis, and Sebastian Gallegos.** 2017. "Dynamic Complementarity or Substitutability? Parental Investment and Childcare in the Production of Early Human Capital." <https://ssrn.com/abstract=2910167>.
- Gilraine, Michael.** 2016. "School Accountability and the Dynamics of Human Capital Formation." https://pdfs.semanticscholar.org/ce2d/184ed9ea45b145ef7574320e2cedadd1008c.pdf?_ga=2.187460207.674734485.1568384816-88954882.1568384816.
- Heckman, James J.** 2007. "The Economics, Technology, and Neuroscience of Human Capability Formation." *Proceedings of the National Academy of Sciences* 104 (33): 13250–255.
- Heckman, James J., and Stefano Mosso.** 2014. "The Economics of Human Development and Social Mobility." *Annual Review of Economics* 6: 689–733.
- Heckman, James, Rodrigo Pinto, and Peter Savelyev.** 2013. "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes." *American Economic Review* 103 (6): 2052–86.
- Howell, Penny L., and Barbara B. Miller.** 1997. "Sources of Funding for Schools." *Future of Children* 7 (3): 39–50.
- Hoxby, Caroline M.** 2001. "All School Finance Equalizations Are Not Created Equal." *Quarterly Journal of Economics* 116 (4): 1189–1231.

- Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond. 2016. "Long-Run Impacts of Childhood Access to the Safety Net." *American Economic Review* 106 (4): 903–34.
- Jackson, C. Kirabo. 2018. "What Do Test Scores Miss? The Importance of Teacher Effects on Non-test Score Outcomes." *Journal of Political Economy* 126 (5): 2072–2107.
- Jackson, C. Kirabo, Rucker C. Johnson, and Claudia Persico. 2014. "The Effect of School Finance Reforms on the Distribution of Spending, Academic Achievement, and Adult Outcomes." NBER Working Paper 20118.
- Jackson, C. Kirabo, Rucker C. Johnson, and Claudia Persico. 2016. "The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms." *Quarterly Journal of Economics* 131 (1): 157–218.
- Johnson, Rucker C. 2019. *Children of the Dream: Why School Integration Works*. New York: Basic Books and Russell Sage Foundation Press.
- Johnson, Rucker C., and C. Kirabo Jackson. 2019. "Reducing Inequality through Dynamic Complementarity: Evidence from Head Start and Public School Spending: Dataset." *American Economic Journal: Economic Policy*. <https://doi.org/10.1257/pol.20180510>.
- Kline, Patrick, and Christopher R. Walters. 2016. "Evaluating Public Programs with Close Substitutes: The Case of Head Start." *Quarterly Journal of Economics* 131 (4): 1795–1848.
- Knudsen, Eric I., James J. Heckman, Judy Cameron, and Jack P. Shonkoff. 2006. "Economic, Neurobiological, and Behavioral Perspectives on Building America's Future Workforce." *Proceedings of the National Academy of Sciences* 103 (27): 10155–62.
- Levine, Phillip B., and Diane Schanzenbach. 2009. "The Impact of Children's Public Health Insurance Expansions on Educational Outcomes." *Forum for Health Economics and Policy* 12 (1): 1–28.
- Lipsey, Mark W., Dale Clark Farran, and Kerry G. Hofer. 2015. *A Randomized Control Trial of a Statewide Voluntary Prekindergarten Program on Children's Skills and Behaviors through Third Grade: Research Report*. Nashville: Vanderbilt University, Peabody Research Institute.
- Lochner, Lance, and Enrico Moretti. 2004. "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports." *American Economic Review* 94 (1): 155–89.
- Lubotsky, Darren, and Robert Kaestner. 2016. "Do 'Skills Beget Skills'? Evidence on Dynamic Complementarities in Cognitive and Non-cognitive Skills in Childhood." *Economics of Education Review* 53: 194–206.
- Ludwig, Jens, and Douglas L. Miller. 2007. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design." *Quarterly Journal of Economics* 122 (1): 159–208.
- Magnuson, Katherine A., and Jane Waldfogel. 2005. "Early Childhood Care and Education: Effects on Ethnic and Racial Gaps in School Readiness." *Future of Children* 15 (1): 169–96.
- Malamud, Ofer, Cristian Pop-Eleches, and Miguel Urquiola. 2016. "Interactions between Family and School Environments: Evidence on Dynamic Complementarities?" NBER Working Paper 22112.
- McLeod, Jane D., and Karen Kaiser. 2004. "Childhood Emotional and Behavioral Problems and Educational Attainment." *American Sociological Review* 69 (5): 636–58.
- Murray, Sheila E., William N. Evans, and Robert M. Schwab. 1998. "Education-Finance Reform and the Distribution of Education Resources." *American Economic Review* 88 (4): 798–812.
- Neidell, Matthew, and Jane Waldfogel. 2010. "Cognitive and Non-cognitive Peer Effects in Early Education." *Review of Economics and Statistics* 92 (3): 562–76.
- Newport, Elissa L. 1990. "Maturational Constraints on Language Learning." *Cognitive Science* 14 (1): 11–28.
- Panel Study of Income Dynamics. 1968–2015. Restricted use data. Survey Research Center, Institute for Social Research, University of Michigan.
- Pinker, Steven. 1994. *The Language Instinct*. New York: Harper Perennial Modern Classics.
- Rossin-Slater, Maya, and Miriam Wüst. 2016. "What Is the Added Value of Preschool? Long-Term Impacts and Interactions with a Health Intervention." IZA Discussion Paper 10254.
- Sacerdote, Bruce. 2014. "Experimental and Quasi-experimental Analysis of Peer Effects: Two Steps Forward?" *Annual Review of Economics* 6: 253–72.
- Shonkoff, Jack P., and Deborah A. Phillips, eds. 2000. *From Neurons to Neighborhoods: The Science of Early Childhood Development*. National Research Council and Institute of Medicine Committee on Integrating the Science of Early Childhood Development. Washington, DC: National Academies Press.
- US Census Bureau. 2018. Current Population Survey, October Supplement, 1964–2010.
- Zigler, Edward, Walter S. Gilliam, and W. Steven Barnett, eds. 2011. *Current Debates and Issues in Prekindergarten Education*. Baltimore: Brookes Publishing.
- Zigler, Edward, and Sally J. Styfco. 2010. *The Hidden History of Head Start*. New York: Oxford University Press.