REDUCING INEQUALITY THROUGH DYNAMIC COMPLEMENTARITY: EVIDENCE FROM HEAD START AND PUBLIC SCHOOL SPENDING

RUCKER C. JOHNSON AND C. KIRABO JACKSON*

February 14, 2017

We explore whether early childhood human-capital investments are complementary to those made later in life. Using the Panel Study of Income Dynamics, we compare the adult outcomes of cohorts who were differentially exposed to policy-induced changes in pre-school (Head Start) spending and school-finance-reform-induced changes in public school spending during childhood, depending on place and year of birth. Difference-in-difference instrumental variables and sibling-difference estimates indicate that, for poor children, increases in Head Start spending and increases in public K12 spending each individually increased educational attainment and earnings, and reduced the likelihood of both poverty and incarceration in adulthood. The benefits of Head Start spending were larger when followed by access to better-funded public K12 schools, and the increases in K12 spending were more efficacious for poor children who were exposed to higher levels of Head Start spending during their preschool years. The findings suggest that early investments in the skills of disadvantaged children that are followed up with sustained educational investments over time can effectively break the cycle of poverty. (*JEL* I20, J20)

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, Hendren, and Katz, 2016). 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, 1999; McLeod and Kaiser, 2004; Heckman and Mosso, 2014). Accordingly, efficiently breaking the cycle of poverty may require early investments in the skills of disadvantaged children that are followed up with sustained investments over time. In this paper, we study whether early childhood investments for disadvantaged children that are followed up with increases in public school spending are particularly effective at improving their long-run outcomes.

* Johnson: University of California, Berkeley and NBER; Jackson: Northwestern University and NBER. 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 University of Michigan, UC-Berkeley, Vanderbilt University, University of Wisconsin/IRP Summer Workshop, and UC-Davis for helpful comments; and thank Martha Bailey & Andrew Goodman-Bacon, and Doug Miller & Jens Ludwig, for sharing data on 1960 county poverty rates and data from the National Archives & Records Administration. Johnson and Jackson are equal co-authors. This research was supported by the National Institutes of Health and the Russell Sage Foundation (to Johnson), and the National Science Foundation (to Jackson).

Theory supports the prescription for early and sustained investments in human capital for disadvantaged youth. Cunha and Heckman (2007) present a model of human capital production in which early human capital investments are complementary to those made later in life. If such complementarities exist, then early human capital investments make subsequent human capital investments more productive and may be ineffective if not followed by subsequent investments.[1] However, there has been little experimental or quasi-experimental exploration into how the efficacy of policies that promote human capital in early life are affected by policies that promote human capital later in life and vice versa.[2] To fill this gap, we exploit two exogenous human capital investment "shocks" that occur at different points in the life course to explore whether early childhood human capital investments are complementary to those made later in childhood—that is, we explore whether human capital investments exhibit dynamic complementarity.

The first independent exogenous shocks to human capital investment are the roll-out of Head Start, which increased access to early childhood education and pediatric care for low-income children. The second independent exogenous shocks to human capital investment are the implementation of court-ordered school finance reforms (SFRs) which (on average) increased the level of per-pupil spending at public K12 schools (Jackson, Johnson, and Persico, 2014). The first policy, Head Start, is the largest early-childhood intervention program in the US. Head Start is a comprehensive, national, federally-funded early childhood program that was established in 1964 as part of Lyndon B. Johnson's "War on Poverty" to provide education, health care, nutrition, and other services to poor children before kindergarten entry. The second policy we exploit is court-ordered school finance reforms (SFRs). In the 1960s, the majority of public school spending was funded through local property taxes so that low-wealth areas tended to have lower per-pupil K12 spending levels than more affluent high-wealth areas. The court-ordered SFRs that began in the early 1970s (and continue to present day) changed the parameters of spending formulas. These changes reduced inequality in school spending and weakened the relationship between the level of public school spending and the wealth and income level of the district (Card and Payne, 2002;

[1] If early childhood investments for disadvantaged youth make subsequent investments in these children more productive, one can justify policies that redistribute resources toward disadvantaged children during their early years on *efficiency* grounds without any appeal to equity issues, fairness, or social justice (Heckman and Mosso, 2014).

[2] Existing studies on dynamic complementarities compare the marginal effects of educational interventions across individuals with different incoming stocks of human capital (Aizer and Cunha, 2012; Malamud and Pop-Eleches, 2015; Lubotsky and Kaestner, 2015). In contrast, we study the interaction between two human capital policies over the life cycle rather than explore the effectiveness of a single policy across different populations.

Murray, Evans, and Schwab, 1998; Hoxby, 2001; Jackson, Johnson, and Persico, 2014). Each of these two policies led to dramatic changes in the structure of public education funding in the United States. We explore the combined effect of the two.

To isolate the effect 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 2013 using the Panel Study of Income Dynamics (PSID). Test scores have been the traditional focus of evaluations of Head Start and K12 spending. However, the effects of interventions on long-run outcomes may go undetected by test scores (e.g., Heckman, Pinto, and Savelyev, 2014; Deming, 2009; Jackson, 2012, 2016; Chetty et al., 2011; Ludwig and Miller, 2007). As such, we explore effects on an array of adult outcomes including educational attainment, earnings, adult poverty, and incarceration.[3]

We seek to identify the potential *interactive,* or synergistic, effects between early and later human capital investments. Identifying such interaction effects requires that we credibly identify the effects of each human capital investment individually. To identify the causal effect of early childhood investments, we exploit geographic variation in the timing of the roll-out of Head Start across counties. In our preferred difference-in-difference models, we compare the adult outcomes of individuals from the same childhood county who were exposed to different levels of Head Start spending, because some were four years old when Head Start spending levels were low (or non-existent) while others were four years old when Head Start spending levels were higher. We present several empirical tests to show that the identifying variation in Head Start spending is unrelated to family, community, and policy changes. We also show that our Head Start effects are robust to (a) instrumental variables models that only use variation in Head Start spending due to the timing of Head Start rollout in the childhood county and (b) using within-family across-sibling variation in Head Start spending exposure.

To identify the causal effect of public K12 school spending, we exploit geographic variation in the timing of court-ordered SFRs. Following Jackson, Johnson, and Persico (2016), we predict the spending change that each district would experience after the passage of court-mandated SFR based on the type of reform and the characteristics of the district before reforms.

[3] In a recent related study Carneiro and Ginja (2014) study the long-run effects of Head Start participation on health and behavioral problems. In our study we explore a wider array of adult outcomes such as earnings, family income, and child outcomes. Also, a key focus of our paper is on how the effectiveness of Head Start spending varies by the quality of the public schools students subsequently attend.

Using instrumental variables models, we see if 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 at passage) in districts predicted to experience larger reform-induced spending increases. We present several empirical tests showing that the within-district variation in per-pupil spending induced by SFRs is exogenous to other family, community, and policy changes in the district. We also show that our K12 school spending effects are robust to using within-family across-sibling variation in SFR-induced K12 public-school spending.

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 Head Start spending and public K12 spending. We can test for dynamic complementarities based on two sources of variation. Namely, some districts experienced increases in school spending due to a SFR when Head Start was available in the county, while other districts experienced similar K12 spending increases when Head Start was not available. This fact allows one to test if the effects of K12 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 K12 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 K12 spending due to the passage of a court-ordered SFR.

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. Consistent with Jackson, Johnson, and Persico (2016), increases in public school K12 spending improved this same array of outcomes in adulthood. We also find robust evidence of dynamic complementarity. Intent-to-treat estimates indicate that for children from low-income families, on average, increasing Head Start spending by $1,000 per poor four-year-old (in the county) increases educational attainment by 0.096 years, increases adult wages by 1.9 percent, and reduces the likelihood of incarceration by 0.75 percentage points. However, in districts at the 90th percentile of the K12 spending distribution, the same increase in Head Start spending increases educational attainment by 0.3 years, increases wages by 8 percent, and reduces the likelihood of incarceration by 3.4 percentage points (after accounting for the direct effect of higher levels of K12 spending). The fact that the long-run benefits of Head Start spending depend on the subsequent level of K12 spending may help explain why some studies find positive effects

of Head Start.[4] Looking at the marginal effects of K12 spending, for low-income children, increasing public K12 spending by 10 percent has small effects on educational attainment, adult wages, or incarceration when not preceded by Head Start. However, among low-income children exposed to Head Start spending at the 90th percentile of the distribution, the same 10 percent increase in K12 per-pupil spending increases educational attainment by 0.52 years, increases earnings by 17 percent, and reduces the likelihood of incarceration by 16.6 percentage points. The patterns of positive interaction effects between Head Start and K12 spending are robust across several empirical models (including sibling comparisons) and are only present among children from low-income families. The effect of K12 spending was unrelated to the level of Head Start spending among more affluent children–for those children, increasing K12 spending by 10 percent increased years of education by 0.2 and earnings by 11.7 percent.

We find substantial long-run benefits of public early childhood investments during the 1970s and 1980s and robust evidence of complementarities between early and later human capital investments for low-income children. The results imply, as suggested by Heckman and Mosso (2014), that investments in the early-childhood development of low-income children may not exhibit an equity-efficiency tradeoff, and that early and sustained investments in the skills of low-income children can be a cost-effective strategy to break the cycle of poverty.

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

II. 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, non-cognitive, and health formation (Shonkoff and Phillips, 2000). Evidence suggests that common developmental processes operate in the formation of cognitive, non-cognitive, and health capacities. Informed by this research and others, Cunha and Heckman (2007), theorize that skill development is an interactive, multistage process such that

[4] For positive effects see Deming (2009), Ludwig and Miller (2007), Currie and Thomas (2000), Garces, Currie, and Thomas, (2002), Carneiro and Ginja (2014). For mixed effects see Zigler et al., (2011), Farran and Lipsey (2015).

early-life human capital investments may cross-fertilize human capital investments made at later childhood stages, leading to developmental synergy effects. From this perspective, if the early childhood human capital investments provided by Head Start improve school readiness, they may facilitate better learning in the K12 system. If so, insofar as increased spending improves school quality, spending on Head Start and public K12 schools would be synergistic and would exhibit dynamic complementarities. We formalize this logic below.

Following the notation of Heckman (2007), the technology of skills production is dynamic. Skills acquired when a child is *t* years old is [1] below

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

where *t*=1,2,...T, θ_t is a vector of skills at time *t*, parental capabilities are connoted by h_t, and investments during time *t* is connoted by I_t. Investments in time *t* (I_t) is 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 skill production arises when the stocks of capabilities acquired by period *t-1* (θ_t) make investments in period t (I_t) more productive, i.e.,

[2] $$(\partial\theta_{t+1})/(\partial\theta_t\partial I_t\ ) > 0.$$

It is important to note that unlike other studies, we do not seek to identify $(\partial\theta_{t+1})/(\partial\theta_t\partial I_t\ )$.[5] Instead, we explore the closely-related ***policy*** question of how public human capital investments made in early childhood affect the efficacy of those made in later developmental stages of childhood. Consider that $\theta_t = f_{t-1}(h_{t-1}, \theta_{t-1}, I_{t-1})$. Because $\partial f_t/\partial I_t > 0$ from above, if [2] holds, then equation [3] below must also hold.

[3] $$(\partial\theta_{t+1})/(\partial I_{t-1}\partial I_t\ ) > 0.$$

In words, dynamic complementarity in skill production implies that there is dynamic complementarity in human capital investments.[6]

[5] Equation [2] makes clear that to credibly test for dynamic complementarity in skill production, one requires exogenous variation in both skills (θ_t) and investments (I_t). Existing studies (Aizer and Cunha, 2012; Malamud and Pop-Eleches, 2015; Lubotsky and Kaestner, 2015) compare the marginal effects of exogenous educational interventions (I_t) across individuals with different incoming stocks of human capital. The empirical challenge faced in these papers is that variation in measured human capital may reflect a variety of factors other than θ_t.

[6] Note that if early investments can increase the efficacy of later investments through mechanisms other than increasing skills, the converse does not *necessarily* hold. While unlikely, it is a theoretical possibility. For example, suppose an increase in Head Start spending led to an increase in the supply of pre-K and K12 teachers (perhaps due to the creation of a new teacher training program to meet the new demand), then increased Head Start spending would reduce the amount of money it would require to attract quality K12 teachers –thus making each marginal dollar in the K12 system

In this paper, we test empirically for dynamic complementarity between early and later childhood human capital investments; in particular, between public investments in spending on early education for poor children (i.e., Head Start) in the public K12 system. We hypothesize that these two human capital policies may exhibit dynamic complementarity through a direct skill acquisition channel, and an indirect spillover channel.

The direct channel is what we call the "alignment" channel. This channel is predicated on the idea that the sequence of when skills are taught matters (Knudsen et al. 2006, Newport 1990; Pinker 1994) and the fact that K12 systems target students with a specific incoming skill level. Students above the target skill level may benefit less from the K12 system (the K12 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 K12 system (the instruction may assume required skills that 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 K12 education system. Furthermore, access to pediatric care (provided to Head Start participants) may promote this skill development (Levine and Shanzenbach 2009; Cohodes, Grossman, Kleiner, Lovenheim, 2015).

Through this alignment channel, 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 K12 schools. In sum, through this channel, on average, the effects of Head Start spending may be larger in well-funded K12 districts and could be negligible in poorly funded public school districts.

The indirect channel is through "spillover effects." In general, research has found that higher shares of low-performing peers or disruptive peers, and high levels of heterogeneity in

more effective. Any such non-skills mechanism would generate spillover effects to non-Head Start participants that we can test for empirically. We present several empirical tests that such mechanisms are not a factor in section X.

ability levels in the classroom have deleterious impacts on student outcomes (see Sacerdote 2014) for an overview of this literature). By increasing the human capital of poor children, increases in Head Start spending may affect the subsequent peer composition of the K12 classrooms for *all* children in the county. If higher levels of Head Start spending reduce the likelihood of having low-achieving peers or disruptive peers or lessen the degree of heterogeneity in the classrooms, it could make it easier for the K12 school system to translate additional resources into improvements in outcomes. Neidell and Waldfogel (2010) provide evidence of this channel by documenting spillover effects from preschool between Head Start and non-Head Start children on math and reading achievement. Moreover, if teachers in the K12 system alter the alignment of their instruction toward an incoming higher-ability student (in light of a lower share of low-achieving students due to Head Start spending), the quality of K12 instruction could be affected for all students. Importantly, these spillover effects need not occur in the same classroom because lower shares of students requiring remediation or special services may allow schools to allocate resources toward more productive inputs, which may benefit all students in the school. [7]

Through the hypothesized direct "alignment channel," both the potential direct effects of Head Start spending and dynamic complementarity will only be experienced by Head Start participants. However, through the indirect "spillover effects" channel, *all* children in K12 schools with former Head Start participants may experience the indirect dynamic complementarity effects, in addition to the direct effects experienced by Head Start participants.

III. BACKGROUND AND OVERVIEW OF HEAD START AND SCHOOL FINANCE REFORMS

In this paper, we study the combined effect of two well-known broad-reaching publicly funded human capital interventions that were targeted largely to low-income children at different ages (Head Start and K12 public school spending). We present background on each in turn.

III.A. *Background on Head Start*

Head Start was established in 1964 as part of Lyndon B. Johnson's "War on Poverty," and

[7] One of the program components of Head Start is in teaching parenting skills; thus, another possible indirect channel is changes in parental quality. However, we argue that this potential mechanism is unlikely to be the dominant channel at work in this study because existing evidence provides much stronger support for the primary mechanisms operating through (a) enhancing literacy, numeracy, reasoning and problem solving, and decision-making skills; (b) access to pediatric care; and (c) improved nutrition (Zigler, XXXX; Currie and Neidell, 2007). We also present a test of this using within family variation. In such tests, we find no indication that siblings of those exposed to higher-levels of Head Start spending have improved outcomes (Appendix E). This run counter to the parental quality mechanism.

is a national, federally-funded, early-childhood program with the aim of improving the human capital of poor children. Head Start was designed to increase the human capital of poor children in several domains; 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.[8,9] Head Start also provides first-time parents with parenting strategies (Zigler et al., 2011). Head Start currently operates more than 19,200 centers and serves more than 900,000 children. [10] Current Head Start expenditures average about $8,500 per child (in 2014 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). However, per pupil Head Start spending levels are on the same order of magnitude as the average public K12 per-pupil spending which is about $10,700 (in 2014 dollars).[11]

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.[12] To open a new Head Start center, local organizations (typically non-profit 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.[13] Each grantee must comply with student-to-teacher ratio guidelines and other standards outlined in the

[8] http://www.acf.hhs.gov/programs/ohs/about/head-start

[9] An OEO report of 1967 documents Head Start accomplishments in the first two years on child health that include 98,000 eye defects treated; 900,000 cases of dental problems addressed (5 cavities per child); 740,000 without polio vaccinations received vaccines; and 1,000,000 were given measles vaccinations.

[10] See appendix **Table A1** for the national, annual enrollment in Head Start between 1965 and 2013.

[11] There is considerable variability around this national average in individual states. States spending the least per pupil included Utah ($6,555), Idaho ($6,791), Arizona ($7,208), Oklahoma ($7,672) and Mississippi ($8,130).

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

[13] As documented in Ludwig and Miller (2007), the poorest 300 counties initially received grant assistance to apply for funding at the program's inception, so that funding was targeted to communities with the highest poverty concentrations.

Head Start Act. At inception(?), the average student-to-teacher ratio in a Head Start classroom was roughly 17 (Zigler, 2010).[14] During the first 15 years of the program, the majority of Head Start children were enrolled in part-day centers (as opposed to full-day programs which are 6 or more hours per day such as Abecedarian), and often part-year (GAO report, 1981).[15]

To be eligible for Head Start participation children had to be four years old. At each center, at least 90% of enrolled children had to be from families whose income is below the federal poverty line, and at least 10% of children had to have a disability.[16] **Figure 1** plots the national Head Start enrollments as a percentage of the number of income eligible four-year-olds between 1960 and 1994. **Figure 1** also plots the share of 3 and 4 year olds who are reported as being enrolled in full-time daycare over time (in the Current Population Survey?). This figure reveals key patterns that help put our empirical work in perspective. First, the ratio of children enrolled in Head Start to the number of poor four-year-olds was as high as 90 percent in the very early years and then stabilized around 60 percent (current levels are estimated at 55 percent). This ratio is important to keep in mind as we interpret the magnitudes of our intent-to-treat estimates (presented in Section VI). It is also important to note that 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. This illustrates that 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 enrollment between 1966 and 1970. 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).

[14] This student to teacher ratio is higher than the prevailing student-to teacher ratios in the model preschool programs of the Perry Preschool (5.7 children per teacher), the Abecedarian Project (6 children per teacher), and Chicago Child Parent Center and Expansion Program (8-12 children per teacher) (Cunha, Heckman, Lochner and Masterov, 2006; Fuerst and Fuerst, 1993; Carneiro and Ginja, 2014). Note also the much smaller scale of these model programs as the Perry and Abecedarian programs each served just over 100 disadvantaged children.

[15] We are unable to identify which of these options a local Head Start center offered children who attended (part-day vs full-day; part-year vs. full-year). Summer-only programs were phased out by 1981 (Gibbs et al., 2012).

[16] Children who are 4 years old and live in poverty (i.e., family income below the federal poverty guidelines, or family is on public assistance programs AFDC or SSI) are eligible to be enrolled in the program; and beginning in 1972 (as part of the Economic Opportunity Act Amendment) at least 10 percent of children per center must have a disability (without any income cap on the eligibility of these children). In 1969, a provision was added allowing children from families above the poverty level to receive Head Start services for a fee. A fee schedule for non-poor participants in Head Start was required; fees were prohibited for families below the poverty line. The eligibility criteria was mostly unchanged during the period of the program we analyze (Source: 45 CFR (Code Federal Regulations), Parts 1301 to 1311, Early Childhood Learning and Knowledge Center: http://eclkc.ohs.acf.hhs.gov/hslc; www.eric.ed.gov; Zigler and Valentine, 1979).

Figure 1 also plots the share of 3 and 4-year-olds who are reported as being enrolled in full-time daycare over time. This figure highlights that the rollout of Head Start coincides with a period in which most children were not in formal full-time pre-school, and also coincides with a general increase in the proportion of children aged 3 to 4 enrolled in full-time pre-school. In the context of the estimated effects of Head Start during this rollout period, the counterfactual option in the early years is primarily home care, as opposed to some other full-time pre-K program (as might be the case with present-day public pre-K expansions).[17]

Because Head Start programs vary in quality, size, and scope, 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 the enrollment margin.[18] As such, because the target eligible population for Head Start is poor four-year-olds, our measure of Head Start spending is federal Head Start spending per *poor* four-year-old in the county. By the early 1970s, the average county with a Head Start center spent about $2,000 per poor four-year-old and about $5,300 (2015 dollars) per enrollee. By comparison, the average total K12 per-pupil spending in the mid-1970s was quite similar at about $5,900 (in 2015 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.[19] The geographic variation in the timing of the roll-out of Head Start is central to our empirical strategy to isolate exogenous variation in Head Start spending across birth cohorts within a county.

[17] The expansion of Head Start involved both increases in the number of enrolled children and increases in the spending levels for enrolled children. **Figure A2** shows the evolution of national Head Start spending per enrollee during the early period. Spending per enrollee (Black line) increased gradually during the first five years of the program (1965-1971) and stabilized at around $5,600 per enrollee thereafter.

[18] In Figure 2, changes in this county-level measure track the national trajectory. However, Head Start spending per enrollee increases do not capture increases in the total number of children affected by Head Start, so that spending per four-year-old in the county is a more appropriate measure. To illustrate this point, we collected data on Head Start spending per enrollee and Head Start spending per poor income eligible 4-year old at the state level between 2003 and 2014 (years for which both sets of data are available). Using within-state changes in spending over time, on average a 10 percent increase in spending per income eligible four-year-old is associated with only a 0.243 percent increase in spending per enrollee. However, it is also associated with a 1.2 percent increase in the number of Head Start participants, and a 6.5 percentage-point increases in the percentage of income eligible poor four-year olds enrolled (**Table A1**). While spending per income eligible 4-year old is sensitive to both increases in funding per enrollee and increases in total enrollment, increases in spending per enrollee are unrelated to increases in enrollment. These patterns make clear that for studying the rollout of Head Start, spending per enrollee would be an inappropriate measure, and that spending per income eligible 4-year old is a much better measure.

[19] Appendix **Figure A2** presents a map of each county in the United States color-coded by the year of its first Head Start center.

III.B. *Background on School Finance Reforms*

The other major human capital interventions we study are the increases in public K12 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.[20] 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.

As pointed out in Hoxby (2001), the effect of a SFR on school spending depends on (a) the type of school funding formula introduced by the reform and, (b) how the funding formula introduced interacts with the specific characteristics of a district. To capture some of this complexity, we follow Jackson, Johnson, and Persico (2016) and categorize reforms into four types. Foundation plans guarantee a base level of per-pupil spending and are designed to increase per-pupil spending for the lowest-spending districts. Spending limit plans prohibit per-pupil spending levels above some predetermined amount. Such plans tend to reduce spending for high spending districts and may reduce long-run spending for all districts. Reward-for-effort plans match locally-raised funds for education with additional state funds (often with higher match rates for lower-income areas). Such plans tend to increase spending for all districts with larger increases in low-income districts. Equalization plans typically tax all districts and redistribute funds toward lower-wealth and lower-income districts. These reform/formula types are not mutually exclusive.

[20] The first of these successful cases is the California case, *Serrano v. Priest*, decided in 1971. Challenges to state school finance systems were argued on either equity or adequacy grounds. The early challenges (1971- mid 1980s) were won on equity grounds. For "equity cases," local financing was found to violate the responsibility of the state to provide a quality education to all children. "Equity cases" sought to weaken the relationship between the quality of educational services and the fiscal capacity of the district. The more recent challenges (late 1980s onwards) were mounted on adequacy grounds. "Adequacy cases" rely on the fact that most states have a constitutional provision requiring the state to provide some minimum "adequate" level of quality schools for all children (Lindseth, 2004) and were argued on the grounds that low per-pupil spending levels in certain districts meant that the state had failed to meet this obligation. Between 1970 and 1990, 10 and 4 states had court-ordered reforms argued on equity and adequacy grounds, respectively.

To illustrate how the introduction of different formula types affected districts by pre-reform income levels and pre-reform spending levels, **Figures 2 and 3** present event-study plots of the natural log of per-pupil spending at the district level (after removing both district and year fixed effects). Year 0 is the first year of the first court order in the state, year "-5" is five years before the first court order, and year "5" is five years after the initial court order. For each court order, we link all formula changes that occurred within three years to that court-ordered SFR. **Figure 2** shows the evolution of per-pupil spending for districts in the bottom and top quartiles of per-pupil spending in 1972 (the year preceding the first court-ordered SFR) after court orders that led to the implementation of different kinds funding formula plans. **Figure 3** presents similar plots for districts in the top and bottom quartiles of the state income distribution in 1963. **Figures 2 and 3** show that court-ordered SFRs that lead to the implementation of different funding formulas had different effects on districts by pre-reform income and spending levels. That is, reforms that lead to "reward-for-effort" formulas tended to increase per-pupil K12 spending in all districts; spending limits had the most 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 K12 school spending will increase in each district as a function of the reform type introduced at the state level and the pre-reform characteristics of the district. Because these relationships are unrelated to the 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 K12 spending increases and students' longer-run outcomes.

IV. DATA

We compiled data on annual Head Start spending at the county level, and public K12 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 (ICPSR) and Surveillance, Epidemiology, and End Results (SEER) population data. These are combined to form a county-level panel of Head Start spending per poor 4-year old in the county between 1965 and 1980. Public K12 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 linked to a database of SFRs from Jackson Johnson and Persico (2016).[21] To avoid confounding nominal with real changes in spending, we convert both Head Start and K12 school spending across all years to 2000 dollars using the Consumer Price Index (CPI).

Our data on long-run outcomes come from Panel Study of Income Dynamics (PSID) sample members born between 1950 and 1976 and were followed into adulthood through 2013. These PSID cohorts straddle both the roll-out of Head Start programs across the country and the implementation of the early waves of court-ordered SFRs.[22] We include all information on PSID individuals between 1968 and 2013.[23] We linked persons in the PSID using their census blocks during childhood to school spending data, SFR data, and Head Start spending data.[24] 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 outline the algorithm in **Appendix A**. 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, roll-out of other "War on Poverty" initiatives and expansion of safety net programs—described in Section V) from multiple data sources.[25]

[21] The Census of Governments has been conducted every five years since 1972 and records school spending for every school district in the US. The Historical Database on Individual Government Finances (INDFIN) contains district finance data annually for a sub-sample of districts from 1967, and 1970 through 1991. After 1991, the CCD School District Finance Survey (F-33) includes data on school spending for every school district in the US. Additional details on how these databases were compiled and the coverage of districts in these data are in Appendix B.

[22] The share of individuals potentially exposed to Head Start expenditures at age 4 increases significantly with birth year over the 1950-1976 birth cohorts analyzed in the PSID sample. Two-thirds of the sample grew up in a state that was subject to a court-mandated SFR between 1971 and 2000 (the first court order was in 1971).

[23] The PSID maintains high wave-to-wave response rates of 95-98%. Studies have concluded that the PSID sample of heads and wives remains representative of the national sample of adults (Gottschalk et al, 1999; Becketti et al, 1997).

[24] The PSID began interviewing a national probability sample of families in 1968. These families were re-interviewed each year through 1997, when interviewing became biennial. All persons in PSID families in 1968 have the PSID "gene," which means that they are followed in subsequent waves. When children with the "gene" become adults and leave their parents' homes, they become their own PSID "family unit" and are interviewed in each wave. The original geographic cluster design of the PSID enables comparisons in adulthood of childhood neighbors who have been followed over the life course. Moreover, the genealogical design implies that the PSID sample today includes numerous adult sibling groupings who have been members of PSID-interviewed families for more than four decades. 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.

[25] The data we use include measures from 1968-1988 Office of Civil Rights (OCR) data; 1960, 1970, 1980, and 1990 Census data; 1962-1999 Census of Governments (COG) data; Common Core data (CCD) compiled by the National Center for Education Statistics; Regional Economic Information System (REIS) data; the comprehensive case inventory of court litigation regarding school desegregation over the 1955-1990 period (American Communities Project), and major plan implementation dates in large districts (compiled by Welch/Light); and American Hospital Association's Annual Survey of Hospitals (1946-1990) and the Centers for Medicare Provider of Service data files (dating back to 1960s) to identify the precise date in which a Medicare-certified hospital was established in each county of the US (an accurate marker for hospital desegregation compliance).

We define low-income children as those whose average parental income (between ages 12 through 17) fell in the bottom quartile.[26] Among cohorts born between 1963-1976 for whom parental income at age four is observed, roughly 80 percent of those that we classify as low-income were below the official federal poverty line at age four, and 93 percent of those who were below the poverty threshold at age four are classified as low-income by our definition. The analytic sample includes 13,381 individuals from 4,684 childhood families, 1,431 school districts, 1,070 counties, across all 50 states. About 87 percent of the sample is white, and 44 percent is female. 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 1) educational outcomes—whether graduated from high school, years of completed education; 2) labor market and economic status outcomes (in real 2000 dollars)— log wages, family income, annual incidence of poverty in adulthood[27] (ages 20-50); and 3) 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.

V. EMPIRICAL STRATEGY

We aim to uncover the causal effects of spending on Head Start programs, public school K12 spending, and the effects of the interaction between the two. To this aim, we exploit policy-induced changes in Head Start and public K12 education spending that are unrelated to child family- and neighborhood-level determinants of adult outcomes. Due to the complexities of the causal effect of each kind of education spending (early childhood versus K12), we describe each source of variation in turn, and then later discuss how we combine the two in our empirical models.

V.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). Our research design takes advantage of the staggered introduction across geographic areas of Head Start programs and the resulting spending

[26] 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.

[27] based on the family income-to-needs ratio and federal poverty thresholds by family structure and household size.

increases during the program's rollout. Before the rollout of Head Start to an area there is no spending on Head Start. However, after the introduction of Head Start in a county, spending levels typically increase for several successive years. **Figure 4** shows an event-study plot of Head Start spending per poor-four-year-old before and after roll-out 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.

In the high-spending counties, once the first center is established, spending per poor four-year-old increases rapidly. As expected, the increase is much larger in the high-spending counties (from zero to about 5000 per poor 4-year old) than the low-spending counties. However, spending is highest during the first year and then falls after that. This initial increase and subsequent fall is an artifact of the large national enrolment in Summer-only programs that were phased out in the following years (Figure 1). The initial increase in Head Start spending due to the Summer-only programs is also evident in counties with low Head Start spending in 1980. In essence, almost all counties experienced a transitory increase in Head Start spending due to the ubiquitous introduction of Summer-only programs that falls over time. However, high-spending counties expanded enrollment (and spending) in full-year programs that is sustained over time, while the low spending counties did not increase spending on full year program and reverted to near zero head start spending within four years.

The left panel of **Figure 4** reveals that among children born within a few years of each other in the same county, some were four years old when there was no Head Start spending in their county, and others would have been four years old at the end of the phase-in stage when spending levels were high. If higher levels of Head Start spending improve outcomes, one should observe that (a) the post-rollout cohorts should have better outcomes than the pre-rollout cohorts, and (b) improvements between pre and post rollout cohorts should be larger in those counties with larger sustained increases in Head Start spending. The right panel of Figure 6 reveals exactly this pattern for years of educational attainment (measured in adulthood). The event study shows that areas that had small increases (dashed grey line) and those with large increases in Head Start spending (solid black line) were largely on the same trajectory among cohorts who were older than four years old when the first Head Start center was established (i.e. years -8 through year 0). However, the post-rollout cohorts have much better outcomes in high Head Start spending counties than in the low-spending counties. This figure provides a graphical representation of our empirical strategy.

Our preferred difference-in-difference (DiD) strategy uses all 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 4, to the differences in outcomes across the same birth cohorts within other childhood counties that experienced small (or no) increases in Head Start spending at age 4. These DiD type comparisons are implemented in a regression setting by estimating [4] by Ordinary Least Squares (OLS).

[4] $$Y_{icb} = \beta \cdot HS_{cb}^{age\,4} + \gamma \cdot C_{icb} + \theta_c + \tau_b + \varepsilon_{icb}.$$

In [4], Y_{icb} is the outcome of individual *i*, from childhood county *c*, in birth cohort *b*. The variable of interest ($HS_{cb}^{age\,4}$) is Head Start spending per poor four-year old in county *c*, when birth cohort *b* was age 4. To rely only on within-county variation in Head Start spending across cohorts, [4] includes childhood county fixed effects (θ_c), and to account for cohrt 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 below. The idiosyncratic error terms is ε_{icb}.

There are two identifying assumptions. First, counties that experienced increased Head Start spending over time (where most of the variation occurs at rollout) 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 student outcomes. Figure 5 suggests that the first condition is satisfied. However, in Section VI.B we present several pieces of evidence that these identifying assumptions are valid.

V.B. *Identifying the effects of K12 School Spending*

Our measure of K12 public school spending during childhood, ppe_{idb}^{5-17}, is the *natural log* of average public K12 school spending per pupil (in real 2000 dollars) during school-age years (ages 5 through 17) in an individual's childhood school district.[28] We refer to this as K12 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

[28] The average level of district per-pupil spending across all school-age years provides a summary measure of the level of financial resources available in the individual's childhood school district during all their school-going years (ages 5 through 17 corresponding to expected grades K12). We use the natural log of this average measure to capture the fact that school spending likely exhibits diminishing marginal product.

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 did. Accordingly, we 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 K12 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 Jackson, Johnson, and Persico (2016), we quantify the relationship between K12 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 [5] by 2SLS. All common variables are defined as in [1].

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

To rely only on variation across birth cohorts within districts, we include school district fixed effects (θ_d), and to account for time effects we include birth-cohort fixed effects (τ_b). 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, $SFRExp_{icb}$, 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 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 ($\widehat{dose}_d$).[29]

[29] More formally, the first stage regression is as in [a] below

It is important that we use an exogenous predictor of $\widehat{dose}_d$ that is unrelated to the potentially endogenous decisions made by districts after reforms. To form such a prediction, we use the full universe of school districts and regress per-pupil spending on (a) indicators for years of SFR exposure, interacted with reform type, interacted with pre-reform spending levels in 1972; and (b) 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. We then take the fitted values from this regression to obtain a predicted reform-induced spending change for each district (based on reform type implemented by the state, district spending prior to reforms, and district income levels prior to reforms). See **Appendix C** for details. As such, $\widehat{dose}_d$ is a weighted average of reform type, pre-reform district income levels, pre-preform district spending levels and their interactions. By construction, $\widehat{dose}_d$ is unrelated to potentially endogenous decisions made by districts after reforms. Because $\widehat{dose}_d$ is estimated using *all* school districts while we estimate effects using districts represented in the PSID sample, our approach is a two-sample-2SLS.[30]

To show that $\widehat{dose}_d$ captures meaningful variation in K12 spending caused by court-mandated SFRs, **Figure 5** shows the evolution of K12 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. $\widehat{dose}_d\leq 0$).[31] 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*. The "-5" cohort are individuals who were 22 years old at the passage of a court-ordered SFR, the "-1" cohort was 18 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 K12 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 K12 spending due to the passage of a court-ordered SFR. Because the outcome is in logs, the values

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

[30] This approach was popularized by Angrist and Krueger (1992) and has been used in several other settings (e.g., Bjorklund and Jantti, 1997; Currie and Yelowitz, 2000; Dee and Evans, 2003; Borjas, 2004).

[31] Roughly two-thirds of districts in reform states are predicted to experience spending increases due to court-ordered SFRs. Districts predicted to increase spending were predicted to increase by 10 percent due to the reforms, on average. Districts predicted to decrease spending were predicted to decrease by 8 percent due to the reforms, on average. As shown in Figure II, the relationship between predicted increases and actual increases is monotonic, but non-linear. This motivates our flexible parametrization of predicted spending increases.

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.

Consistent with the timing of court-ordered SFRs being exogenous to underlying trends in school spending, both districts with lower- and higher predicted dosage were on similar pre-reform trajectories as similar districts in non-reform states. Consistent with $\widehat{dose}_d$ isolating real variation in dosage, cohorts that turned 5 years old during the year of the initial court order (cohort 12) in districts with $\widehat{dose}_d > 0$ experience a 19 percent increase in school-age per-pupil spending, while the same cohorts in districts with $\widehat{dose}_d \leq 0$ experience a 5 percent increase. 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) likely isolates exogenous variation in school spending. We present additional tests to support the validity of this approach in Section VI.B.

If our identification strategy is valid and K12 spending affects outcomes, outcome differences across exposed and unexposed cohorts should follow similar patterns to those of K12 spending. The right panel of Figure 5 shows this for years of educational attainment. Areas that had small (gray line) and large (black line) reform-induced increases in K12 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 K12 spending increase districts than in the low predicted K12 spending increase districts. This figure depicts graphically the variation that undergirds our identification strategy.

V.C. *Testing for Dynamic Complementarity*

To test whether the marginal effect of increased Head Start spending varies by the level of K12 spending and vice versa, we estimate the effects of public pre-K and K12 spending on adult outcomes with the inclusion of the interaction between Head Start spending at age 4 ($HS_{icb}^{age\ 4}$) and the natural log of public K12 spending between the ages of 5 and 17 (ppe_{idb}^{5-17}). We define $\text{INT}_{idb} = \left(HS_{icb}^{age\ 4} \times ppe_{idb}^{5-17}\right)$. We use the DiD variation (i.e. the within county variation) in Head Start spending ($HS_{icb}^{age\ 4}$), and instrument for both K12 spending (ppe_{idb}^{5-17}) and the interaction between Head Start and K12 spending (INT_{idb}). Our excluded instruments are all the two-way and three-way interactions between (a) the number of years of exposure to a court-ordered SFR, (b) predicted dosage, and (c) Head Start spending per four-year-old when the individual was

age 4.[32] Because a school district is a smaller unit of observation than a county, all models include district fixed effects (which subsumes the childhood county fixed effects). The resulting model is as in [6], where $\ln\left(\widehat{PPE}_{5-17}\right)_{idb}$ and $\widehat{INT}_{idb}$ are fitted values from first-stage regressions.

[6] $$Y_{icb} = \beta_{HS} \cdot HS_{cb}^{age\ 4} + \beta_{k12} \cdot \widehat{ppe}_{idb}^{5-17} + \beta_{int} \cdot (\widehat{INT}_{idb}) + \gamma \cdot C_{icb} + \theta_d + \tau_{bdi} + \varepsilon_{idb}.$$

The interaction effect between pre-K and K12 spending can be identified in [6] because (a) among counties that faced similar increases in Head Start spending, some were located in school districts that experienced larger (or smaller) increases in K12 spending due to the passage of a court-ordered reform, and (b) among cohorts from districts that faced similar increases in K12 spending due to the passage of a court-ordered reform, some were located in counties that had higher (or lower) levels of Head Start spending when those cohorts were age 4.

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, occupational status and income, mother's marital status at birth, birth weight, child health insurance coverage, and gender. 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 county-by-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, 2013; Chay, Guryan, & Mazumder, 2009; Hoynes, Schanzenbach, and Almond, 2012). Standard errors are clustered at the state level.

VI. Results

We present results from specification [6] that exploit all the within-county across-cohort variation in Head Start spending and instruments for K12 public school spending at the district level using the SFR instruments. To facilitate interpretation of the base effects of K12 spending

[32] Specifically, $\widehat{ppe}_{idb}^{5-17} = \pi_{11}(SFRExp_{idb} \times dose_c) + \pi_{12}(SFRExp_{idb}) + \pi_{13}(SFRExp_{idb} \times dose_c) \cdot HS_{cb}^{age\ 4} + \pi_{14}(SFRExp_{idb}) \cdot HS_{cb}^{age\ 4} + \pi_{14}(dose_c) \cdot HS_{cb}^{age\ 4} + C_{idb} + \theta_{d1} + \tau_{b1}$.
and $\widehat{INT}_{idb} = \pi_{21}(SFRExp_{idb} \times dose_c) + \pi_{22}(SFRExp_{idb}) + \pi_{23}(SFRExp_{idb} \times dose_c) \cdot HS_{cb}^{age\ 4} + \pi_{24}(SFRExp_{idb}) \cdot HS_{cb}^{age\ 4} + \pi_{24}(dose_c) \cdot HS_{cb}^{age\ 4} + C_{idb} + \theta_{d2} + \tau_{b2}$.

and Head Start spending when the interaction between the two is included, both K12 spending and Head Start spending are centered around their respective means. Thus, the coefficient on Head Start is the marginal effect of Head Start spending at the average level of K12 spending, and the coefficient on K12 spending is the marginal effect of K12 spending at the average level of Head Start spending. To organize our discussion, we first discuss the base effects of K12 spending (in logs) and Head Start spending, present empirical evidence that these base effects are valid, and then discuss the estimated interaction effects. We present our estimated effects on education outcomes, followed by adult economic outcomes, and finally incarceration.

VI.A. *Estimating the Base Effects of Head Start and K12 Spending*

Table 2 presents the estimates from [6] for low-income (bottom income quartile during childhood) children. Column 1 presents the effects on the probability of high school graduation. The coefficient on Head Start spending per poor four-year-old is 0.01984 (p-value<0.01). This indicates that increasing Head Start school per poor 4-year old in the county by $1,000 (roughly a 25 percent increase) increases the likelihood of high school graduation by 1.9 percentage points for a poor child exposed to the average level of K12 spending. Given that the average level of Head Start spending conditional on having any Head Start program in the county is about $4,000, this implies that having access to the average Head Start program would increase high school graduation of poor children by roughly 7.6 percentage points.

Increases in Head Start spending can affect outcomes through increasing Head Start participation, increases in the quality and scope of the head start services, and also indirectly through peer effects in the K12 system due to having better-prepared schoolmates. While existing studies have focused on the effect of *enrolling in* Head Start on participants, we estimate the effect on Head Start spending on *all* eligible children. Because there are multiple channels through which spending effect may emerge, we attempt to provide a sense of how our spending effects relate to the participation effects in the extant literature. Unfortunately, the PSID survey did not collect information regarding Head Start participation during the years in question.[33] However, we can approximate this margin using national data. Between 1965 and 1980, Head Start enrollment (full-time or part year) accounted roughly 66 percent of all eligible four-year-olds (see Figure 1).

[33] The survey data employed in Currie et al are retrospective data collected starting in 1995 for individuals born after 1965 (i.e. who were 4 years old after 1979). As such, Head Start enrolment data were not recorded for the ramp up period during which most of our variation in derived. See Appendix **F** for further discussion.

Because, centers can enroll 10 percent of non-poor enrollees and must have had disables children, the participation rate among income-eligible could have been as low as 60 percent. Roughly 80 percent of eligible individuals 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 individuals with a Head Start center in their local area will participate, this implies a Head Start participation rate of about 0.6/0.8=0.75 conditional on having a Head Start center in the county. We take this as our ballpark estimate of the increase in Head Start enrollment rate (among low-income children) due to the rollout of the average Head Start center in the county during our study period.

If all of our intention-to-treat effect was due to Head Start enrollment (and there were no spillover effects to other low-income children), our assumed participation margin effect implies a treatment-on-the-treated effect of 0.076/.75=0.105, or 10.5 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 only considers the effect of full-year Head Start, the average enrollment rate among eligible children was about 40 percent between 1965 and 1980. This implies a full-year Head Start participation rate of about 0.4/0.8=0.5 conditional on having a Head Start center in the county at age 4. If all of our intention-to-treat effect was due to *full-year* Head Start enrollment, this assumed full-year Head Start participation margin effect implies a treatment-on-the-treated high-school completion effect of 0.076/.5=0.152, or 15.2 percentage points. This is in line with the larger of the participation margin effects in the literature, or with the smaller participation effects in the literature if there are modest spillovers. In sum, our spending effects can be explained entirely by an enrolment effect using the range of estimates in the existing literature. However, we cannot rule out that some modest portion of our effects are driven by (a) improvements in the quality and scope of Head Start centers (full day versus half day, full time versus summer only, better teachers, etc.), and (b) spillovers from Head Start participants to low-income non-participants in the K12 schooling system.

As expected, the coefficient estimates for K12 spending are very similar to those presented in Jackson, Johnson, and Persico (2016). The coefficient on the log of K12 spending during the

[34] For example, Curie et al find that participating in Head Start increases the high school graduation rates for white by 20 percentage points, with no statistically significant effect for blacks. Deming 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) find that the average effect is 14 percentage points.

school age years is 0.5956 (p-value<0.01). That is, increasing K12 school spending (across all 12 school-age years) by 10 percent increases the likelihood of high school graduation by 5.9 percentage points for a poor child exposed to the average level of Head Start spending. Relative to baseline, this is an 8.3 percent increase. The estimates indicate that increasing Head Start spending by $3,000 would have roughly the same effect on high school graduation as increasing K12 spending by 10 percent across all school-age years.[35] Calculations indicate that on average, $1 spent on Head Start (per eligible child in the county) had about the same effect as $1.43 spent on K12 education (per enrolled student in the district).[36] Given that Head Stat spending is targeted to poor children while K12 spending is not, this is not surprising.

Column 2 presents a similar pattern for completed years of education for low-income children. Increasing Head Start spending per poor 4-year old in the county by $1,000 increases the years of educational attainment by 0.096 years for a poor child exposed to the average level of K12 spending. As such, at average Head Start spending levels of about $4000, a Head Start center is estimated to increase years of education by roughly 0.38 (just over a third of a year). Increasing school-age K12 spending by 10 percent increases the number of years of completed education by 0.237 years for a low-income child exposed to the average level of Head Start spending. For years of education attained, the relative benefits of early education spending are even larger than for high school graduation. On average, $1 spent on Head Start (per poor child in the county) had a similar effect as $1.74 spent on K12 education (per enrolled student in the district).

Results for non-poor children (top 3 income quartiles during childhood) are in Table 3. The estimated K12 spending effects on the education outcomes are positive and sizable, but not statistically significant. The point estimates indicate that increasing K12 spending in the district by 10 percent increases the likelihood of high school graduation by 3.2 percentage points, and increases years of educational attainment by about 0.2 years for a non-poor child exposed to the average level of Head Start spending. These estimated effects are similar in magnitude to the effects for poor children, but are less precise. Accordingly, while one cannot reject that the

[35] During the sample period, a 10 percent increase in K12 spending is roughly equal to increasing per-pupil K12 spending by $480 each year over 12 years (about $4300 in present value terms assuming a 7 percent interest rate).
[36] The 10 percent K12 spending increase was about $4300 in present value terms and it led to a 5.95 percentage point increase in high school graduation. This is 1.38 percentage points per thousand dollars spent on K12 education. From Table 2, the increase in High school graduation from a 1000 increase in Head Start spending is 1.98 percentage points. As such each dollar spend on Head Start (per poor four year old) has the same effect as 1.98/1.39=1.43 dollars spending on K12 spending per student.

marginal effects of K12 spending on the education outcomes are different for low-income and non-poor children, one can also not reject that the estimated effects are zero for non-poor children.

While the estimated effects of K12 spending are similar across the two groups, the estimated effects of Head Start are very different (as one might expect). In contrast to the large effects for low-income children, the coefficients indicate that increasing Head Start school per poor four-year-old in the county has very small effects on a non-poor child exposed to average levels of K12 spending. For both these outcomes, one *can* reject that the marginal effect of Head Start is the same for low-income and non-poor children, one cannot reject that the effect on the non-poor is zero, and one can reject that the effect for low-income children is zero. This suggests that (a) there no spillover effects of Head Start on non-poor children, and that (b) increases in Head Start spending are not associated with other policies that improve the outcomes of non-poor children. This bolsters the credibility of the research design. Note that no effect for non-poor children does not rule out spillover effects on low-income children who did not attend Head Start.

The economic outcomes we examine in this paper are adult wages, annual family income, and the annual incidence of poverty between the ages of 20 and 50. Our models that analyze adult economic outcomes (such as wages and annual family income) 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 3, 4, and 5 in Table 2 present these results for children from low-income families. Looking at wages, the coefficient on the log of public K12 school spending is 0.927 (p-value<0.1) and that on Head Start spending per poor 4-year-old is 0.0193 (p-value<0.01). That is, for children from low-income families exposed to average levels of Head Start spending, increasing K12 spending by 10 percent is associated with about 9.3 percent higher adult wages. Similarly, for these same children, increasing Head Start spending by $1,000 per poor 4-year-old is associated with 1.93 percent higher wages in adulthood. At the average level of Head Start spending following the program's rollout, this implies an average Head Start rollout effect of 7.75 percent higher wages due to the expansion of public pre-K availability for poor children. Using the same calculations as for the education outcomes, for low-income children, $1 spent on Head Start (per eligible child in the county) had about the same effect as $0.90 spent on K12 education (per enrolled student in the district). This suggests that for the economic outcomes, *among low-income children*, spending on per pupil K12 education (on the margin) had a similarly sized effect as spending on Head Start per eligible 4-year-old (on the margin).

Column 3 of Table 3 presents the effects on adult wages for non-poor children. Similar to the education outcomes, there are positive effects of K12 spending, but no effect of Head Start spending on the wages of adults who were from non-poor families. The coefficient on the log of K12 public school spending is 1.173 (p-value<0.05) and that on Head Start spending per poor 4-year-old is 0.00617 (p-value>0.1). That is, for children from non-poor families exposed to average levels of Head Start spending, increasing K12 spending by 10 percent is associated with 11.7 percent higher earnings between the ages of 20 and 50. However, consistent with no Head Start spillover effects on non-poor children, increasing Head Start spending is associated with no difference in earnings between the ages of 20 and 50.

The pattern of estimates for family income and poverty in columns 4 and 5 of Tables 2 and 3 mirror those for adult wages. Head start spending is associated with large and statistically significant improvement in the economic outcomes of children from low-income backgrounds (Table 2) and has small and statistically insignificant effects on the outcomes of children who did not grow up in poor households (Table 3). However, increases in public K12 spending are associated with sizable improvement in the economic outcomes of all children. The one exception to this pattern is adult poverty for which neither Head Start spending nor K12 spending has an effect among non-poor children (likely because baseline rates are low for this population). Indeed, the incidence of poverty at age 30 is only 5 percent for children from non-poor families compared to 18 percent for children from low-income households.

The last outcome we examine is the probability that an individual has ever been incarcerated (Column 6 of tables 2 and 3). As with poverty, we find no effect of either kind of spending on the adult incarceration rate for non-poor children. We, therefore, focus our discussion on the outcomes of children from low-income families. For these children, a $1,000 increase in Head Start spending per poor four-year-old reduces the likelihood of being incarcerated by 0.7 percentage points (p-value<0.05). At the average levels of Head Start, this implies an average Head Start rollout effect of about three percentage points. If one were to ascribe all of this effect to the participation margin for full-year Head Start, it would imply a participation effect of 6 percentage points. Effects of this magnitude are in line with the results from Garces, et al. (2000). Column 6 also shows that increasing K12 per-pupil spending by 10 percent (at average Head Start spending levels) reduces the likelihood of adult incereration by 12 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 Moreti 2003), 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 K12 spending and reduced crime.

In sum, increases in Head Start spending improve the adulthood outcomes for children from low-income families and have no effect on the outcomes of non-poor children. In contrast, increases in K12 spending improve outcomes both groups. These patterns lend credibility to the research design because K12 spending increases are experienced by all children, while the direct effects of Head Start spending are experienced by low-income children. Moreover, because students tend to be segregated by parental income, any indirect spillover effects of Head Start on non-Head Start enrollees will likely be experienced by low-income children.

VI.B. *Establishing the validity of the base effects*

It is important for us to establish that the variation we use in both Head Start spending and K-12 spending is exogenous. Here we summarize empirical tests that support the validity of each source of variation (for a detailed discussion of each test, see Appendix G).

Spending Effects by Child Age: *No confounding policies.*

One may worry that the timing of Head Start rollout or the timing of SFRs coincided with other policies that also improve adult outcomes. One test of this would be to determine whether the effects of the spending increases occur only among those who were of the appropriate age. If counties or districts adopted other policies to improve outcomes for low-income children (that were not targeted to the exact same age range as that in question) one would observe improvements for other age ranges also. To test this for Head Start, we estimated the marginal effect of the level of Head Start spending that prevailed when individuals were different ages, conditional on the level of Head Start spending when they were four-years-old. The marginal effects of Head Start spending on years of education and wages by age (conditional on spending at age 4), are presented in Figure 6. Higher level of Head Start spending when the individual was four years old are associated with improved adult outcomes while the spending levels at ineligible ages (age 1 through 3 or 5 through 10) are not. This is consistent with our hypothesized mechanisms.

We conduct a similar test for K12 spending. If the spending increases we exploit operate through improved K12 education, one should see improvements for those who were between the ages of 5 and 17 when there was a SFR, but no effect for individuals from the same districts who

were 18 or older at the time. To test this, we instrument for the K12 spending levels that prevailed in an individual's childhood school district when they were between the ages of 18 to 22, and we find no effect on adult outcomes. Both tests show positive effects of spending levels that prevailed when individuals were of the appropriate age, and no effect of the spending levels that prevailed when individuals were either too young or too old to be affected through the hypothesized mechanisms. This suggests, the effects are not driven by confounding policies, the effects emerge through the hypothesized channels, and areas that saw spending increase were not on a pre-existing trajectory of improving outcomes. These tests support a causal interpretation of the patterns.

Instrumenting for Head Start Spending

Because we do not instrument for Head Start spending, as we do for K12 spending, one may still worry that the changes in Head Start spending could be related to other changes that influence outcomes. To address this, we instrument for Head Start spending with an indicator denoting whether any Head Start center had been established in one's county of birth by the year the child was age 4. In such instrumental variables models that rely only on the *timing* of the establishment of the first Head Start center in a county, the effects or Head Start spending are similar to those in Table 2. This also supports a causal interpretation of the patterns in Table 2.

Accounting for Unobserved Family Characteristics or Selection

Another concern one may have with the estimates is that due to selective migration, the characteristics of the individuals exposed to different levels of K12 spending or Head Start spending are not the same. To rule out the possibility that our results are driven by differences across treated and untreated families, 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] to include sibling fixed effects. In such models, the effects are similar to those in Table 2. This suggests that family selection cannot explain the main pattern of results.

Addressing Bias due to Endogenous Mobility

This sibling tests outline above also address concerns regarding endogenous mobility driving the results because individuals in the same family have the same residential address. However, as an additional check on endogenous mobility, we re-estimated all models limiting the analysis sample to those who lived at their (earliest) childhood residence prior to the enactment of

Head Start programs in their respective county or SFR in their district. We find similar results to Table 2, so that that endogenous residential mobility is not a source of bias in our analysis.

VI.C. *Evidence of Dynamic Complementarity Effects*

Before presenting the magnitudes of any dynamic complementarity effects, we first establish whether such effects exist. Specifically, in estimation of [6] we test for whether the coefficient on the interaction is positive and statistically significantly different from zero. In Table 2, for kids from low-income families, the coefficient on the interaction is positive and statistically significant at the one percent level for high school graduation, years of education, and adult wage, and positive and significant at the 5 percent level for family income. For the two adverse outcomes, adult poverty and ever being increased, the coefficients on the interaction terms are negative and statistically significant at the 10 percent level. That is across all outcomes for low-income children, increases in Head Start spending raise the marginal effect of K12 spending and vice versa. 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 close to statistically significant, and the sign of the coefficients across outcomes do not go in the same direction. That is, Head Start spending had no direct or indirect effect on the outcomes of children who were from non-poor families.

Because the coefficients of interaction terms can be difficult to interpret directly, we also present the interaction effects graphically. Using the regression estimates, we compute the marginal effect of increasing Head Start spending per poor four-year-old by $1000 at each percentile of the K12 spending distribution (between the 20th and 90th percentile).[37] We also present the 90 percent confidence interval for each estimated marginal effect. Similarly, we compute the marginal effect of increasing K12 spending by 10 percent at each percentile of the Head Start spending distribution (between the 0thth and 90th percentile). The marginal effects are presented for both poor and non-poor children. Because the interactions are not significant for non-poor children, confidence intervals are provided for low-income children only.

The left panel of Figure 7 presents the estimated marginal effects of Head Start spending by the percentile of K12 spending on the likelihood of graduating high school. If dynamic complementarity exists between early childhood and K12 spending, the plots will be upward sloping. As expected, for non-poor children (dashed line), the marginal effect of Head Start

[37] We do not include estimated marginal Head Start effects below the 20th percentile or above the 90th percentile of the K12 distribution because this is likely outside the range at which the marginal effects are estimated.

spending is flat and indistinguishable from zero at all levels of K12 spending. In contrast, for children from low-income families (solid line), there is a clear positive gradient. Head Start spending has small statistically insignificant effects on children in public school districts below the 30th percentile of per-pupil K12 spending level. However, at the 90th percentile of K12 spending, increasing Head Start spending per poor four-year-old by $1000 increases the likelihood of graduating high school by about 6.5 percentage points (*p*-value<0.01). The marginal effect of the same Head Start spending increase on the high school graduation rates of low-income children is roughly twice as large in districts at the 90th percentile of the K12 spending distribution than those at the median. The right panel presents the marginal effects of increases in K12 spending at different points in the Head Start spending distribution. As expected, the marginal effect of K12 spending increases on the high school graduation rates of non-poor children is positive and is unrelated to the level of Head Start spending. However, the marginal effect of K12 spending increases on the high school graduation rates of low-income children is larger among those from childhood counties with higher levels of Head Start funding. The K12 spending effects indicate that for a low-income child, a 10 percent increase in K12 spending increase high school graduation rates by about two percentage points at the 5th percentile of the Head Start spending distribution (*p*-value>0.1) and as much as 12 percentage points at the 90th percentile (*p*-value<0.05). The marginal effect of the same increase in K12 spending is almost twice as large when Head Start spending is at the 90th percentile than at the median.

Figure 8 presents very similar patterns for years of completed education. Because the coefficient on the interaction term in predicting the years of education of non-poor children has a *p*-value greater than 0.5 (and is the opposite sign as that for high school graduation), we focus the discussion on low-income children. For low-income children, increasing Head Start spending per poor four-year-old by $1000 increases the years of education completed by about 0.03 at the 20th percentile of K12 spending, by 0.12 years at the median of the K12 spending distribution (*p*-value<0.01), and by 0.31 at the 90th percentile of the K12 spending distribution (*p*-value<0.01). Looking at the marginal effect of K12 spending, increasing K12 spending by 10 percent increases years of education completed by about 0.1 at the 10th percentile of the Head Start spending distribution (*p*-value>0.5), by 0.24 years at the median (*p*-value<0.1), and by 0.52 at the 90th percentile (*p*-value<0.05). In sum, these patterns suggest important dynamic complementarity between early childhood education spending and public K12 spending for the educational

outcomes of low-income children. The pattern of results indicates that in areas with no Head Start center, increases in K12 spending may have increased educational attainment gaps. However, in areas with well-funded Head Start programs, increases in K12 spending both increased outcomes for all students and simultaneously reduced educational attainment gaps.

Commensurate with the education outcomes, there is evidence of complementarity between Head Start spending and public K12 spending in the production of adult economic outcomes for children from low-income families. Figure 9 presents the marginal effect on adult wages of each spending type by the percentile of the other spending type by childhood poverty status. For individuals from non-poor families, the marginal effect of Head Start spending is indistinguishable from zero at all levels for K12 spending levels, and a 10 percent increase in K12 spending increases the adult wage of non-poor children by about 12 percent at all levels of Head Start funding (p-value<0.05). This is consistent with K12 spending increases improving the outcomes of all children (including non-poor children) and Head Start spending being unrelated to the outcomes of non-poor children. We now focus on the magnitudes of the positive interaction effects for low-income children. Increasing Head Start spending per poor four-year-old by $1000 has no appreciable effect on the adult wage at the 20th percentile of the K12 spending distribution, but increases the adult wage by about 2.5 percent at the median (p-value<0.01), and by about 8 percent at the 90th percentile (p-value<0.01). The dynamic complementarities are sufficiently large that the marginal effect of increases in Head Start spending on the adult wage is more than three times as large when K12 spending is at the 90th percentile than at the median. We now turn to the marginal effect of increases in K12 spending. Increasing K12 spending by 10 percent increases the wage for low-income children by about 6.5 percent at the 10th percentile of Head Start spending (p-value>0.1), by 7.5 percent at the median (p-value<0.05), and by 17 percent at the 90th percentile (p-value<0.01). Similar to the education outcomes, the marginal effect of increases in K12 spending (on the adult wage for low-income children) is about twice as large when Head Start spending is at the 90th percentile than at the median. The pattern of results for adulthood family income are similar to those for adult wages and are presented Appendix H.

We also examine effects on the annual incidence of poverty in adulthood. A family is poor if their income-to-needs ratio is below the federally-determined threshold for poverty. As such, while adult poverty is *related* to family income, it is a measure of hardship. The marginal effects on the annual incidence of adult poverty are presented in Figure 10. For children from non-poor

families, neither spending increases in K12 spending nor increases in Head Start spending affect the likelihood of adulthood poverty. We therefore, focus our discussion on effects for low-income children. For children from low-income families, increasing Head Start spending per poor four-year-old by $1000 reduces the annual likelihood adulthood poverty by about 0.5 percentage points at the 20th percentile of the K12 spending distribution (p-value>0.1). Consistent with dynamic complementarity, that same increase in Head Start spending reduced adult poverty by 1.2 percentage points for those from K12 school districts at the median of the K12 spending distribution (p-value<0.05), and 2.4 percentage points at the 90th percentile (p-value<0.05). Relative to the baseline poverty rate, this marginal effect at the 90th percentile of the K12 spending distribution represents a 13 percent reduction. The differences in the marginal effect of K12 spending on adult poverty by Head Start spending are more noisily estimated (even though the interaction term is statistically significant at the 5 percent level). However, the general pattern of larger benefits to K12 spending at higher levels of Head Start spending is maintained.

The final outcome we examine is the likelihood of ever being incarcerated. The marginal effects are presented in Figure 11. As with adulthood poverty, we focus our discussion on effects for children from low-income families. As with other outcomes, there is evidence of strong dynamic complementarity between Head Start spending and public K12 spending. For children from low-income families, increasing Head Start spending by $1000 had no effect on incarceration among children in public school districts at the 20th percentile of K12 spending distribution, but reduced it by 1.3 percentage points for students in districts at the median (p-value<0.05), and reduced it by 3.4 percentage points for students in districts at the 90th percentile (p-value<0.05). Looking to the effect of K12 spending on the likelihood of being incarcerated, the marginal effects are larger at higher levels of Head Start spending (notice the difference in scale), but the K12 spending effects are rather imprecise. However, the point estimates indicate that a 10 percent increase in K12 per pupil spending reduces the likelihood of being incarcerated by 12 percentage points at the 5th percentile of the Head Start spending distribution (p-value<0.05), by 14 percentage points at the median (p-value<0.05), and 16.6 percentage points at the 90th percentile (p-value<0.05). As with the other outcomes, the reduction in crime associated improvement in early education is larger when there are greater investments in subsequent education and *vice versa*.

VI.D. *Is Parenting Quality Part of the Story?*

In Section II, we posited that the Head Start effects are driven by the programmatic

components targeted to children. However, 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 who shared the same parents, we can test for improvements in parent quality directly. Specifically, we can 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 parent 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 such models (Appendix I), older siblings are unaffected by Head Start exposure of the younger sibling. This suggests that parental quality is not part of the story and that our Head Start spending effects reflect real investments in the human capital of low-income children.

VII. Summary Discussion & Conclusions

This study provides new evidence on the life-cycle effects of Head Start spending and K12 school spending for low-income children. That is, we explore dynamic complementarities between human capital investments made in pre-school and those that subsequently occur in the K12 system. In particular, we use children's differential exposure to Head Start spending (at age 4) and court-ordered school finance reforms (between the ages 5 through 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 K12 spending increases. We present extensive tests to document that the policy-induced variation in Head Start spending and K12 public school spending we exploit is unrelated to other childhood family, community, or policy changes.

For non-poor children, SFR-induced K12 spending increases led to significant improvements in educational and economic outcomes, while increases in Head Start spending had no effect. However, for low-income children, both Head Start spending increases and SFR-induced K12 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 K12 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 born in K12 school districts that spent at the 90th percentile of the distribution than those born in K12 school districts that spent at the median. Similarly, the benefits of K12 school-spending increases on adult outcomes were larger among low-income children who were exposed higher levels of Head Start spending during their pre-school years. For low-income children, the combined benefits of growing up in counties with *both* greater Head Start spending and attending school in districts with higher K12 per-pupil spending are significantly greater than the sum of the parts (i.e., the independent effects of increases in Head Start or K12 school spending in isolation).

There are important two caveats to our work. First, because the counterfactual childcare and pediatric care may be better today than in the late 1960s and 70s, the marginal effect of Head Start may be smaller today than in the earlier period that we study.[38] 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 the profound implications for policy. However, they do suggest that the contemporary magnitude of the effects may be smaller than those we present here.

The cumulative nature of skill development as modeled by James Heckman and coauthors[39] is likely responsible for the pattern of results. Our findings highlight the importance of modeling early and later educational investments jointly and may provide an explanation for 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 socioeconomically-advantaged populations (Currie and Duncan, 1995) and the fade out of the effects of Head Start on test scores as students age (Currie and Duncan, 2000). The key policy implication of our findings is that human capital investments made in concert throughout child developmental stages (pre-school; elementary/middle school; adolescence) may yield greater returns than separate,

[38] 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 (Kline and Walters, 2015; Feller et al., 2015). The proportion of three- and four-year-olds in school has increased from roughly 10 percent in 1964 to almost 60 percent by 2000 (source: US Census Bureau, CPS October Supplement, 1964-2010; see Appendix Figure). 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.

[39] Insert a long list of the papers here!

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, but they also highlight the importance of sustained investment in the skills of disadvantaged youth.

References

Almond, Douglas and Janet Currie (2010). *Handbook of Labor Economics,* chapter "Human Capital Development Before Age Five", edited by Orley Ashenfelter and David Card.

Almond, Douglas, Hilary Hoynes, and Diane Whitmore Schanzenbach (2011). "Inside the War on Poverty: The Impact of the Food Stamp Program on Birth Outcomes". *Review of Economics and Statistics*, Vol. 93, No. 2: 387-403.

Ben-Shalom, Yonatan, Robert Moffitt, and John Karl Scholz. *An Assessment of the Effectiveness of Anti-Poverty Programs in the United States*. No. 17042. National Bureau of Economic Research, Inc, 2011.

Card, David, and Krueger, Alan B. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." *Journal of Political Economy*, February 1992, 100(1), pp. 1-40.

Card, D., and A. A. Payne. "School Finance Reform, the Distribution of School Spending, and the Distribution of Student Test Scores." *Journal of Public Economics*, 83(1) (2002): 49–82.

Cascio, Elizabeth (2009). "Do Investments in Universal Early Education Pay Off? Long-term Effects of Introducing Kindergartens into Public Schools". NBER working paper #14951.

Chay, K. Y., J. Guryan, and B. Mazumder. "Birth Cohort and the Black-White Achievement Gap: The Roles of Access and Health Soon After Birth." *NBER Working Paper No. 15078* (2009).

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schazenbach, and Danny Yang (2010). "How Does Your *Kindergarten* Classroom Affect Your Earnings? Evidence From Project STAR". *NBER working paper #16381.*

Currie, Janet and Duncan Thomas (1995). "Does Head Start make a difference?" *American Economic Review* 85(3): 341-364.

Currie, Janet and Mathew Neidell (2007). "Getting Inside the "Black Box" of Head Start Quality: What Matters and What Doesn't". *Economics of Education Review* 26(1): 83-99.

Currie, Janet and Duncan Thomas (2000). "School Quality and the Long-term Effects of Head Start." *Journal of Human Resources* 35(4): 754-774.

Cutler, David and Adriana Lleras-Muney (2008). "Education and Health: Evaluating theories and evidence". Published in *Making Americans Healthier: Social and Economics Policy as Health Policy*, Robert F. Schoeni, James S. House, George Kaplan and Harold Pollack, editors, New York: Russell Sage Foundation.

Deming, David. "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start." American Economic Journal: Applied Economics 1, no. 3 (July 1, 2009): 111–34.

Dobbie, Will, and Roland G. Fryer, 2011, "Are high-quality schools enough to increase achievement among the poor? Evidence from the Harlem Children's Zone," *American Economic Journal: Applied Economics*, Vol. 3, No. 3, July, pp. 158–187.

Fryer, Roland, and Steven Levitt (2006). "The Black-White Test Score Gap Through Third Grade." *American Law and Economic Review*, 8(2): 249-281.

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

Hanushek, Eric A. Spending on Schools. Stanford, Calif.: Hoover Press, 2001. http://hanushek.stanford.edu/sites/default/files/publications/Hanushek%202001%20PrimerAmerEduc.pdf.

_____., "School Resources and Student Performance," in Gary Burtless, ed., Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success, Brookings Institution, Washington, D.C., 1996, pp. 43-73

_____. "The Economics of Schooling: Production and Efficiency in Public Schools." Journal of Economic Literature, 1986, 1141–77.

Heckman, James J. (2008). "*Schools, Skills, and Synapses*," *Economic Inquiry*, vol. 46(3), pages `289-324.

_____. (2007). The economics, technology, and neuroscience of human capability formation. *PNAS*, 104(33):13250–13255.

Heckman, J., R. Pinto, and P. Savelyev. "Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes." *American Economic Review* (forthcoming).

Hoxby, C. M. "All School Finance Equalizations Are Not Created Equal." *The Quarterly Journal of Economics* (2001): 1189–1231.

Hoynes, Hilary, Marianne Page, and Ann Stevens (2011). "Can Targeted Transfers Improve Birth Outcomes? Evidence from the Introduction of the WIC Program", *Journal of Public Economics*, 95: 813–827.

Jackson, Kirabo, Rucker C. Johnson, Claudia Persico (2015). "The Effects of School Spending on Educational & Economic Outcomes: Evidence from School Finance Reforms". Forthcoming in the *Quarterly Journal of Economics*.

_____. The Effect of School Finance Reforms on the Distribution of Spending, Academic Achievement, and Adult Outcomes. NBER Working Paper 20118, May (2014).

Johnson, Rucker C. (2015). "Follow the Money: School Spending from Title I to Adult Earnings". Special edited volume, ESEA at 50, forthcoming in *The Russell Sage Foundation Journal of the Social Sciences.*

_____. 2015. "Can Schools Level the Intergenerational Playing Field? Lessons from Equal Educational Opportunity Policies". Forthcoming in Federal Reserve volume on mobility.

_____. "Long-run Impacts of School Desegregation & School Quality on Adult Attainments."*NBER Working Paper No. 16664* (2011) , updated August 2015.

_____. (2010). "The Health Returns of Education Policies: From Preschool to High School & Beyond." *American Economic Review Papers and Proceedings* (May), 100(2): 188-94.

Loeb, Susanna, and John Bound. "The Effect of Measured School Inputs on Academic Achievement: Evidence from the 1920s, 1930s and 1940s Birth Cohorts." *The Review of Economics and Statistics* 78, no. 4 (November 1, 1996): 653–64. doi:10.2307/2109952.

Ludwig, Jens, and Douglas L. Miller. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design." *The Quarterly Journal of Economics* 122, no. 1 (February 1, 2007): 159–208. doi:10.1162/qjec.122.1.159.

Murray, S. E., W. N. Evans, and R. M. Schwab. "Education-Finance Reform and the Distribution of Education Resources." *American Economic Review*, 88(4) (1998): 798–812.

U.S. Department of Education, National Center for Education Statistics. Public School Finance Programs of the United States and Canada: 1998–99. NCES 2001–309; Compilers Catherine C. Sielke, John Dayton, and C. Thomas Holmes, of the University of Georgia and Anne L. Jefferson of the University of Ottawa. William J. Fowler, Jr., Project Officer. Washington, DC: 2001.

Table 1:
Summary Statistics of the Analytic Dataset

	All (N=13,381)	Poor Child (N=5,623)	Non-Poor Child (N=7,758)
Adult Outcomes:			
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
Childhood school variables:			
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*	0.04	0.19	0.02
Child attended any pre-school program	0.23	0.31	0.23
School District Per-pupil spending (avg, ages 5-17)	$4,277	$4,013	$4,363
Any court-ordered school finance reform, age 5-17	0.13	0.11	0.14
Cond'l on any: # of exposure yrs to school finance reform	7.37	6.90	7.50
1960 District Poverty Rate (%)	21.52	28.25	19.35
Childhood family variables:			
Income (avg, ages 12-17)	$54,488	$22,520	$65,130
Income-to-needs ratio (avg, 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

Note: All descriptive statistics are sample weighted to produce nationally-representative estimates of means. Dollars are CPI-U deflated in real 2000 $. "Poor kid" is defined here as children whose parents were in the bottom quartile of the income distribution (approximately 80% of whom were below the poverty line). Analysis sample includes 13,381 individuals (191,613 person-year observations ages 20-50), from 4,684 childhood families, 1,431 school districts, 1,070 childhood counties and all 50 states.

*Child-specific pre-K attendance & Head Start program participation info collected retrospectively in 1995 survey IW.

Table 2:

2SLS-Difference-in-Difference Estimates of Early and K12 education Spending on Adult Outcomes

	Children from Poor Households					
	1	2	3	4	5	6
	Prob(High School Grad)	Years of Education	Ln(Wage), ages 20-50	Ln(Family Income), ages 20-50	Poverty, ages 20-50	Prob(Ever Incarcerated)
Head Start Spending$_{(age\ 4)}$/1000	0.0198***	0.0967***	0.0193***	0.0359***	-0.009***	-0.007541**
	(0.0039)	(0.0191)	(0.0050)	(0.0075)	(0.0028)	(0.003285)
Ln(K12 Per-pupil Spending)$_{(age\ 5\text{-}17)}$	0.595**	2.3733*	0.9271*	0.4476	-0.01846	-1.2751**
	(0.2370)	(1.3883)	(0.5277)	(1.0880)	(0.3331)	(0.5804)
Interaction	0.118***	0.4936***	0.150***	0.1120**	-0.03302*	-0.06045*
	(0.0455)	(0.1424)	(0.0330)	(0.0534)	(0.0187)	(0.03597)
Number of Person-year Observations			49,282	54,064	54,309	
Number of Children	5,385	5,378	5,378	5,378	5,378	4,317

Robust standard errors in parentheses (clustered at childhood state level)

*** p<0.01, ** p<0.05, * p<0.1

Data: PSID geocode Data (1968-2013), matched with childhood school and neighborhood characteristics. Analysis sample includes all PSID individuals born 1950-1976 whose parents were in the bottom quartile of the income distribution, and who have been followed into adulthood.

Models: Head Start spending per poor 4-year old in the county is centered around $4,230 (and measured in 000s) and instrumented ln(school district per-pupil spending during ages 5-17) is centered around 1.6, to facilitate interpretation of the main effects evaluated at roughly the respective means; the average SFR-induced increase in school-age spending is about 10%. Results are based on 2SLS-Difference-in-Difference models that include: parent's relative rank in income distribution, 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, roll-out of "War on Poverty" & related safety-net programs (community health centers, food stamps, medicaid, AFDC, UI, Title-I (average during childhood yrs)), timing of state-funded Kindergarten intro and timing of tax limit policies; controls for 1960 county characteristics (poverty rate, percent black, education, percent urban, population size, percent voted for Strom Thurmond in 1948 Presidential election*race (proxy for segregationist preferences)) each interacted with linear cohort trends; and controls for childhood family characteristics (parental income/education/occupation, mother's marital status at birth, birth weight, gender). The first-stage model include 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. There exists a significant first-stage.

Table 3:

2SLS-Difference-in-Difference Estimates of Early and K12 education Spending on Adult Outcomes

	Children from Non-Poor Households					
	1	2	3	4	5	6
	Prob(High School Grad)	Years of Education	Ln(Wage), ages 20-50	Ln(Family Income), ages 20-50	Poverty, ages 20-50	Prob(Ever Incarcerated)
Head Start Spending$_{(age\ 4)}$/1000	0.003162	0.02351	0.006178	0.001183	1.64E-04	-0.001766
	(0.0029)	(0.0159)	(0.0052)	(0.0096)	(0.0014)	(0.001613)
Ln(K12 Per-pupil Spending)$_{(age\ 5\text{-}17)}$	0.3159	1.9886	1.1730**	1.1293*	-0.0192	0.1775
	(0.2891)	(1.6299)	(0.4565)	(0.5952)	(0.0994)	(0.1372)
Interaction	0.02439	-0.1684	0.03147	-0.03185	0.01019	-0.01048
	(0.0405)	(0.2551)	(0.0416)	(0.0891)	(0.0131)	(0.01284)
Number of Person-year Observations			79,239	81,598	81,737	
Number of Children	7,718	7,729	7,729	7,729	7,729	4,364

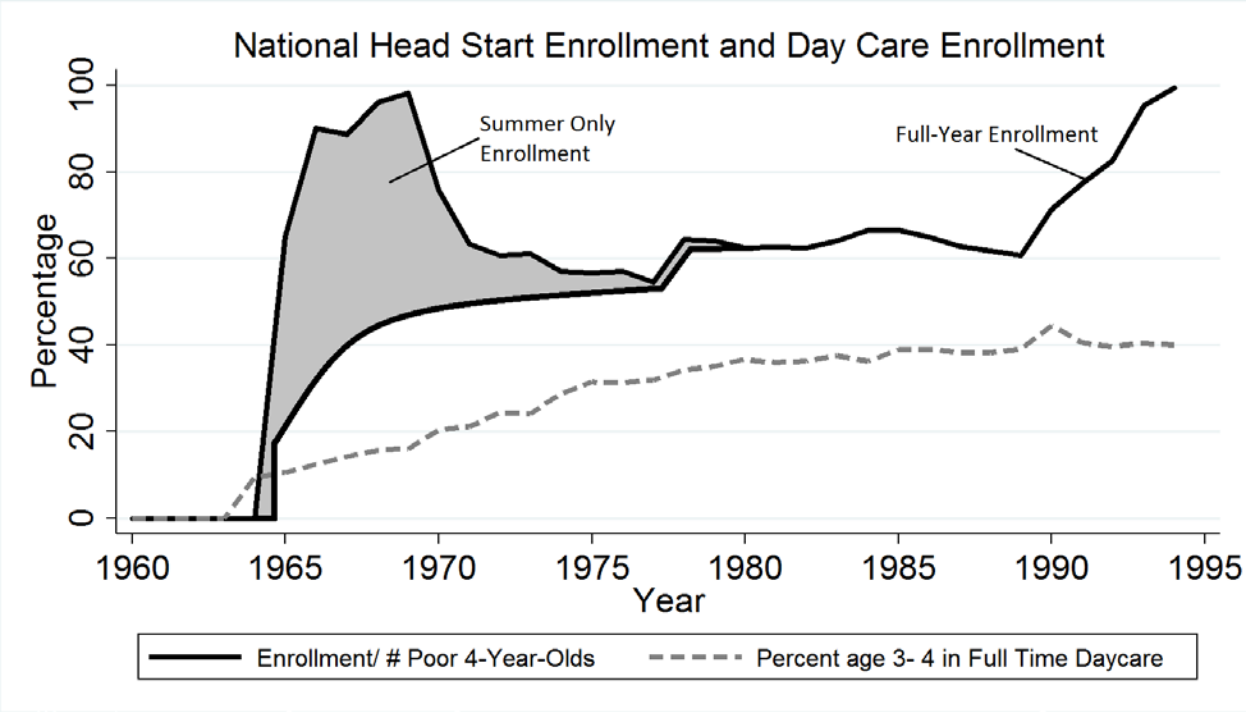
Robust standard errors in parentheses (clustered at childhood state level)

*** p<0.01, ** p<0.05, * p<0.1

Data: PSID geocode Data (1968-2013), matched with childhood school and neighborhood characteristics. Analysis sample includes all PSID individuals born 1950-1976 whose parents were in the bottom quartile of the income distribution, and who have been followed into adulthood.

Models: Head Start spending per poor 4-year old in the county is centered around $4,230 (and measured in 000s) and instrumented ln(school district per-pupil spending during ages 5-17) is centered around 1.6, to facilitate interpretation of the main effects evaluated at roughly the respective means; the average SFR-induced increase in school-age spending is about 10%. Results are based on 2SLS-Difference-in-Difference models that include: parent's relative rank in income distribution, 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, roll-out of "War on Poverty" & related safety-net programs (community health centers, food stamps, medicaid, AFDC, UI, Title-I (average during childhood yrs)), timing of state-funded Kindergarten intro and timing of tax limit policies; controls for 1960 county characteristics (poverty rate, percent black, education, percent urban, population size, percent voted for Strom Thurmond in 1948 Presidential election*race (proxy for segregationist preferences)) each interacted with linear cohort trends; and controls for childhood family characteristics (parental income/education/occupation, mother's marital status at birth, birth weight, gender). The first-stage model include 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. There exists a significant first-stage.

Figure 1:



Notes:

Figure 2: *Effect of Formula Type by District Spending (1972)*

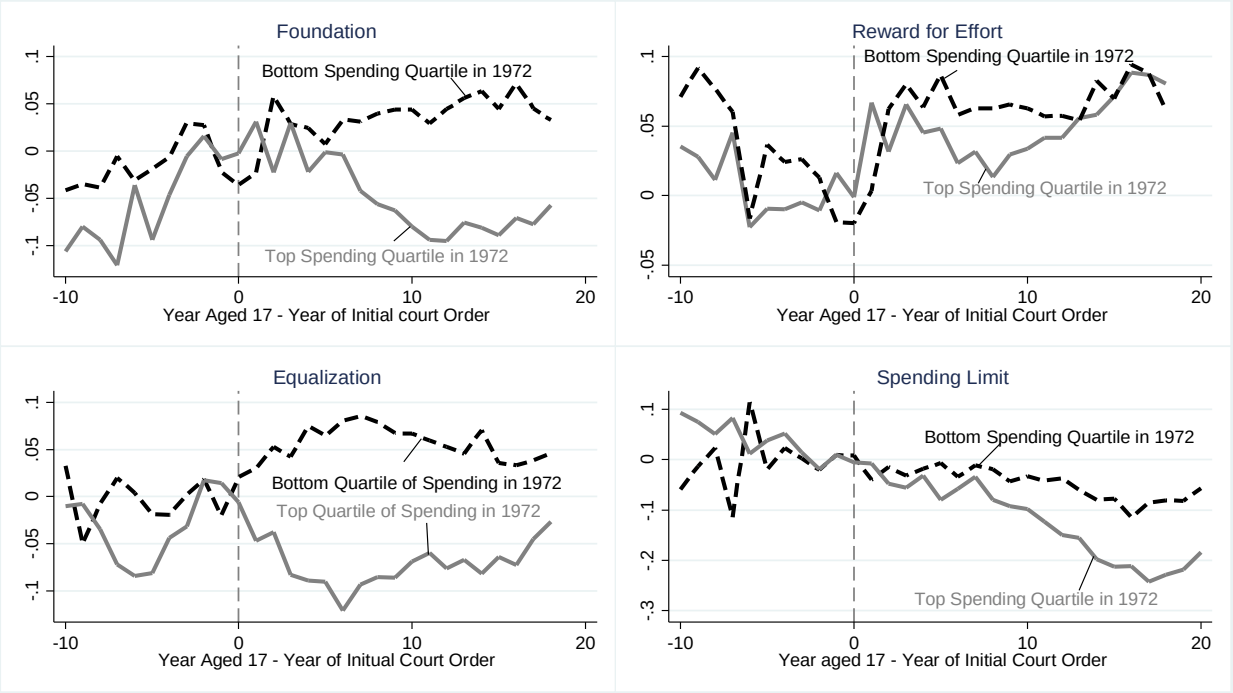


Figure 3: *Effect of Formula Type by District Income (1963)*

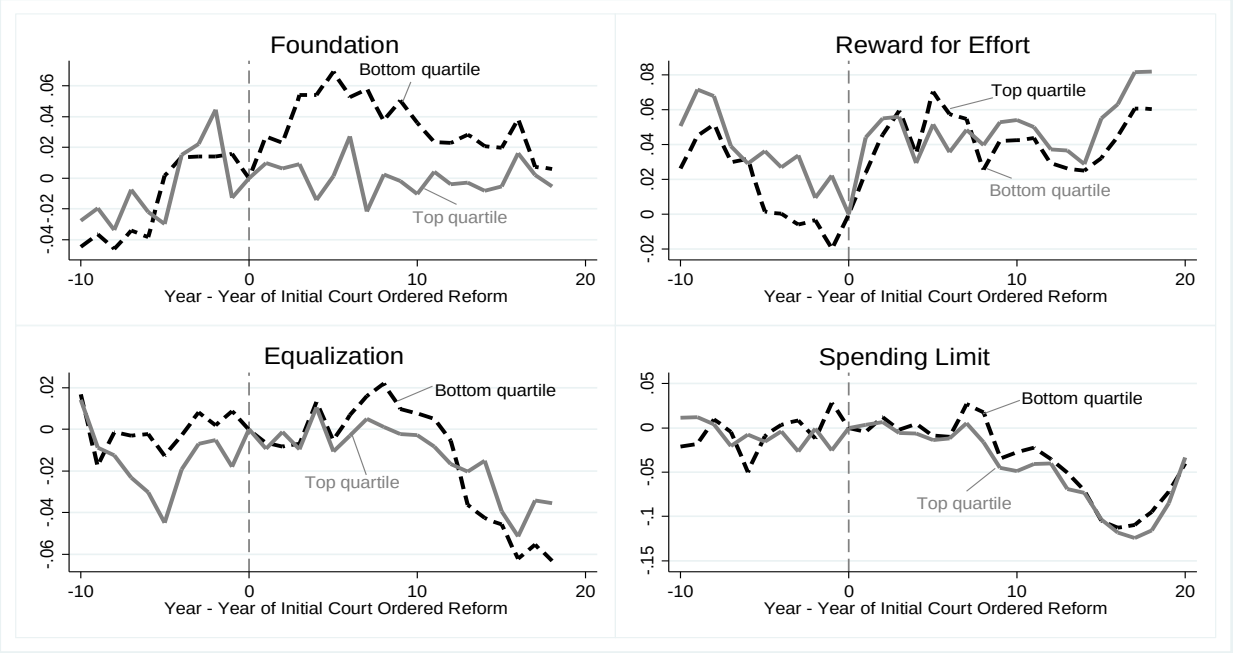
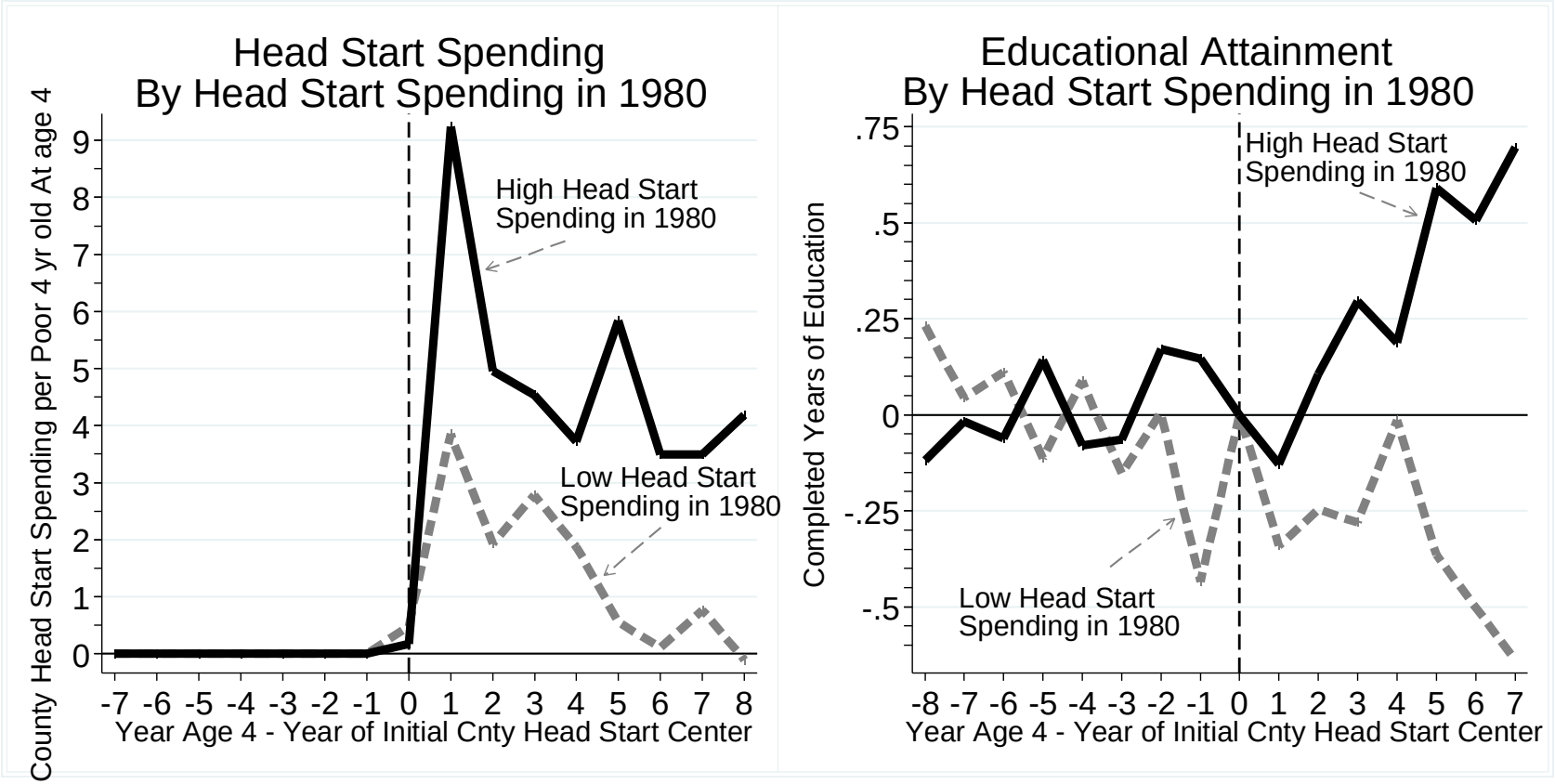
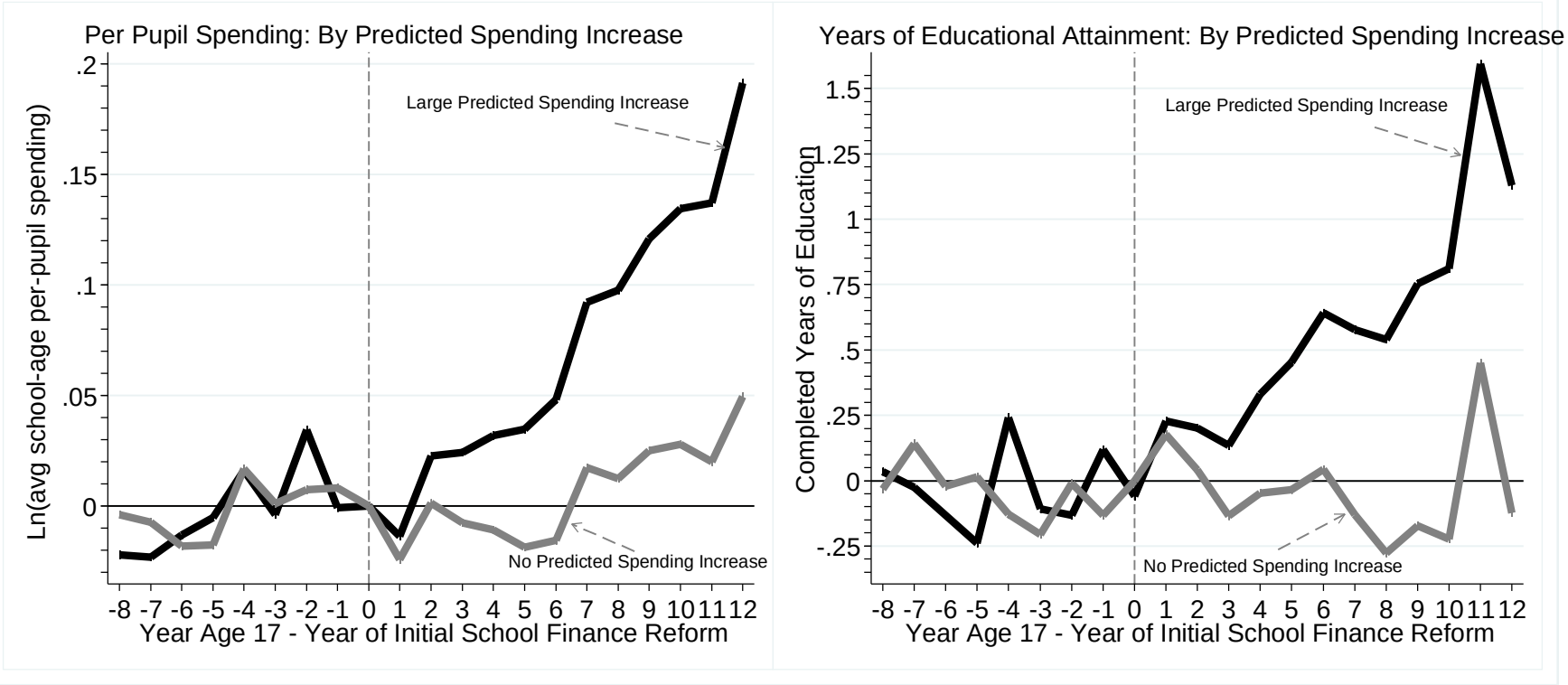


Figure 4: *Evolution of Head Start Spending at Roll-Out*



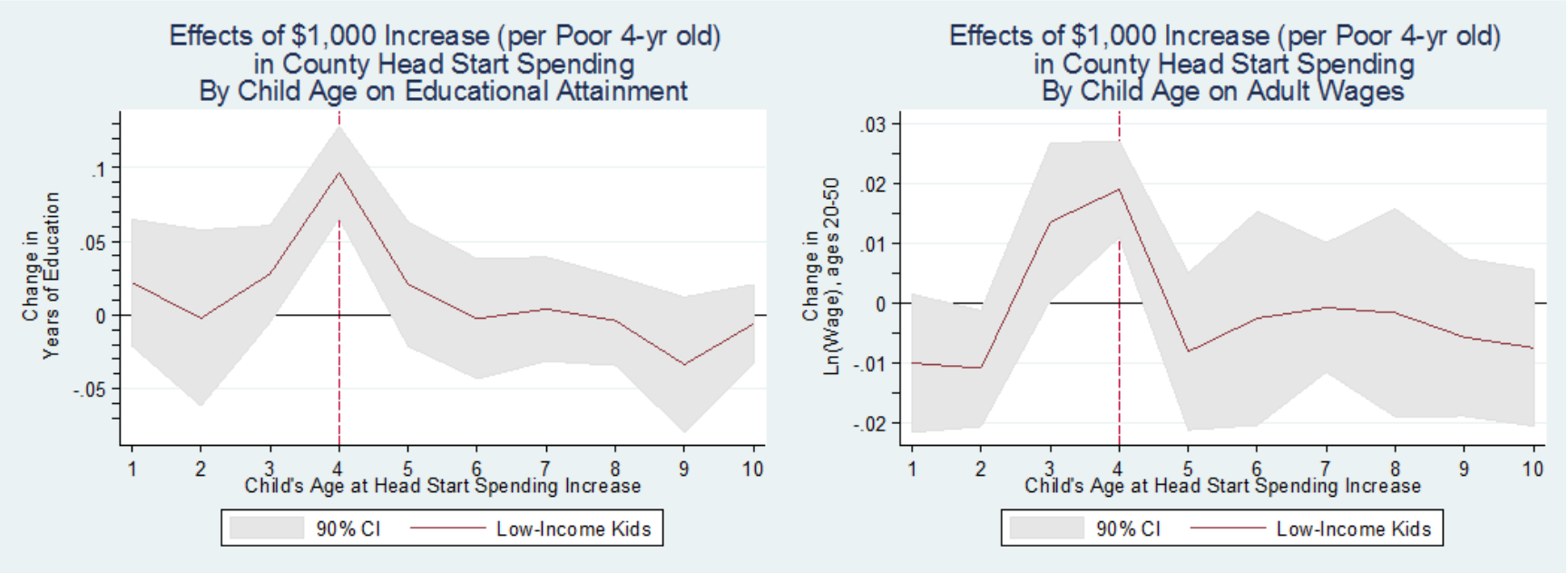
Note: High Head Start spending is determined by ……. The 90 percent confidence interval (grey) is presented around the estimate point estimate (black).

Figure 5



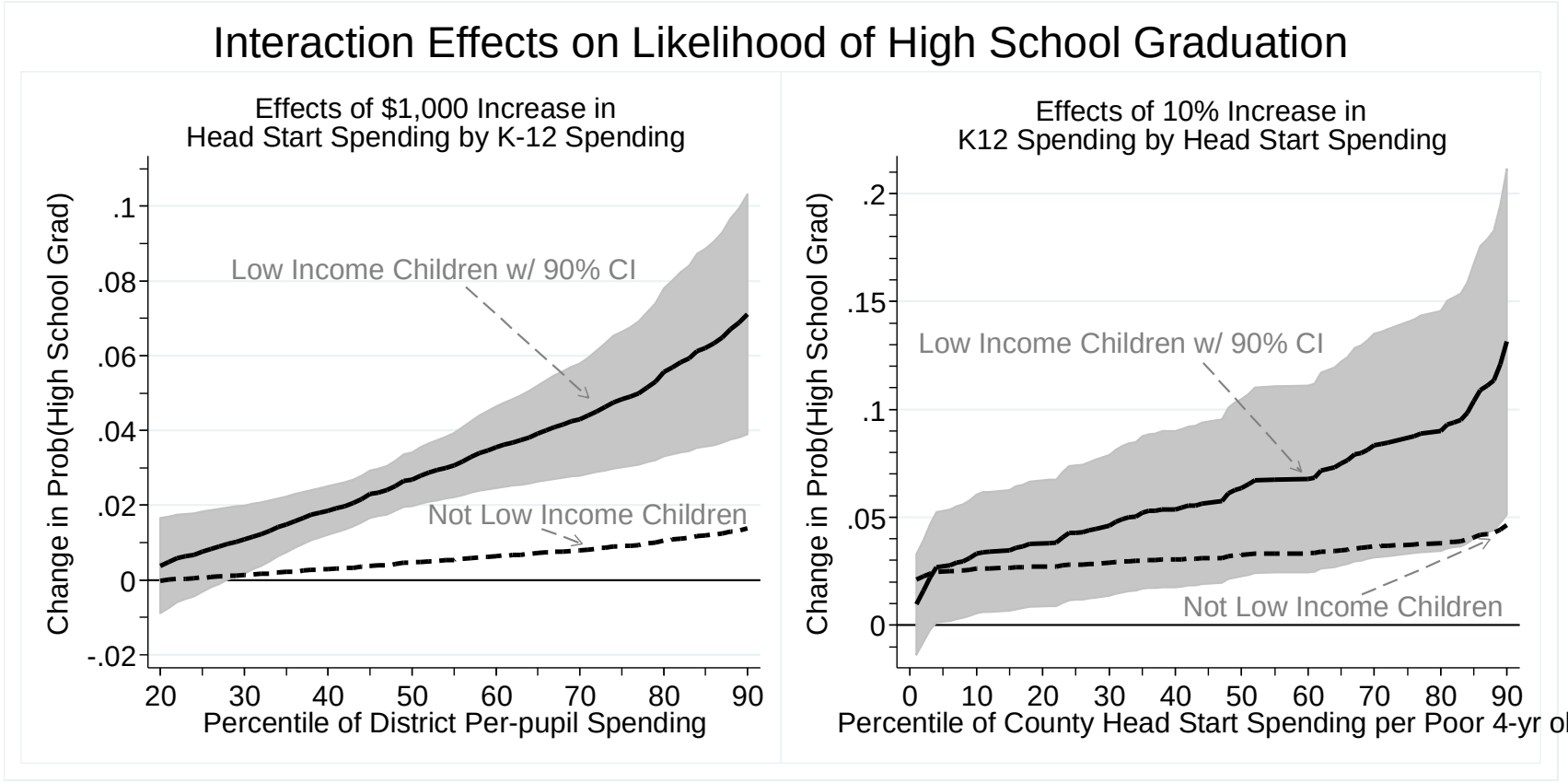
Notes:…….

Figure 6: *Effects of head Start Spending: by Age of Spending* [placeholder]



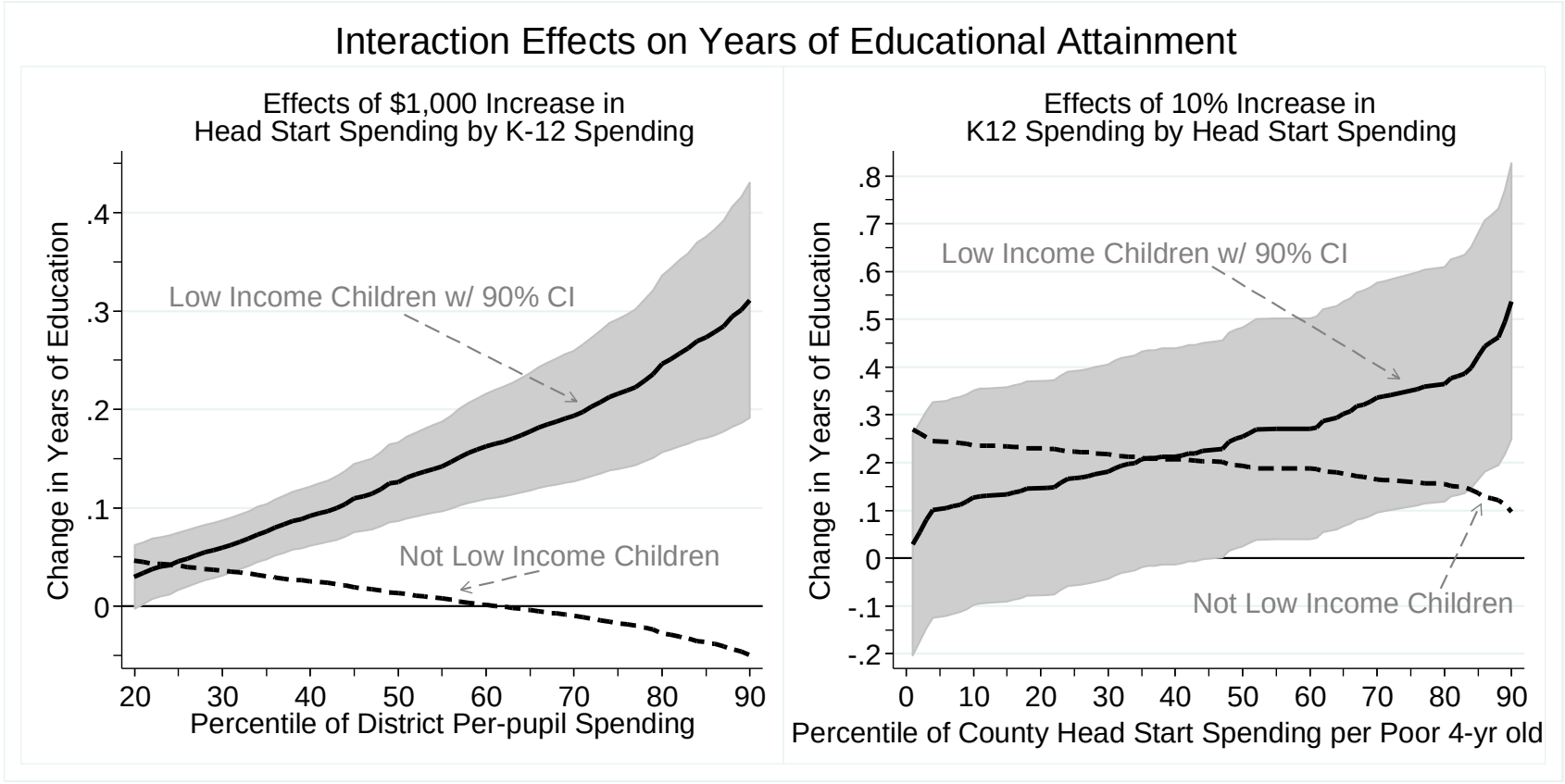
These figures present the marginal effects of Head Start spending in an individual's childhood country at different ages, conditional on the level of Head start spending in the childhood county at age 4 (when such spending should have an effect). The coefficients on the non-eligible years 1 through 3 and 5 through 10, are all conditional on spending at age 4. The coefficient for spending at age 4 is based on model with no other ages included. For the regression estimates underlying this model for all adult outcomes see **Appendix table X**.

Figure 7: *Effect of Head Start Spending by K12 spending Levels and vice Versa on High School Graduation*



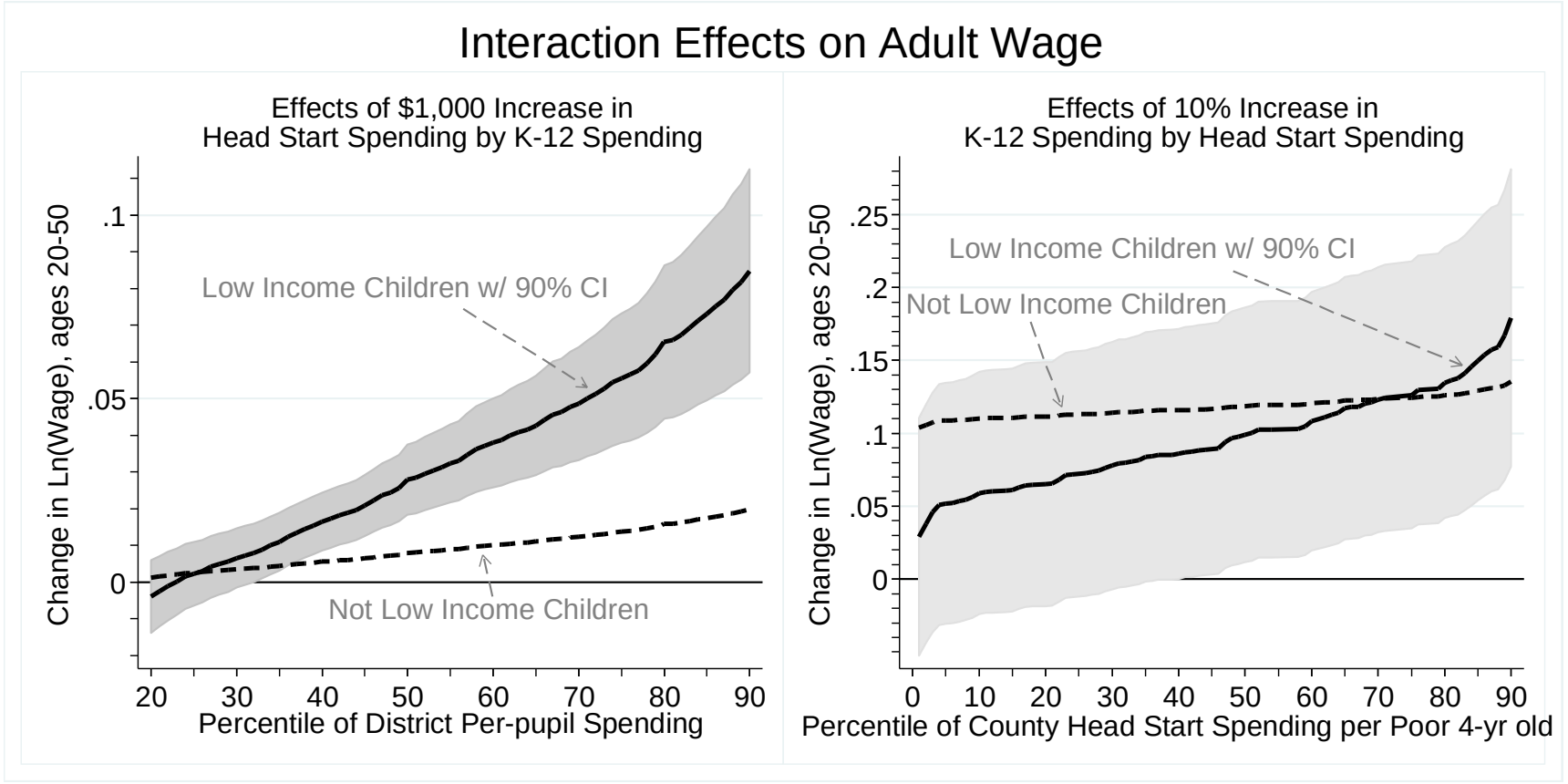
Notes:……

Figure 8: *Effect of Head Start Spending by K12 spending Levels and vice Versa on Years of Education*



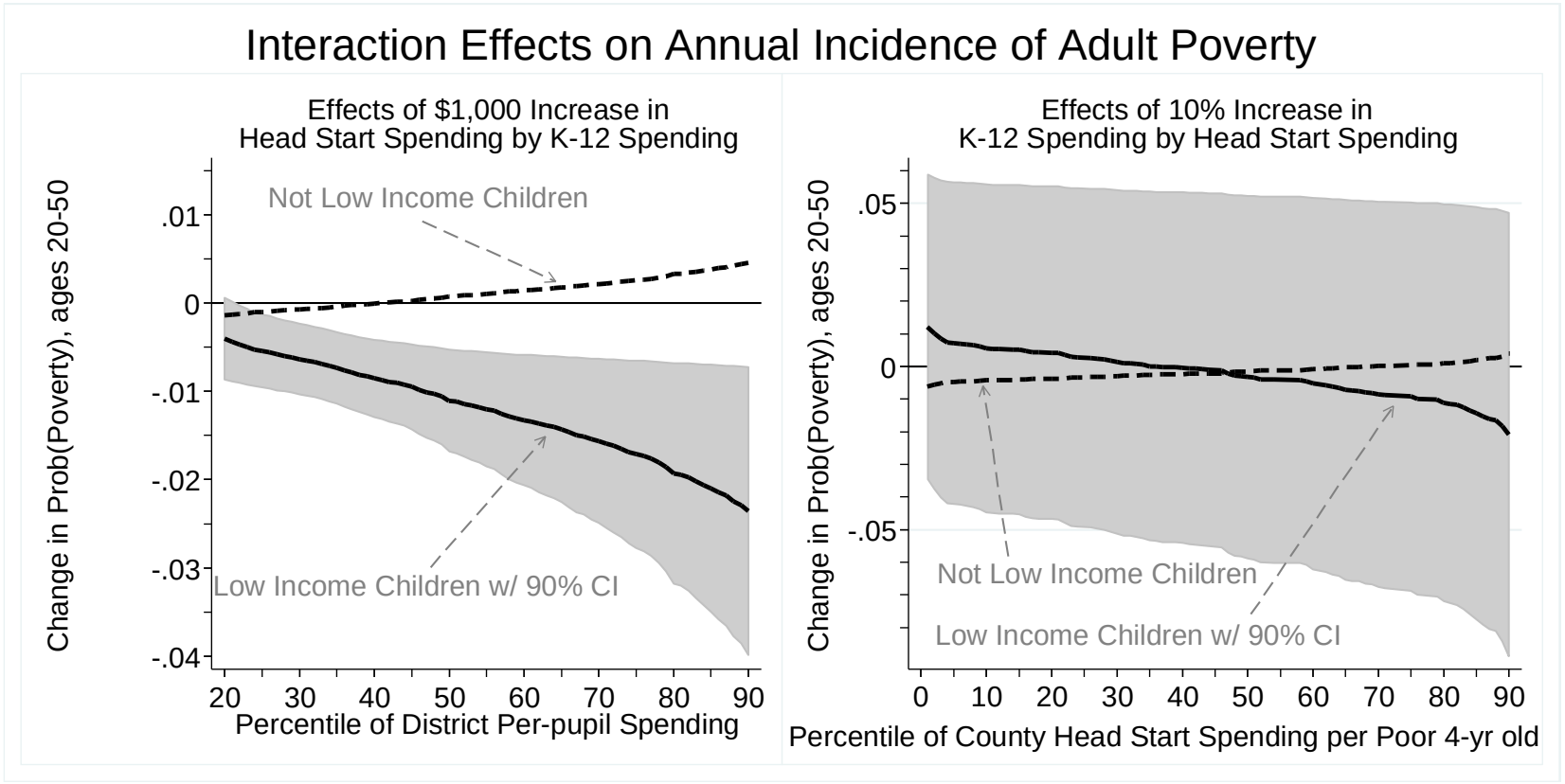
Notes:……

Figure 9: *Effect of Head Start Spending by K12 spending Levels and vice Versa on Adult Wage (ages 20 -50)*



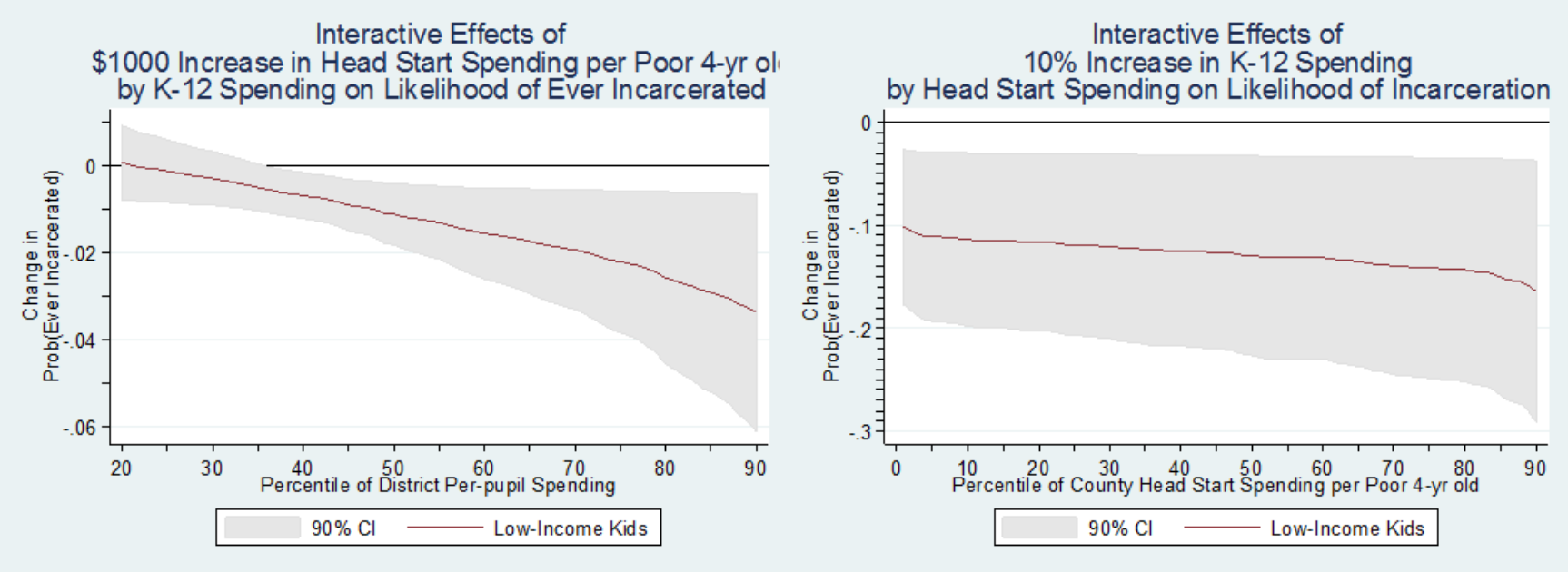
Notes:……

Figure 10: *Effect of Head Start Spending by K12 spending Levels and vice Versa on Likelihood of Adult Poverty (ages 20 -50)*



Notes:……

Figure 11: *Effect of Head Start Spending by K12 spending Levels and vice Versa on ever Being Incarcerated* [Placeholder]



Notes:……

Appendix A
Additional Tables

Figure A1

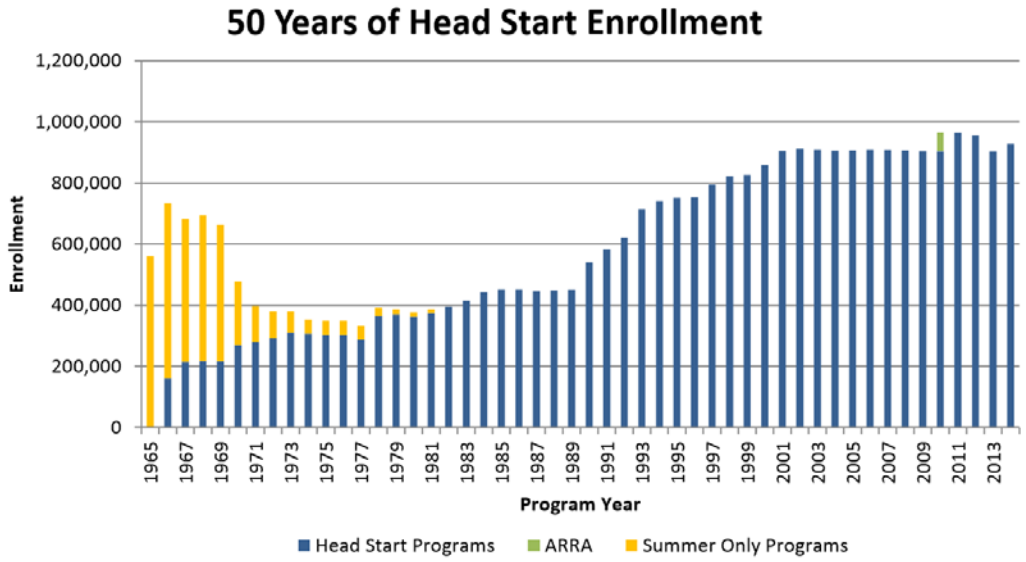
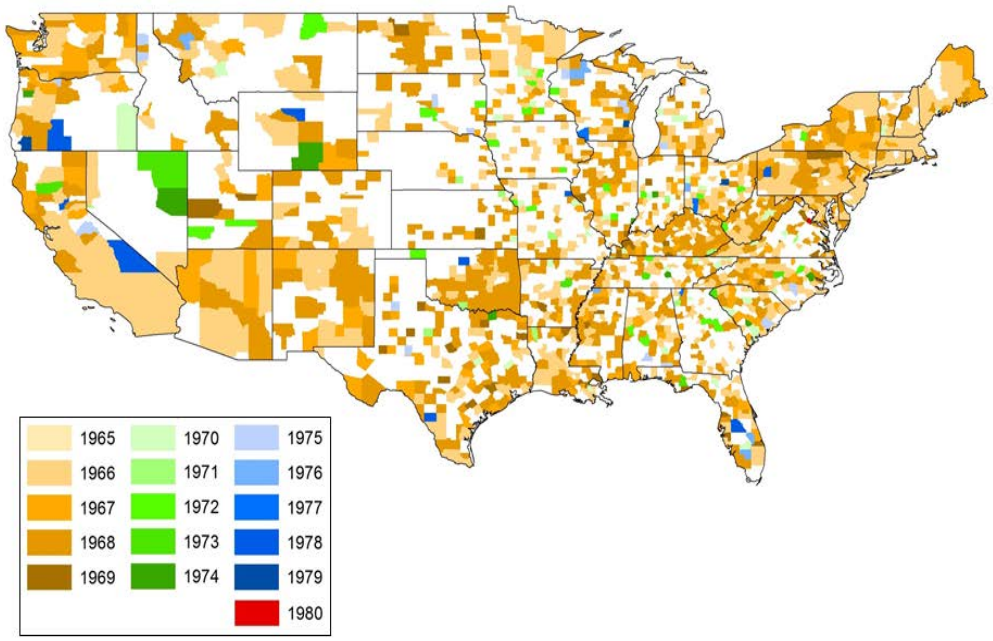


Figure A2



Appendix B

Spending per enrollee versus spending per eligible

Table 2:

	1	2	3	4	5	6	7	8
	Spending per Enrollee		Log of Spending per Enrollee		Log of Head Start Enrollment		Share of Income Eligible 4 year olds enrolled	
Spending per poor 4-year-old	0.0174	0.0379*						
	[0.0359]	[0.0143]						
Log Spending per poor 4-year-old			-0.0192	0.0243*	0.0810+	0.121**	0.648**	
			[0.0271]	[0.00930]	[0.0482]	[0.0438]	[0.0506]	
Log Spending per Enrollee								-0.13
								[0.208]
Year FX	N	Y	N	Y	N	Y	Y	Y
State FX	Y	Y	Y	Y	Y	Y	Y	Y
Observations	612	612	612	612	612	612	612	612
R-squared	0.759	0.927	0.79	0.93	0.996	0.996	0.984	0.909

Robust standard errors in brackets adjusted for clustering at the state level

** p<0.01, * p<0.05, + p<0.1

Notes:

Appendix C:

Matching PSID Individuals to their Childhood School Districts

In order to limit the possibility that school district boundaries were drawn in response to pressure for SFRs, we utilize 1969 school district geographies. The "69-70 School District Geographic Reference File" (Bureau of Census, 1970) relates census tract and school district geographies. For each census tract in the country, it provides the fraction of the population that is in each school district. Using this information, we aggregate census tracts to 1970 district geographies with Geographic Information Systems (GIS) software. We assign census tracts from 1960, 1980 and 1990 to school districts using this resulting digital map based on their centroid locations.

To construct demographic information on 1969-1970-definition school districts, we compile census data from the tract, place, school district and county levels of aggregation for 1960, 1970, 1980 and 1990. We construct digital (GIS) maps of 1970 geography school districts using the 1969-1970 School District Geographic Reference File from the Census. This file indicates the fraction by population of each census tract that fell in each school district in the country. Those tracts split across school districts we allocated to the school district comprising the largest fraction of the tract's population. Using the resulting 1970 central school district digital maps, we allocate tracts in 1960, 1980 and 1990 to central school districts or suburbs based on the locations of their centroids. The 1970 definition central districts located in regions not tracted in 1970 all coincide with county geography which we use instead.

Appendix D
Detaildes dicussion of Datasets

Appendix E
Predicting Dosage

The prediction of $dose_c$ is obtained in two steps. First, using district-by-birth-cohort data for the full universe of districts (not only those represented in the PSID), we estimate [6] where all variables are defined as in [4] and [5], $I_{F,d}$ is an indicator for the type of reform (*F*) (i.e−*foundation plans, spending limits, reward for effort plans, equalization plans, and equity cases*) introduced by the court order in the state containing district *d*, $Q_{ppe72,d}$ are indicators for the quartile of district *d* in the state distribution of per-pupil spending in 1972, and $Q_{inc69,d}$ is the quartile of district *d* in the state distribution of median income in 1969.

[6] $$\ln(PPE_{5-17})_{idb} = \sum_{Q_{ppe=1}}^{4}\sum_{T=-20}^{20}\left(I_{T_{idb}=T}\times I_{Q_{ppe72,d}=Q_{ppe}}\right)\cdot\alpha_{T,Q_{ppe}} + \sum_{F=1}^{5}\sum_{Q_{ppe=1}}^{4}\sum_{T=-20}^{20}\left(I_{T_{idb}=T}\times I_{Q_{ppe72,d}=Q_{ppe}}\times I_{F,d}\right)\cdot\alpha_{T,Q_{inc},F} + \Pi C_{idb} + \theta_{d3} + \theta_{b3} + \varphi_{idb}.$$

The coefficients $\alpha_{T,Q_{ppe}}$ map out the effect of *T* years of exposure to a court-ordered SFR for those from districts in the Q^{th} quartile of the state distribution of per-pupil spending in 1972. Similarly, the coefficients $\alpha_{T,Q_{inc},F}$ map out the effects on school-age per-pupil spending of *T* years of exposure to a court-ordered SFR that introduced reform type *F* for those from districts in the Q^{th} quartile of the state distribution of median income in 1969. Second, for each district we use the predicted spending change (based on reform type implemented by the state and district spending and district income levels prior to reforms) for those who were between the ages of 10 and 15 in the year of the initial court-ordered SFR. As such, our predicted effect from [6] is

[7] $$\text{dose}_c = \left[\sum_{Q_{ppe=1}}^{4}\sum_{T=-20}^{20}\left(I_{T_{idb}=T}\times I_{Q_{ppe72,d}=Q_{ppe}}\right)\cdot\hat{\alpha}_{T,Q_{ppe}} + \sum_{F=1}^{5}\sum_{Q_{ppe=1}}^{4}\sum_{T=-20}^{20}\left(I_{T_{idb}=T}\times I_{Q_{ppe72,d}=Q_{ppe}}\times I_{F,d}\right)\cdot\hat{\alpha}_{T,Q_{inc},F}\right]/6$$

By using the predicted values, dose_c, from [6] from the full universe of school districts as an instrument in a 2SLS regression on the PSID sample, we are implementing a two sample instrumental variables (2S-2SLS) strategy where our excluded instruments are the exposure indicator variables interacted with a function of the reform type implemented by the state, the district income level prior to reforms, and the spending level of the district prior to reforms.[40] This approach captures meaningful variation in K12 spending due to the reforms but removes any variation in spending that is determined by local factors that also influence outcomes.

[40] The two-sample 2SLS estimator was popularized by Angrist and Krueger (1992) and has been used successfully in several other empirical settings (e.g. (e.g., Bjorklund and Jantti, 1997; Currie and Yelowitz, 2000; Dee and Evans, 2003; Borjas, 2004).

Appendix F

Estimated Effects on Head Start Participation

Go into detail about problems here!!!!!

1. 40.4% of low-income children attended Head Start if there was a Head Start Center at age 4 in their county (unweighted). The corresponding sample-weighted estimate is 30%.

2. 1.8% of low-income children attended Head Start if there was no Head Start Center at age 4 in their county (unweighted). The corresponding sample-weighted estimate is 2%.

3. Implies a simple estimate of about 28 percentage points.

4. What about recall bias? With such bias we could have an underestimate!!!!!

5. Simply using national participation rates, 60 percent would not be crazy! But it would be lower in the first period, so that we could use 40 percent as a ball park for the early years. Plus 10 percent of them may be higher income and another bunch may be disabled, so that 33 percent seems like a pretty good estimate.

To gain a sence of the magnitude of the particvation effect we examine information on whether individuals attended preschool, type of program (Head Start vs private pre-K), duration (number of months; part-day/full-day), as collected retrospectively in the special early education module of the 1995 PSID survey wave. Because there are multiple channels through which Head Start roll out in a county can affect student outcomes, **this is not a first stage**. As shown in **Table 2**, on average, among low-income children in our PSID sample, having a Head Start in the county when the child is age 4 increases the likelihood of reporting being in Head Start by 23 percentage points (p-value<0.01). In models that include both having any center and also the spending in the county, and each additional $1,000 increase in program spending increases the likelihood of Head Start attendance by 1.6 percentage points (p-value<0.01). We find no relationship between Head Start spending and preschool attendance among non-poor children.

Appendix G

Robustness Checks and Tests of Validity

Because one of the parameters of interest is the marginal effect of the interaction between Head Start Spending and public K12 spending, it is important for us to establish that the variation we use in each of these is exogenous and will yield causal relationships. Here we present a series of empirical tests that support the validity of each source of variation.

Head Start Spending Effects by Child Age: *No confounding policies*. As a falsification/placebo test, we investigate the effects of Head Start spending increases by the child's age at which these increases occur. If the results are consistent with a causal interpretation of Head Start spending, then we would expect to find significant effects of that spending only for children who are age-eligible (age 4), and not for children who were already school-age at the time of the spending increase. Furthermore, even though our models control for a variety of other policies, and we find no Head Start effects on non-poor children, one may still worry that the timing of Head Start rollout or the timing of SFRs coincided with other policies that also improve adult outcomes. One test of this would be to determine whether the effects of the spending increases are experienced only among those who were of the appropriate age. If counties or districts adopted other policies to improve outcomes for low-income children (that were not targeted to the exact same age range as that in question) one would observe improvements for other age ranges also. To test this for Head Start spending, we estimated the marginal effect of the level of Head Start spending that prevailed when the individual was different ages. To test whether Head Start spending at other ages predicts student outcomes, conditional on Head Start spending at age 4, we estimate the following regressions where all variables are as defined in [4] and [6].

[G1] $$Y_{icb} = \beta_w \cdot HS_{cb}^{age\,W} + \beta \cdot HS_{cb}^{age\,4} + \gamma \cdot C_{icb} + \theta_c + \tau_b + \varepsilon_{icb}.$$

We estimate models such as [G1] where we include our repressor of interest (β_w) is the marginal effect of Head Start spending at age W on individual outcomes, conditional ion the effect of Head Start spending at age 4. In principle, one should see that Head Start sending per poor four-year old has effects when the individual was four years old but not at other ages. This is exactly what we find across every one of the adult outcomes of poor children we analyze. In figure 6 and appendix figures **G1**, we plot the marginal effect of head start spending by age conditional on spending at age 4. Note that the estimated effect for age 4 is not conditional on spending at other ages. However, the marginal effect of spending at age 4 is largely the same in models that include spending at other ages. The **figures** all shows that increases in the Head Start spending level that prevailed when the individual was four years old are associated with significantly improved adult outcomes while the corresponding spending level at ineligible ages (1-3;5-10) are not.

Even though we instrument for K12 spending levels, it is important to establish that the identifying variation we use is valid. If the spending increases we exploit operate through improved K12 education, one should see improvement for those who were between the ages of 5 and 17 when there was a school finance reform, but no effect for individual from the same districts who were 18 or older at the time. Figures 5 shows that only those individuals who were of school-going age at the time of a reform induced spending increase experience improved outcomes. These figures also reveal that outcomes in districts that saw increases in K12 spending were not on a positive or negative trajectory – indicating that the timing of the SFR were exogenous to the underlying trends in outcomes in affected districts. To show this more formally, we estimate models that instrument for K12 spending in an individual's childhood districts when they were

between the ages of 18 and 22. Results are in table E1. If the effects are real, we should see effects for reforms induced spending increases when an individual was between the ages of 5 and 17 but not for increases that occurred when an individual was between the ages of 18 and 22. As in Jackson, Johnson and Persico (2016), K12 spending levels between ages 18 to 22 have no effect on outcomes.

Instrumenting for Head Start Spending

Even though we present several empirical tests that suggest that the within county variation in Head Start identified in [1] is valid, one may worry that some of the variation in Head Start spending is determined by county-level choices that may be endogenous to student outcomes. For example, counties that decided to implement generous Head Start funding levels may also have implemented other policies that affect student outcomes or may differ in unobserved ways from those that decided on less generous funding levels. Though we find no evidence that this was the case, to address this possibility we proposed an instrumental variables strategy that removes the variation in the level of Head Start spending that is driven by county level decisions and relies exclusively on variation in the timing of Head Start rollout.

Figure 4 shows that most of the within-county variation in Head Start spending over time occurs before and after rollout. As such, we instrument for the level of Head Start spending per poor 4-year old in an individual's county when they were 4 years old with an indicator for whether Head Start was available in the individual's childhood county when they were 4 years old, and the number of years a Head Start center had existed in childhood county at age 4. This instrumental variables strategy compares the adult outcome of individuals who were 4 years old before Head Start was available in their county (unexposed cohorts with zero spending), to the outcomes of adult from the same childhood county were 4 years old soon after rollout (early exposed cohort with moderate levels of Head Start spending, on average) and also those from the same childhood county were 4 years old long after rollout (late exposed cohorts with high levels of Head Start spending, on average). To account for underlying cohort level differences on can use the difference across the same birth cohorts in non-roll-out counties as a basis for comparison. The identifying assumption is that the timing of Head Start rollout in a county is unrelated to other changes at the county level that exert an independent effect on adult outcomes. Under this identifying assumption, if, *among individuals from the same childhood county*, the late exposed cohorts have better outcomes than the early exposed cohort, and the early exposed cohorts have better outcomes than the unexposed cohorts (relative to the differences across the same birth cohorts in counties that did not rollout Head Start across the same time period), it would imply a real casual effect of Head Start spending on adult outcomes. Figure 4 already presents some visual evidence of this.

The Instrumental Variables model that exploits the variation described above is implemented by estimating [G2] and [G3] by Two Stage Least Squares (2SLS) where [G2] is a first stage regression based on the Head Start rollout instruments.

[G2] $\widehat{HS}_{icb}^{age\,4} = \rho_1 \cdot HSExp_{icb} + \rho_2 \cdot Years(HSExp_{icb}) + \rho_3 \cdot C_{icb} + \theta_{c1} + \theta_{b1} + \varphi_{icb}.$

[G3] $Y_{icb} = \beta \cdot \widehat{HS}_{icb}^{age\,4} + \gamma \cdot C_{icb} + \theta_c + \theta_b + \varepsilon_{icb}.$

All common variables are defined as in [3]. $\widehat{HS}_{cb}^{age\,4}$is the first stage prediction of Head Start spending per 4-year old, $HSExp$ is an indicator for whether individual *i* from birth cohort *b* was age four (or younger) when Head Start was first established in their childhood county *c*, and $Years(HSExp_{icb})$ is the number of years that Head Start had been established in childhood county *c* for individual *i*, at the time birth cohort b was 4 years old.

To test or interaction model also, we implement models that instrument for both K12

spending and the interaction, and also for Head Start spending. Specifically, we estimate [G4]] below, where $\widehat{HS}_{cb}^{age\ 4}$ is estimated in a first stage using the number of years that Head Start existed in an individual's county as our exogenous predictor of Head Start spending per four-year-old at age 4, and $\widehat{INT}_{idb}$ is estimated in a first stage based on the interactions between the Head Start rollout instruments and the SFR-reforms instruments.

[G4] $Y_{icb} = \beta_{HS} \cdot \widehat{HS}_{cb}^{age\ 4} + \beta_{k12} \cdot \ln\left(\widehat{PPE}_{5-17}\right)_{idb} + \beta_{int} \cdot (\widehat{INT}_{idb}) + \gamma \cdot C_{icb} + \theta_d + \theta_b + \varepsilon_{idb}$.

In [G4], all of the exogenous shifters in K12 spending ($(SFRExp_{idb} \times dose_c)$ and $(SFRExp_{idb})$) are interacted with the exogenous shifters of Head Starts spending ($(HSExp_{icb})$ and $Years(HSExp_{icb})$), and are used to create the first stage predictions of K12 spending ($\ln\left(\widehat{PPE}_{5-17}\right)_{idb}$), Head Start spending ($\widehat{HS}_{cb}^{age\ 4}$) and the interaction between the two ($\widehat{INT}_{idb}$). While this model yields noisier estimates, if the results between model [G4] are similar to those in [6] it will be compelling evidence that our estimated effects are real.

Results from this strategy that instruments for K12 spending, Head Start spending and the interactions between the two is presented in **Table G1**. While this model yields less precise point estimates than those in Table 2, the pattern of results is very similar – this suggest that changes in Head Start spending are unrelated to other changes. Because these 2SLS models that use the timing of Head Start rollout rely on the exogeneity of the timing of Head Start rollout, we also present evidence that this alternate variation is likely valid. To show this, we present event study figures for the effect of Head Start rollout for counties with high spending level versus other counties in **Figure 4**. As discussed previously, there is no evidence of any pre-existing time trends, suggesting that the timing of rollout was exogenous.

No selection or endogenous mobility: Another concern one may have with the estimates is that due to selective migration or neighborhood change, the characteristics of the individuals exposed to different levels of K12 spending or Head Start spending are not the same. We address this possible concern in two ways. First, we demonstrate that the spending changes we exploit are unrelated to observed family and neighborhood characteristics. Specifically, we regress year of educational attainment and the adult wage on several observable characteristics (list here) and then take the fitted values from those regression as our predicted outcomes. The predicted educational attainment and predicted wage are weighted averages of the observed covariates. We then regress our predicted outcomes on the spending changes (excluding all of these same observable characteristics). If the spending changes are unrelated to those observable characteristics that predict the adult outcome, the estimated coefficients will be zero. Indeed, this is what we find (See Table **G2**).

Even though our spending changes are unrelated to observed characteristics, one may worry about selection on unobserved characteristics. To rule out the possibility that our results are driven by differences across treated and untreated families, 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. This approach accounts for all observed and unobserved shared family characteristics that predict outcomes. We achieve this by augmenting [6] to include sibling fixed effects. As one can see in Table G4, the estimated coefficients for low-income children are very similar to those without sibling fixed effects. This suggests that family selection cannot explain the main pattern of results. This sibling tests also addresses any potential lingering concerns regarding endogenous mobility driving the results

because individuals in the same family have the same residential address. As an additional check on endogenous mobility, we re-estimated all models limiting the analysis sample to those who lived at their (earliest) childhood residence prior to the enactment of Head Start programs in their respective county. The results are presented in Appendix Table G5. We find nearly identical results as those in the full sample. This indicates that endogenous residential mobility is not an important source of bias in this analysis.

PLACEHOLDER—UPDATE and ADD INCERATION!!!!

:Appendix Table G1:

2SLS/IV Estimates of Court-Ordered School Finance Reform Induced Effects of Per-Pupil Spending on Long-Run Outcomes: Placebo Tests for Non-school Ages (All children. All adult outcomes are measured between ages 20-45)

	Years of Education	Prob(High School Grad)	Ln(Wage)	Ln(Family Income)	Prob(poverty)	Prob(incarcerated)
	1	2	3	4		5
$Ln(PPE_d)_{(age\ 5\text{-}17)}$	3.2957**	0.6063***	0.8750***	1.0694***	-0.3739***	-0.3739***
	(1.2963)	(0.1509)	(0.2029)	(0.2761)	(0.0976)	(0.0976)
$Ln(PPE_d)_{(age\ 20\text{-}24)}$	-0.1700	-0.2878	-0.0705	-0.2499	-0.0959	-0.0959
	(2.5971)	(0.3954)	(0.4407)	(0.5954)	(0.2436)	(0.2436)
Number of Individuals	15353	15353	13183	14730	14737	14737
Number of Districts	1409	1409	1395	1414	1414	1414

Robust standard errors in parentheses (clustered at school district level) *** p<0.01, ** p<0.05, * p<0.10

Data: PSID geocode Data (1968-2011), matched with childhood school and neighborhood characteristics. Analysis sample includes all PSID individuals born 1955-1985, followed into adulthood through 2011. Sampling weights are used so that the results are nationally representative.

Models: The key treatment variable, $Ln(PPE_d)_{(age\ 5\text{-}17)}$, is the natural log of average school-age per pupil spending. All models include school district fixed effects, birth cohort fixed effects, and the additional controls listed below. The excluded instruments from the second stage are (the number of years of exposure to a court-ordered SFR) and (the number of years of exposure to a court-ordered SFR) × (the quartile of the district in the distribution of $Spend_d$) and (the number of years of between the ages of 20 and 24 that occur after a court-ordered SFR) and (the number of years of between the ages of 20 and 24 that occur after a court-ordered SFR) × (the quartile of the district in the distribution of $Spend_d$).

Additional controls: childhood family characteristics (parental income/education/occupation, mother's marital status at birth, birth weight, gender). Also race × census division × birth cohort fixed effects; controls at the county-level for the timing of school desegregation by race, hospital desegregation × race, roll-out of community health centers, county expenditures on Head Start (at age 4), food stamps, medicaid, AFDC, UI, Title-I (average during childhood years), timing of state-funded Kindergarten intro and timing of tax limit policies; controls for 1960 county characteristics (poverty rate, percent black, education, percent urban, population size, percent voted for Strom Thurmond in 1948 Presidential election*race (proxy for segregationist preferences)) each interacted with linear cohort trends.

PLACEHOLDER: Table G2:

2SLS-Difference-in-Difference Estimates of Early and K12 education Spending on PREDICTED Adult Outcomes Children from Poor Households

	1	2	3	4	5	6
	Predicted Prob(High School Grad)	Predicted Years of Education	Predicted Ln(Wage), ages 20-50	Predicted Ln(Family Income), ages 20-50	Predicted Poverty, ages 20-50	Predicted Prob(Ever Incarcerated)
Head Start Spending$_{(age\ 4)}$/1000	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)
Ln(K12 Per-pupil Spending)$_{(age\ 5\text{-}17)}$	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)
Interaction	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)
Number of Person-year Observations			49,282	54,064	54,309	
Number of Childhood Families	2,024	2,023	2,016	2,098	2,098	1,693
Number of Children	5,385	5,378	5,378	5,378	5,378	4,317

Robust standard errors in parentheses (clustered at childhood state level)

*** p<0.01, ** p<0.05, * p<0.1

Data: PSID geocode Data (1968-2013), matched with childhood school and neighborhood characteristics. Analysis sample includes all PSID individuals born 1950-1976 whose parents were in the bottom quartile of the income distribution, and who have been followed into adulthood.

Models: Head Start spending per poor 4-year old in the county is centered around $4,230 (and measured in 000s) and instrumented ln(school district per-pupil spending during ages 5-17) is centered around 1.6, to facilitate interpretation of the main effects evaluated at roughly the respective means; the average SFR-induced increase in school-age spending is about 10%. Results are based on 2SLS-Difference-in-Difference models that include: no controls.

Predicted outcomes are fitted values of each outcome son parent's relative rank in income distribution, 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, roll-out of "War on Poverty" & related safety-net programs (community health centers, food stamps, medicaid, AFDC, UI, Title-I (average during childhood yrs)), timing of state-funded Kindergarten intro and timing of tax limit policies; controls for 1960 county characteristics (poverty rate, percent black, education, percent urban, population size, percent voted for Strom Thurmond in 1948 Presidential election*race (proxy for segregationist preferences)) each interacted with linear cohort trends; and controls for childhood family characteristics (parental income/education/occupation, mother's marital status at birth, birth weight, gender). The first-stage model include 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. There exists a significant first-stage.

Table G3:

2SLS-2SLS Estimates of Early and K12 education Spending on Adult Outcomes

	Children from Poor Households					
	1	2	3	4	5	6
	Prob(High School Grad)	Years of Education	Ln(Wage), ages 20-50	Ln(Family Income), ages 20-50	Poverty, ages 20-50	Prob(Ever Incarcerated)
Instrumented: Head Start Spending$_{(age}$	0.09010***	0.2768**	0.03779*	0.05518**	-0.02969***	-0.01936*
	(0.03177)	(0.1236)	(0.01966)	(0.02657)	(0.01146)	(0.01117)
Instrumented: Ln(K12 Per-pupil	1.5488***	3.4924*	0.8023*	1.5388**	-0.4070+	-1.5206**
	(0.4422)	(2.0229)	(0.4570)	(0.7830)	(0.3084)	(0.7175)
Instrumented: Interaction	0.3207***	0.4744*	0.1574**	0.2732***	-0.1534***	-0.1411*
	(0.1087)	(0.2587)	(0.06900)	(0.1010)	(0.05388)	(0.08434)
Number of Person-year Observations			49,282	54,064	54,309	
Number of Childhood Families	2,024	2,023	2,016	2,098	2,098	1,693
Number of Children	5,385	5,378	5,378	5,378	5,378	4,317

Robust standard errors in parentheses (clustered at childhood state level)

*** p<0.01, ** p<0.05, * p<0.1

Data: PSID geocode Data (1968-2013), matched with childhood school and neighborhood characteristics. Analysis sample includes all PSID individuals born 1950-1976 whose parents were in the bottom quartile of the income distribution, and who have been followed into adulthood.

Models: Instrumented Head Start spending per poor 4-year old in the county is centered around $4,230 (and measured in 000s) and instrumented ln(school district per-pupil spending during ages 5-17) is centered around 1.6, to facilitate interpretation of the main effects evaluated at roughly the respective means; the average SFR-induced increase in school-age spending is about 10%. Results are based on 2SLS-IV models that include: parent's relative rank in income distribution, 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, roll-out of "War on Poverty" & related safety-net programs (community health centers, food stamps, medicaid, AFDC, UI, Title-I (average during childhood yrs)), timing of state-funded Kindergarten intro and timing of tax limit policies; controls for 1960 county characteristics (poverty rate, percent black, education, percent urban, population size, percent voted for Strom Thurmond in 1948 Presidential election*race (proxy for segregationist preferences)) each interacted with linear cohort trends; and controls for childhood family characteristics (parental income/education/occupation, mother's marital status at birth, birth weight, gender), and age (cubic). The first-stage model include 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 instruments for Head Start spending per poor 4-year old include: a dummy indicator of whether a county Head Start center had ever been established by the time the individual was age 4 and the number of years a Head Start center had existed in childhood county at age 4; and these instruments are interacted with school-age years of exposure to school finance reform*top quartile of the respective school district's predicted reform-induced change in school spending based on the timing and type of court-ordered reform. There exists a significant first-stage.

Table G4:

Within family Model: 2SLS-Difference-in-Difference Estimates of Early and K12 education Spending on Adult Outcomes

	Children from Poor Households					
	1	2	3	4	5	6
	Prob(High School Grad)	Years of Education	Ln(Wage), ages 20-50	Ln(Family Income), ages 20-50	Poverty, ages 20-50	Prob(Ever Incarcerated)
Head Start Spending$_{(age\ 4)}$/1000	0.02044*	0.1153***	0.008247+	0.02255**	-0.002645	-0.01890***
	(0.01096)	(0.02779)	(0.005471)	(0.009970)	(0.003889)	(0.006655)
Ln(K12 Per-pupil Spending)$_{(age\ 5\text{-}17)}$	0.9831**	4.4541***	0.7861+	2.3433**	-0.4527*	-1.3356*
	(0.4508)	(1.7092)	(0.5467)	(0.9161)	(0.2667)	(0.7545)
Interaction	0.09013+	0.5500**	0.01816	0.09604	-0.05668+	0.03364
	(0.06069)	(0.2457)	(0.05940)	(0.09992)	(0.04317)	(0.07735)
Number of Person-year Observations			49,282	54,064	54,309	
Number of Childhood Families	2,024	2,023	2,016	2,098	2,098	1,693
Number of Children	5,385	5,378	5,378	5,378	5,378	4,317

Robust standard errors in parentheses (clustered at childhood state level)

*** p<0.01, ** p<0.05, * p<0.1

Data: PSID geocode Data (1968-2013), matched with childhood school and neighborhood characteristics. Analysis sample includes all PSID individuals born 1950-1976 whose parents were in the bottom quartile of the income distribution, and who have been followed into adulthood.

Models: Head Start spending per poor 4-year old in the county is centered around $4,230 (and measured in 000s) and instrumented ln(school district per-pupil spending during ages 5-17) is centered around 1.6, to facilitate interpretation of the main effects evaluated at roughly the respective means; the average SFR-induced increase in school-age spending is about 10%. Results are based on 2SLS-Difference-in-Difference models that include: parent's relative rank in income distribution, 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, roll-out of "War on Poverty" & related safety-net programs (community health centers, food stamps, medicaid, AFDC, UI, Title-I (average during childhood yrs)), timing of state-funded Kindergarten intro and timing of tax limit policies; controls for 1960 county characteristics (poverty rate, percent black, education, percent urban, population size, percent voted for Strom Thurmond in 1948 Presidential election*race (proxy for segregationist preferences)) each interacted with linear cohort trends; and controls for childhood family characteristics (parental income/education/occupation, mother's marital status at birth, birth weight, gender). The first-stage model include 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. There exists a significant first-stage.

PLACEHOLDER Table G5:

Early Address Sample: *2SLS-Difference-in-Difference Estimates of Early and K12 education Spending on Adult Outcomes*

	Children from Poor Households					
	1	2	3	4	5	6
	Prob(High School Grad)	Years of Education	Ln(Wage), ages 20-50	Ln(Family Income), ages 20-50	Poverty, ages 20-50	Prob(Ever Incarcerated)
Head Start Spending$_{(age\ 4)}$/1000	0.02044*	0.1153***	0.008247+	0.02255**	-0.002645	-0.01890***
	(0.01096)	(0.02779)	(0.005471)	(0.009970)	(0.003889)	(0.006655)
Ln(K12 Per-pupil Spending)$_{(age\ 5\text{-}17)}$	0.9831**	4.4541***	0.7861+	2.3433**	-0.4527*	-1.3356*
	(0.4508)	(1.7092)	(0.5467)	(0.9161)	(0.2667)	(0.7545)
Interaction	0.09013+	0.5500**	0.01816	0.09604	-0.05668+	0.03364
	(0.06069)	(0.2457)	(0.05940)	(0.09992)	(0.04317)	(0.07735)
Number of Person-year Observations			49,282	54,064	54,309	
Number of Childhood Families	2,024	2,023	2,016	2,098	2,098	1,693
Number of Children	5,385	5,378	5,378	5,378	5,378	4,317

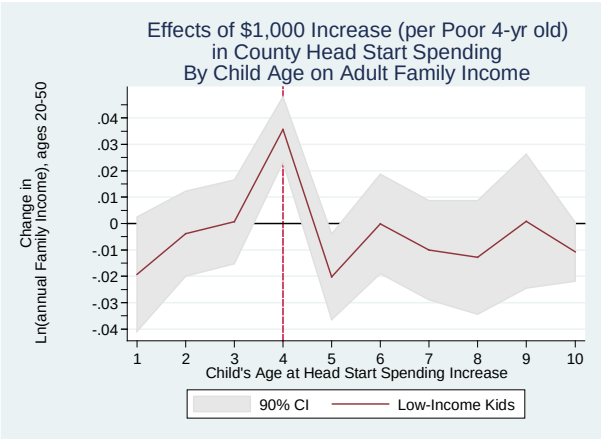
Robust standard errors in parentheses (clustered at childhood state level)

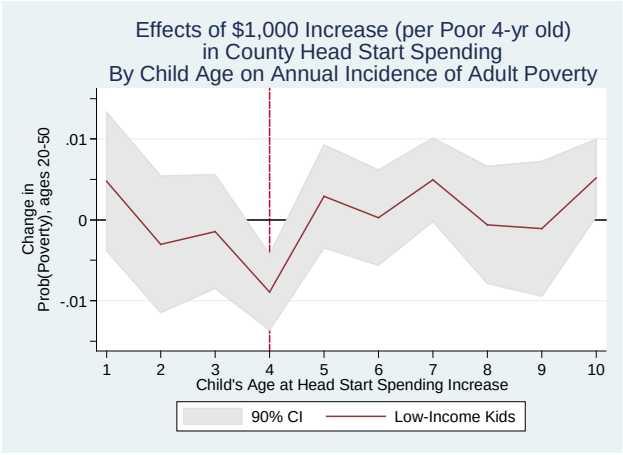
*** p<0.01, ** p<0.05, * p<0.1

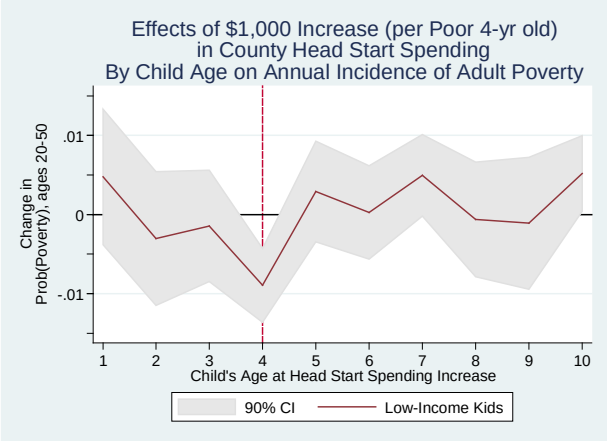
Data: PSID geocode Data (1968-2013), matched with childhood school and neighborhood characteristics. Analysis sample includes all PSID individuals born 1950-1976 whose parents were in the bottom quartile of the income distribution, and who have been followed into adulthood.

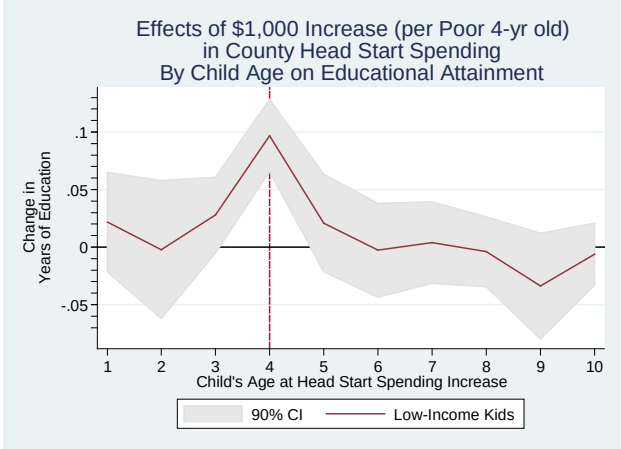
Models: Head Start spending per poor 4-year old in the county is centered around $4,230 (and measured in 000s) and instrumented ln(school district per-pupil spending during ages 5-17) is centered around 1.6, to facilitate interpretation of the main effects evaluated at roughly the respective means; the average SFR-induced increase in school-age spending is about 10%. Results are based on 2SLS-Difference-in-Difference models that include: parent's relative rank in income distribution, 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, roll-out of "War on Poverty" & related safety-net programs (community health centers, food stamps, medicaid, AFDC, UI, Title-I (average during childhood yrs)), timing of state-funded Kindergarten intro and timing of tax limit policies; controls for 1960 county characteristics (poverty rate, percent black, education, percent urban, population size, percent voted for Strom Thurmond in 1948 Presidential election*race (proxy for segregationist preferences)) each interacted with linear cohort trends; and controls for childhood family characteristics (parental income/education/occupation, mother's marital status at birth, birth weight, gender). The first-stage model include 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. There exists a significant first-stage.

PLACEHOLDER: Figure G1: Effect of Head Start Spending by Age (other outcomes)









Appendix H

Additional dynamic complementarity plots for family income

Appendix I

Testing for Improvement in Parent Quality due to Head Start

Regress older sibling outcomes on the Head Start spending of their younger sibling and controlling for their own Head Start spending level on all outcomes. [Can do this without the K12 spending model to keep it simple]