American Economic Journal: Economic Policy 2023, 15(1): 37–69
https://doi.org/10.1257/pol.20200660

The Effect of Early Childhood Education on Adult Criminality: Evidence from the 1960s through 1990s†

By John Anders, Andrew C. Barr, and Alexander A. Smith*

We compare the effects of early childhood education on adult criminal behavior across time periods, using administrative crime data that provide significant precision advantages over existing work. We find that improvements in early childhood education led to large (20 percent) reductions in later criminal behavior, reductions that far exceed those implied by estimates of test score gains in prior studies. While the benefits generated account for a large portion of the costs of the education provided, we find substantial relative gains from the targeting of funds to high-poverty areas and areas without existing access to subsidized care. (*JEL* H75, I21, I26, I28, I32, I38, K42)

Are criminals made or born? This fundamental question not only has important implications for our understanding of criminality but also is central to any efforts aimed at reducing the large costs that crime imposes on society ($2 trillion annually).[1] Most policies address these costs by focusing on the incapacitation or rehabilitation of criminals. Relatively little is known about the factors that influence an individual's likelihood of becoming a criminal in the first place. The concentration of crime among a small number of perpetrators (less than 6 percent of the population commit the majority of crime) provides an opportunity for policy interventions to have outsized effects if they can prevent the development of criminals.[2] In fact, some estimates suggest that preventing the development of a single career criminal could result in more than 100 fewer victims of crime each year.[3]

*Anders: Department of Economics, Trinity University (email: jpaulanders@tamu.edu); Barr: Department of Economics, Texas A&M University (email: abarr@tamu.edu); Smith: Department of Social Sciences, United States Military Academy, West Point (email: alexander.smith@westpoint.edu). C. Kirabo Jackson was coeditor for this article. We thank Jason Lindo, Mark Hoekstra, and participants at the 2020 NBER Children's Program and the 2018 Southern Economics Association for helpful comments and suggestions. We also thank Laura Hitt and the North Carolina Smart Start organization for sharing their funding data. The opinions expressed herein reflect the personal views of the authors and not those of the US Army or the Department of Defense. All errors are our own.

†Go to https://doi.org/10.1257/pol.20200660 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

[1] For context, $2 trillion is 17 percent of annual GDP (*The Costs of Crime, before the United States Senate Committee on the Judiciary*, 107th Cong. (2006) (statement of Jens Ludwig, Professor, Georgetown Public Policy Institute, Georgetown University).

[2] Farrington (2006) generates this statistic by tracking the criminal behavior of a set of boys in London. Given the higher propensity to commit crime among males, 6 percent is likely a substantial overestimate of the share of the population that commits the majority of crime.

[3] Across major crime categories, estimates suggest that a relatively small proportion of individuals (consistently less than 10 percent) account for the majority of crime. These "career criminals" commit hundreds of crimes each year (authors' calculations from Chaiken and Chaiken 1982).

One policy intervention that may influence later criminal behavior is early childhood education. Understanding its relationship to criminality is important because recent attempts to improve the quality and accessibility of early childhood education have been driven in large part by policymakers' belief that those interventions will have large impacts later in life. Positive externalities, such as reductions in adult criminal behavior, are particularly important for thinking about the social return to investments in early childhood education programs: unlike the labor market return to improvements in human capital, these benefits accrue largely to those who are not directly affected by the programs and thus may not be considered in parents' participation decisions. Because of the large social costs of criminal behavior, even modest reductions may warrant large subsidies to early childhood education. Indeed, in one often cited early childhood study, effects on crime account for 40–65 percent of the estimated benefits of the program (Heckman et al. 2010).

Yet, the limited evidence on the effect of early childhood education on later criminal behavior is mixed, imprecise, and limited largely to cohorts from the 1960s and 1970s (online Appendix Figure A1a).[4] The evidence rests entirely on evaluations of three early childhood programs: Perry Preschool, the Abecedarian Project, and Head Start. Perry Preschool, a single evaluation of a small-scale high-intensity intervention, provides some of the more compelling evidence, but even in this case, there is disagreement among researchers regarding the presence of statistically significant effects (Heckman et al. 2010; Anderson 2008). A randomized evaluation of a similar program, the Abecedarian Project, indicates no overall effect of the program on crime (Campbell et al. 2012; Anderson 2008; García, Heckman, and Ziff 2019). Furthermore, while these studies provide rigorous evidence driven by random assignment, both rely on very small samples with substantial attrition to support their conclusions.[5] Evidence on the effects of Head Start on criminal behavior is also mixed and relies on small samples, self-reported crime data, and sibling comparison approaches that may underestimate the effects of community-level program availability in the presence of spillovers (Deming 2009; Garces, Thomas, and Currie 2002). Johnson and Jackson (2018) overcome some of these limitations by leveraging the rollout of Head Start to estimate its effects on incarceration in the Panel Study of Income Dynamics (PSID), but these data have their own limitations. Furthermore, as we discuss in the next section and online Appendix B, in all of these studies, the underlying sample sizes limit the confidence we can place in the resulting estimates. For example, to identify the simple mean effect size of early childhood education on criminality in the literature, the statistical power of the Head Start studies ranges from 0.09 to 0.13. Under balanced priors on the likelihood of an effect, this implies false discovery rates of 42 to 54 percent. Finally, the early childhood context and counterfactual care options have changed dramatically from the 1960s to the present day, suggesting that even if we had a convincing sense of the role of early childhood education in preventing criminal behavior in earlier periods, the same relationship may not exist in recent years.

[4] The construction of Figure A1 is discussed in detail in online Appendix B.

[5] At adult follow-up, the Perry experiment had 123 members and the Abecedarian experiment had 111, both somewhat reduced from initial samples.

We contribute to this literature by using administrative crime data to examine the effects of two large-scale early childhood education programs operating in two different time periods. Our use of the universe of administrative criminal conviction data from North Carolina results in estimates based on the criminal behavior of hundreds of thousands (versus a hundred or thousands), with corresponding increases in precision.[6] As we discuss and support with simulation exercises in online Appendix C, this improved precision reduces the likelihood that any statistically significant estimated effects are false positives (i.e., the false discovery rate) and reduces the expected upward bias in the magnitude of reported effects. Our added statistical power allows us to be more confident that early childhood education reduces criminal behavior, that this reduction is likely considerably smaller than some prior estimates suggest, and that it is concentrated in counties with high poverty levels and limited access to alternative care options.

Our approach also substantially improves external validity over prior estimates of long-run crime effects that focus on small-scale single-site interventions and/or treatments that occurred roughly 50 years ago. It enables a direct comparison of the adult crime effects of two different interventions implemented in very different contexts while adding to the limited evidence on the effectiveness of *modern* early childhood education programs by demonstrating that these effects extend past the schooling years.[7] This is important because, as we discuss below, existing evidence suggests that short-term effects on test scores are not predictive of long-run effects on criminal behavior.

Our unique data provide two additional advantages. First, the increased statistical power enables us to explore the extent of heterogeneity in effects across areas with different levels of poverty and access to existing care. This allows us to provide a potential explanation for the pattern of results in the literature and to inform the debate between universal and more targeted early education policies. Second, the detailed information available in the data allows us to better quantify the crime benefits of the programs by directly estimating the discounted social benefits generated by reductions of various types of crimes.

Our investigation exploits the staggered introduction of the Head Start and Smart Start programs in North Carolina. The Head Start program, funded and administered through the US Department of Health and Human Services, has been an integral part of early childhood education nationwide since the late 1960s. It is the largest early childhood education program in the United States, with an annual enrollment that has grown from 400,000 during its early years to nearly a million participants today. Head Start was designed to focus on the "whole child" by providing a number of wraparound services alongside education (Ludwig and Miller 2007). The Smart Start program was implemented in North Carolina in 1993 to address concerns about

[6] Our samples are based on the behavior of 1.5 million individuals (Head Start) and 1.4 million individuals (Smart Start) versus at most 4,536 individuals in the case of prior studies (see column 4 of online Appendix Table B1 and online Appendix Figure B2 for information on the sample sizes of prior studies). Correspondingly, for example, our estimate of the effect of Head Start is more precise than a precision weighted average of the estimates from the five prior studies combined.

[7] For example, Fitzpatrick (2008, 2010) and Ladd, Muschkin, and Dodge (2014) show short-term effects on test scores and maternal labor force participation.

the preparation of children for school. The program helps parents pay for child care, improves the quality of early care and education programs, provides tools that help parents support their children, and ensures that children have access to preventative health care. The multifaceted nature of Smart Start has many similarities with the early years of the Head Start program; in particular, both program implementations allowed considerable flexibility with respect to how funding was spent. Unlike Head Start, individuals do not enroll directly in the Smart Start program. Instead, the program operates by funding a variety of programs at the community level, and critically, these funds are not explicitly targeted toward poor children the way that Head Start funds are.

To estimate the effect of early childhood education on adult crime, we leverage within-county variation in exposure and funding levels generated by the rollout of each program, along with individual-level administrative data for the universe of convicted criminals in North Carolina between 1972 and 2018. These administrative data are particularly well suited to our estimation strategy because they include each criminal's county of birth. Thus, we are able to link early childhood policy exposure to later criminal outcomes and can overcome a variety of measurement and endogeneity difficulties that likely inhibited earlier research.[8] We combine these data with birth counts to construct birth-county-by-birth-cohort conviction rates, which we then link with information on early childhood education funding and availability in each county and year.

Our estimates indicate that early childhood education reduces adult criminality. In particular, we find that both programs are responsible for reducing the conviction rate by approximately 20 percent in high-poverty areas. Head Start availability reduces by 1.3 percentage points the likelihood of a serious conviction by age 35, but only in high-poverty counties.[9] Smart Start, which operated decades after Head Start and often in different counties, generates reductions in later criminal behavior on the same order of magnitude, reducing the likelihood of a serious criminal conviction by 0.6 percentage points (12 percent) overall by age 24.

The internal validity of these estimates is supported by the nature of the introduction of both programs, which suggests that program availability and funding was likely not related to an individual's propensity for criminal behavior.[10] Consistent with this, there is little relationship between baseline county characteristics and exposure for either program, and our estimates are robust to a variety of standard checks that allow for differential trends in criminality across counties and the inclusion of controls for potential confounders.[11] The legitimacy of the identification strategy

[8] Most administrative crime datasets do not contain county of birth. Using these data would likely require researchers interested in the early childhood environment to make assumptions about the relationship between the location of arrest and earlier residence, thus introducing measurement error that would bias the effect estimates toward zero, as well as potentially introducing bias due to endogenous migration.

[9] This is perhaps not surprising given the focus of the Head Start program on poor children and the resulting concentration of funding among high-poverty counties. Head Start funding per capita is between three and four times larger in high- versus low-poverty counties (online Appendix Figure A3).

[10] For example, Smart Start's pilot partnerships were chosen so as to be representative of North Carolina's diversity and geography.

[11] In the case of Head Start, for example, the estimates are robust to the inclusion of time-varying county-level controls for the availability of other "war on poverty" programs as well as birth county trends. Furthermore, Head Start availability is unrelated to other policy changes shown to affect crime (e.g., the removal of lead from

is further bolstered by event study estimates that show no significant program "impact" in the years prior to its introduction in a given county but a reduction in criminality in the years afterward. While our data only cover crimes committed in North Carolina, we find no evidence of differential out-of-state migration as a result of Head Start availability or Smart Start funding levels in early childhood.[12] Given the modest level of migration out of one's state of birth (and the lack of evidence of any differential out-migration), we view our estimates as providing a lower bound for the overall effect on criminal behavior.

Though not a targeted program, Smart Start's reductions in crime are larger in high-poverty counties (1.2 percentage points, or 23 percent) and are larger for Blacks than Whites, perhaps as a result of differential preexisting access to high-quality care. The reductions generated by Smart Start are also larger in counties without access to Head Start, suggesting that the effectiveness of additional early education funding may diminish as funding increases. Both patterns of differential effects are consistent with the general patterns in the literature. More broadly, the concentration of effects in high-poverty and underserved areas has important implications for the growing number of cities and states considering the provision of universal pre-K. Our results suggest much smaller average effects than those implied by a number of older studies of programs implemented in very high-poverty, underserved settings, suggesting large potential gains from a more targeted approach.

A critical question for evaluating the benefits of more recent programs, where long-run outcomes cannot yet be observed, is the extent to which long-run crime effects are predictable based on earlier impacts on commonly measured short-run outcomes such as test scores. Our work underscores a recent body of evidence suggesting that effects on cognitive measures alone may not be predictive of long-run effects on criminal behavior. Under very conservative assumptions, the estimated effects of Smart Start on criminal behavior are more than seven times the effect implied by the third-grade test score estimates and the conditional correlation between test scores and crime. Effects of Head Start on educational attainment underpredict effects on criminal behavior by an even larger margin. These results suggest that early childhood education likely imparts important noncognitive skills that are not captured by test scores.

Finally, we show that the discounted benefits generated by early childhood education's later crime reduction account for a substantial portion of the costs of the education provided. In fact, these benefits pay for nearly the full cost of the programs in high-poverty counties when accounting for multiple offenses per conviction. This is especially noteworthy because later crime reduction was not the stated objective of either program, and these benefits likely accrue in large part to those not directly affected by the programs. Taken together, our results support recent efforts

gasoline, changes to compulsory schooling law ages in North Carolina, or the legalization of abortion), which occurred at the state level and generally affected different cohorts of individuals.

[12] Across a variety of approaches and subsamples, our estimates indicate a small and nonsignificant relationship between childhood Head Start availability and the likelihood of living in one's state of birth. Assuming similar patterns of criminality among North Carolina leavers and stayers, our upper-bound estimate of additional migration can explain at most 5 percent of our estimated effect. Additional analyses support Ladd, Muschkin, and Dodge's (2014) conclusion that Smart Start funding did not result in differential migration out of North Carolina. We discuss migration concerns further in Section IVC.

to expand and improve early childhood education but further point to large potential gains from targeting these efforts toward areas with the greatest need.

I. Evidence on the Origins of Criminal Behavior

Research on the developmental factors that influence the likelihood that an individual will become a criminal is limited, with many studies focusing solely on adolescence. A number of evaluations of the Moving to Opportunity project provide mixed evidence on the effect of neighborhood environment on criminal behavior, while studies of assignment to foster care do suggest that family environment plays an important role in affecting both contemporaneous and later criminal behavior (Sanbonmatsu et al. 2011; Doyle 2007, 2008).[13] Several studies have focused on the relationship between secondary education and crime, suggesting that additional years of schooling, increases in school quality, and changes in the composition of school peers can affect the likelihood of criminal behavior several years later (Lochner and Moretti 2004; Deming 2011). Because these adolescent treatments occur at an age when individuals typically first decide to engage in crime, they may directly affect the costs or benefits of crime (e.g., through direct exposure to crime or criminal peers) rather than influence an individual's development.[14]

Research on earlier periods of development is somewhat less common, yielding mixed evidence of effects. Emerging evidence suggests an important role for early childhood health and nutrition. Evaluations of the Nurse-Family Partnership program and the Food Stamp Program suggest significant effects of early health interventions on adolescent or adult criminal behavior (Olds et al. 1998; Olds, Sadler, and Kitzman 2007; Barr and Smith 2021).

Still fewer studies examine the role of early childhood education, and the evidence is mixed, imprecise, and limited largely to cohorts from the 1960s and 1970s. Online Appendix Figure A1a consolidates information on effect size and estimate precision from a comprehensive review of studies that contain estimates of the causal effect of an early childhood education program on the likelihood that an individual will become a criminal.[15] The evidence rests largely on evaluations of three early childhood programs: Perry Preschool, the Abecedarian Project, and Head Start.

Perry Preschool and the Abecedarian Project were pilot interventions that enrolled cohorts from the later 1960s and 1970s, respectively. Randomized evaluations of these somewhat resource-intensive programs provide mixed evidence. Heckman et al. (2010) find that Perry Preschool participation led to large

[13] While early evaluations of the program found mixed evidence of effects on involvement with the criminal justice system at different ages (Katz, Kling, and Liebman 2001; Kling, Ludwig, and Katz 2005; Ludwig and Kling 2007), Sanbonmatsu et al. (2011) indicate no clear pattern of significant effects on arrests or delinquent behavior. Any effects that exist appear to be a result of current neighborhood conditions rather than the neighborhood that one grew up in. Doyle (2008) finds that those on the margin of placement are two-to-three times more likely to enter the criminal justice system as adults if they are placed in foster care.

[14] Deming (2011) suggests peer effects as one explanation for the effect of school quality on criminal behavior. Bayer, Hjalmarsson, and Pozen (2009) estimate criminal peer effects more directly, showing that juvenile offenders assigned to the same facility affect each other's subsequent criminal behavior.

[15] Online Appendix B provides the construction and details of these estimates.

reductions in criminal behavior, but the Campbell et al. (2002) evaluation of the Abecedarian program indicates limited effects of the program on crime. While these studies provide rigorous evidence driven by random assignment, both rely on small sample sizes from single sites to support their conclusions. Other studies that adjust for multiple hypothesis testing suggest that neither program had statistically significant effects on adult crime for boys or girls at the 5 percent level (Anderson 2008) or that the effect is robust only for participating girls (García, Heckman, and Ziff 2019). Additionally, it is not clear that these estimates readily generalize to larger-scale programs with relatively less disadvantaged children, lower spending per pupil, and broader potential for important spillover effects (List, Momeni, and Zenou 2019).[16]

A. *The Head Start Program*

The largest early childhood education program in the United States, Head Start, began as a summer program in 1965. It quickly expanded to a year-round program in the following year. Head Start's mission was to "[provide] the children of the poor with an equal opportunity to develop their full potential" (Office of Child Development 1970, 14). To that end, it was designed to focus on the "whole child" by providing a number of wraparound services alongside education (Ludwig and Miller 2007). These additional services included providing nutritious meals and snacks, access to social workers, mental health and dental treatment, immunizations, and health screenings.

Head Start served a decidedly disadvantaged population in the early years of the program; the median family income of children enrolled in Head Start was less than half that of all families in the United States (Office of Child Development 1968). Quasi-experimental evidence focused on this period suggests that Head Start has had important long-term effects for the cohorts of children who participated. Leveraging sibling comparisons and discontinuities in grant-writing assistance and program eligibility, a number of studies have documented increased educational attainment, better health, and higher earnings (Carneiro and Ginja 2014; Deming 2009; Garces, Thomas, and Currie 2002; Ludwig and Miller 2007), even in the presence of short-term test score fade-out (Deming 2009).[17] More recent evidence indicates that the positive effects of the program persist into later ages (Thompson 2017), interact positively with school funding levels (Johnson and Jackson 2018), and even spill over into the next generation (Barr and Gibbs forthcoming).

Two of the earlier studies include criminal behavior in their investigations of the long-run effects of Head Start but yield conflicting evidence. Using the Children of the National Longitudinal Survey of Youth (CNLSY), Deming (2009) finds no effect of Head Start participation on criminal behavior. Using the PSID, Garces, Thomas, and Currie (2002) find that Head Start participation reduces later criminality, but only among Black participants. However, in their replication of this study, Miller, Shenhav, and Grosz (2019) estimate an overall positive effect of Head

[16] As we discuss further below, these types of spillovers either attenuate or elude randomized control trial (RCT) estimates.

[17] Gibbs, Ludwig, and Miller (2011) provide a more comprehensive review of the Head Start literature.

Start participation on criminal behavior and an effect on Black children that is half as large as the earlier study and not statistically significant. When they reweight this estimate to account for nonrandom selection into the identifying sample, the magnitude is halved again and is far from statistically significant. While effects on crime are not the focus of any of these papers, these estimates should be interpreted cautiously given the well-known issues with underreporting in self-reported measures of criminal behavior (Hindelang, Hirschi, and Weis 1981).[18] Moreover, these studies use family fixed effects designs, where the choice to send one child to Head Start and not their sibling may be related to characteristics of the child or to the parents' circumstances at the time, potentially biasing the estimates.

As with the small-scale RCT studies, family fixed effect evaluations of early childhood programs capture only direct effects on participants rather than broader effects on the community. Even these direct estimates may be biased downward if there are important spillovers between siblings (in the case of family fixed effects) or between peers (in the case of single-site RCTs in small communities).[19] These types of spillovers may be particularly important in the context of the noncognitive benefits provided by early childhood education (List, Momeni, and Zenou 2019). To better capture the overall community-level effects, we focus on the plausibly exogenous variation in Head Start access by county and year (see Figure 1) and use unique administrative crime data with the offender's county of birth for everyone convicted of a crime in North Carolina in the last four decades.[20] Because the Head Start program rolled out quickly and grant funds were distributed directly to local grantees, the program became available in different counties at different times. There was substantial variation in the year of adoption among counties with similar baseline characteristics, generating plausibly exogenous variation in Head Start access (Barr and Gibbs forthcoming; Thompson 2017).

This rollout strategy is also used by Johnson and Jackson (2018) in an innovative empirical test for the presence of dynamic complementarities that suggests large effects of Head Start on the likelihood of ever being incarcerated, with these effects substantially magnified by the level of K–12 spending. While not the focus of their study, these incarceration measures have their own set of limitations. Incarceration measures in the PSID rely on either retrospective self-reports of incarceration (in the 1995 crime history module) or whether a family member reports that the individual in question is unavailable to respond because they are in prison or jail. The crime measure may therefore be subject to reporting concerns similar to those in prior Head Start studies. Additionally, the family report data only indicate whether the individual was reported to be in prison or jail at the time of the interview, likely

[18] As Deming (2009, 129) notes, "previous research has found that self-reported crime data (unlike the arrest records used in the Perry Preschool study) are highly unreliable, both in the National Longitudinal Survey of Youth and in other data sources (Lochner and Moretti 2004; Kling, Ludwig, and Katz et al. 2005)."

[19] For example, a back-of-the-envelope calculation for Perry Preschool suggests that nearly all age-eligible children in poverty in Ypsilanti were part of the experimental sample.

[20] This strategy captures peer effects within the same birth cohort but may still underestimate the total effect if there are spillovers to older cohorts or to individuals born in different counties. We believe that the potential for this type of spillover is modest in our context.

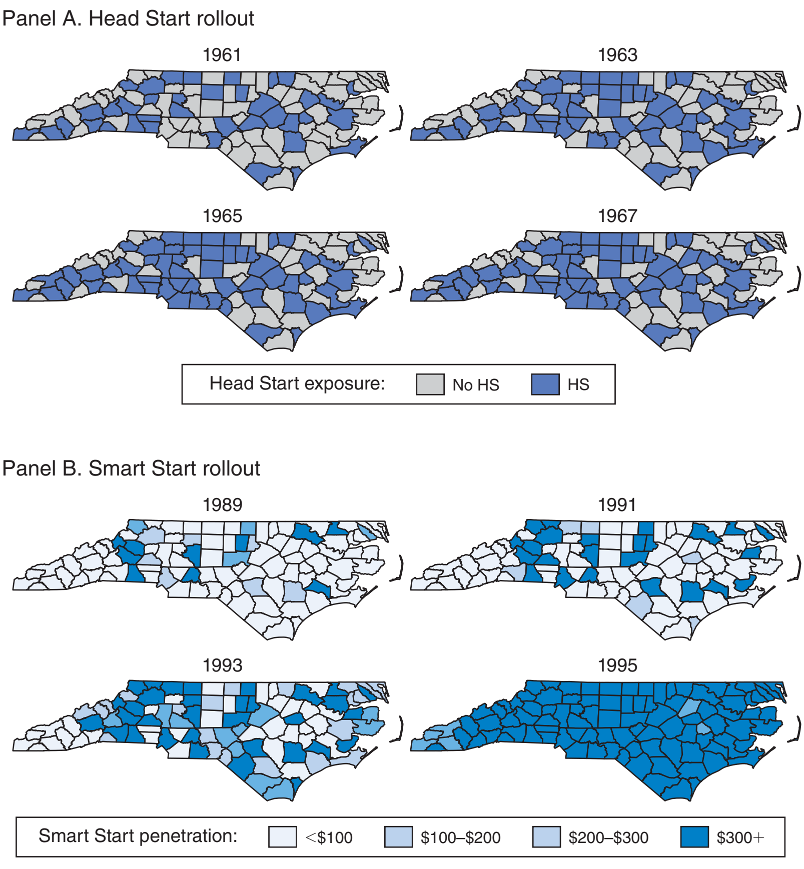

Figure 1. County-by-Birth-Cohort Early Childhood Program Rollout in North Carolina

Notes: The figure shows which birth cohorts born to which counties had Head Start and Smart Start available to them in North Carolina. Panel A: Prior to the 1961 birth cohort, no counties had Head Start available. Head Start availability is identified from county-by-year-level Head Start funding data following Barr and Gibbs (forthcoming). Head Start funding levels are obtained from Head Start Historical Records. Panel B: Prior to the 1989 birth cohort, no counties had Smart Start funding penetration. Smart Start penetration is defined following Ladd, Muschkin, and Dodge (2014) and depicted in nominal terms. Smart Start funding levels are obtained from the Smart Start organization. See the text for additional details.

missing many periods of incarceration, particularly during later years when the PSID was fielded every other year.[21]

Statistical Power for Effects on Criminal Behavior in Previous Studies.—In addition to the mixed nature of the evidence, uncertainty about the internal validity of the approaches, and questions about the effect of measurement error in self-reported criminal behavior, the underlying sample sizes in the survey data limit the extent to which the resulting estimates inform our understanding of the relationship between early childhood education and adult criminal behavior. We illustrate this at length in online

[21] This would attenuate the resulting estimates toward zero but could also bias estimates if there are different crime-age profiles across counties (since the incarceration of later cohorts may be undermeasured).

Appendix C for the studies depicted in Figure A1a. Simulation exercises using the rollout approach demonstrate the wide distribution of estimates one is likely to obtain with samples of the size available in the PSID, the 1979 National Longitudinal Survey of Youth (NLSY79), and the CNLSY, as well as the limited statistical power available to identify effects of the magnitude that we find in our North Carolina population data. This low level of power implies a surprisingly high (40 percent) false discovery rate given balanced priors on the likelihood of an effect.[22] Indeed, if we conduct our same rollout analysis in the NLSY79, we estimate a null effect of Head Start availability with positive coefficients for those subgroups more likely to be eligible.

This lack of statistical power in the existing evidence base is not a function of the empirical approach. In online Appendix Table C3, we summarize the statistical power of each study to identify the average published effect size, following the guidance provided in the meta-analysis literature.[23] For example, to identify the simple mean effect size per $1,000 in the literature, the statistical power of the Head Start studies ranges from 0.09 to 0.13. Under balanced priors on the likelihood of an effect, this implies false discovery rates of 42 to 54 percent. These false discovery concerns regarding effects on criminality are not limited to the Head Start literature; the randomized evaluations of the more intensive interventions are only modestly better powered. This does not imply that prior estimates of the effect of early childhood education on adult criminality are incorrect, but it does suggest that a better-powered study could dramatically narrow the confidence interval for these effects.[24] Our use of administrative data on the universe of criminal convictions in North Carolina results in estimates based on the criminal behavior of hundreds of thousands (versus a hundred or thousands) and corresponding increases in precision.

B. *North Carolina's Smart Start Program*

Our approach also substantially improves external validity over prior estimates of long-run crime effects that focus on small-scale single-site interventions and/or treatments that occurred roughly 50 years ago. While the initial rollout of Head Start in the 1960s and 1970s provides a compelling source of variation, it is not clear whether these early estimates can be generalized to more recent periods. Indeed, the extreme poverty, hunger, and poor early childhood health that were prevalent during this period are uncommon in recent decades, and the counterfactual early childhood education landscape has changed dramatically.[25] To assess the more general potential for early childhood education to reduce criminality, we examine the impact of the Smart Start program in North Carolina.

The Smart Start program was created in 1993 to address concerns about the school readiness of children in North Carolina. The program helps parents pay for

[22] See online Appendix C for further discussion of these points.

[23] This will be an overestimate of the true average effect size in the presence of a publication preference for statistically significant effects. Indeed, as we discuss in online Appendix C, many of these estimates do not replicate across studies, sometimes even when using the same dataset.

[24] This is not a function of the empirical approach; the family fixed effects designs have similar power limitations in these samples.

[25] For example, the enrollment rate of low-income four-year-old children has more than doubled from the 1960s to the 1990s.

child care, improves the quality of early care and education programs, provides tools that help parents support their children, and ensures that children have access to preventative health care. The multifaceted nature of this program is similar in many ways to the early years of the Head Start program. But unlike Head Start, individuals do not enroll directly in the Smart Start program. Instead, the program operates by funding a variety of initiatives at the community level. And critically, the funds are not required to be targeted toward poor children.

Like Head Start, Smart Start was rolled out to different counties at different times, resulting in significant variation in funding across birth cohorts within similar counties (Figure 1). Ladd, Muschkin, and Dodge (2014) use this variation in county–year funding to demonstrate sizable short-term effects of Smart Start funding on third grade test scores (around 5 percent of a standard deviation for the average level of funding in a county at ages 0 to 5). However, we are not aware of any studies of the longer-term effects of Smart Start. We adopt the identification strategy of Ladd, Muschkin, and Dodge (2014) to provide the first estimates of the effect of Smart Start on an adult outcome, criminal behavior.

C. *Do Cognitive Effects in Children Predict Long-Run Criminal Behavior Effects?*

A critical question for evaluating the benefits of more recent programs, where long-run outcomes cannot yet be observed, is the extent to which long-run effects are predictable based on earlier impacts on commonly measured short-run outcomes such as test scores. The nature of the relationship between cognitive skills and criminal outcomes is not well understood. Indeed, we are unaware of any evidence that a program that improves cognitive skills in childhood (as Smart Start and Head Start have) will necessarily produce reductions in adult criminal behavior. In fact, there is a variety of evidence that noncognitive skills are much more closely related to impulsive and criminal behavior than cognitive skills (i.e., test scores) and that each skill type can be influenced independently (Heckman, Stixrud, and Urzua 2006; Hill et al. 2011; Heckman et al. 2010; Jackson 2018).[26] Deming (2009) finds large effects of Head Start participation on measures of cognitive ability (0.13 SD at ages 7–10) but no effect on criminal behavior, and Heckman, Pinto, and Savelyev (2013) decompose the effects of Perry Preschool on eventual criminal behavior and show that they operate entirely through a noncognitive channel (externalizing behavior). Simply put, there is evidence of early childhood education affecting cognitive skills but not criminal behavior, as well as evidence of effects on criminal behavior but not cognitive skills.[27]

This lack of relationship between effects on cognitive measures and effects on adult criminal behavior in prior studies may not generalize to all cognitive measures, such as those used by Ladd, Muschkin, and Dodge (2014) to estimate Smart Start effects. However, Deming (2011) uses the same North Carolina test score measures

[26] Jackson (2018) shows that teachers can have effects on test scores and behaviors and that these value-added measures are only weakly correlated. This suggests that many teachers who raise test scores do not improve behaviors. Along similar lines, he shows that about 75 percent of the variation in his behavior index is unrelated to test scores.

[27] More generally, Gibbs, Ludwig, and Miller (2011) suggest that while test score fade-out is observed in studies of Head Start and Perry Preschool as well as kindergarten class size, the absence of cognitive effects in certain grades does not preclude significant long-run gains.

as Ladd, Muschkin, and Dodge (2014) to demonstrate long-run effects of school choice on criminal behavior despite the gains in school quality generating no effect on test scores, suggesting that the crime effects are not operating through the cognitive channel in this context. Indeed, even the conditional correlation between the North Carolina test scores and criminal behavior, which is likely partially driven by omitted variable bias, suggests a weak relationship.[28] If we use the conditional correlations between test scores and criminal behavior and assume a persistent effect on test scores from third through fifth grade, we can infer a conservative upper bound on the magnitude of Ladd, Muschkin, and Dodge's (2014) implied effect of Smart Start on criminal behavior of −0.088 pp, a reduction of 1.75 percent of the mean.[29] In other words, the effects of Smart Start on the cognitive channel imply an extremely modest reduction in crime.

A natural question is whether the richer set of observed human capital accumulation effects in the Head Start literature might suggest larger impacts on criminal behavior than those implied by test scores alone. Indeed, the positive effect of Head Start on educational attainment and the negative relationship between educational attainment and criminal behavior hint that effects on crime might be expected or even entirely predictable. However, a more careful effort to combine prior estimates of the effect of Head Start on educational attainment with causal estimates of the effect of education on crime suggests a different conclusion. Previously estimated effects on educational attainment imply reductions in crime of under 0.8 percent of the mean, an extremely small effect relative to the observed effects of Head Start on other margins.[30]

The different roles of cognitive and noncognitive skills in determining adult outcomes suggests that the evidence on the test score (or subsequent educational attainment) effects of modern early childhood education programs might not successfully predict these programs' impacts on criminal behavior. To better predict these longer-run behavioral effects, it is likely that researchers will need to develop better short-run behavioral measures. This is an area of future research that will be of great practical benefit to policymakers.

II. Data

Our primary data source is administrative conviction data from the state of North Carolina. We use these data, along with birth records, to calculate rates of conviction by county–year birth cohort. We use Head Start and Smart Start funding information to construct measures of program exposure at the same level. We link the measures

[28] For example, a consistent standard deviation reduction in test scores across grades 3 through 5 generates a relative odds ratio of 1.3, whereas ever being suspended in each of those grades generates an odds ratio of 3.6 (Deming 2011).

[29] We use the estimates in online Appendix Table III of Deming (2011) to calculate odds ratios, which we apply to our measures. We assume Smart Start effects operate by moving individuals in the range between −2 and −1 SD and apply the estimates from Ladd, Muschkin, and Dodge (2014) accordingly.

[30] We combine the mean effect on years of education and high school graduation across published studies of Head Start (scaled to the implied effect given the participation rates in our high-poverty counties) with the respective estimates of the effect of education on crime from Lochner and Moretti (2004).

of program exposure to the conviction rates at the county–year birth cohort level to estimate the effect of each early childhood education program on later adult crime.

A. *North Carolina Data*

We obtained data containing public information on all individuals convicted of a crime in North Carolina between 1972 and 2018 from the North Carolina Department of Public Safety.[31] The administrative data contain information on the type of crime, including the statute of the offense and whether it was a felony, as well as the name, date of birth, sex, and race of the perpetrator. An important advantage of the North Carolina data over other state criminal databases is the inclusion of county of birth for each individual. Combining information on criminals' years and counties of birth with birth counts obtained from the North Carolina Department of Health and Human Services allows us to construct conviction rates for birth cohorts of individuals born in North Carolina. For example, to generate the cohort conviction rate by age 35 for children born in county c in 1961, we divide the number of individuals born in county c in 1961 and convicted by age 35 by the total number of individuals born in county c in 1961.

For our Head Start analyses, we restrict the sample to individuals born between 1955 and 1968, allowing us to leverage the variation in Head Start availability that occurred up to and including 1972 (as Head Start availability is measured four years after birth). For our Smart Start analyses, we focus on individuals born between 1980 and 1994, which allows us to observe criminal convictions through age 24 for all cohorts in the sample (we observe convictions through 2018).[32]

Summary statistics are contained in Table 1. Slightly less than 5 percent of individuals born between 1955 and 1968 were convicted of a serious crime by age 35; likewise, slightly more than 5 percent of individuals born between 1980 and 1994 were convicted of a serious crime by age 24. Our definition of a serious crime is based on the Federal Bureau of Investigation's (FBI's) Part I offenses.[33]

While the data contain the universe of individuals convicted of a crime in North arolina during this time period and allow us to link these individuals to their counties of birth, they are limited in that they do not allow us to observe convictions for individuals who are born in North Carolina and then leave the state. While most likely criminals remain in their state of birth (and the likelihood of criminal behavior is lower for those who leave), this may be a concern for the interpretation of our estimates if Head Start availability or Smart Start spending affects out-of-state migration, especially if

[31] For the cohorts in our sample, this includes offenders aged 16 and older. Until a law change in 2017 (not fully implemented until 2019), 16-year-olds were charged as adults in North Carolina. No cohorts in our sample were affected by this law change.

[32] We find similar effects when we focus our estimates on convictions through older ages, which necessitates a reduction in our identifying variation (e.g., through age 25 (using cohorts through 1993) and through age 26 (using cohorts through 1992)). See online Appendix Table A10. We focus on convictions by age 24 because it provides a balance between variation in Smart Start funding exposure across birth cohorts and our capacity to observe these cohorts at ages with the highest rates of criminal behavior.

[33] We largely follow the convention of the FBI's Uniform Crime Reporting (UCR) Statistics for Part I offenses. Violent crimes are defined as offenses containing the words "murder," "assault," or "robbery." Property crimes are defined as offenses containing the words "burglary" or "larceny."

Table 1—Descriptive Statistics

	All	High poverty	Low poverty
Panel A. Head Start sample			
First cohort with Head Start	1962.3	1962.3	1962.3
Head Start funding (2015$ per 4-year-old)	893.8	2,061.1	605.7
Criminal conviction	0.0469	0.0462	0.0471
Observations (cells)	882	308	574
Individuals represented	1,487,225	444,848	1,042,377
Panel B. Smart Start sample			
First calendar year of Smart Start	1995.5	1996.7	1995.0
Smart Start penetration (2015$)	818.9	750.0	838.3
Criminal conviction	0.0516	0.0512	0.0517
Observations (cells)	1,500	750	750
Individuals represented	1,407,042	666,073	740,969

Notes: Panel A contains summary statistics of crime outcome variables for the sample of birth cohorts born from 1955 to 1968, which is used in the Head Start analysis. Each observation (cell) is at the county–birth cohort level. Head Start availability and funding are reported in the first two rows. Funding levels are given in 2015 dollars and averaged over exposed county cohorts only, so that only nonzero values are included. Criminal conviction by adulthood is the fraction of individuals in a given birth-county-and-birth-year cohort that are later convicted of a UCR Part 1 crime in North Carolina by age 35. UCR Part 1 crimes include violent crimes (those in which the description of the offense contains the words "murder," "assault," or "robbery" (rape not being included)) and property crimes (those in which the description of the offense contains the words "burglary" or "larceny"). All variables are further broken down by county-level poverty status. Those counties for which the poverty rate in 1960 was above the median in North Carolina (40.2 percent poverty) are called "High poverty," while those below the median are called "Low poverty." The sample is restricted to counties that ever received Head Start between 1965 and 1976. Panel B contains summary statistics of crime outcome variables for the sample of birth cohorts born from 1980 to 1994, which is used in the Smart Start analysis. Each observation (cell) is at the county–birth cohort level. Smart Start availability and funding penetration are reported in the first two rows of panel B. Penetration measures are defined following Ladd, Muschkin, and Dodge (2014), reported in 2015 dollars, and averaged over exposed county cohorts only, so that only nonzero values are included. Criminal conviction by adulthood is the fraction of individuals in a given birth-county-and-birth-year cohort that are later convicted of a UCR Part 1 crime in North Carolina by age 24. All variables are further broken down by county-level poverty status. Those counties whose poverty rate in 1980 was above the median in North Carolina (17.3 percent poverty) are called "High poverty," while those below the median are called "Low poverty."

Sources: North Carolina Department of Corrections, Head Start historical records, and Smart Start records

it leads individuals with a higher propensity for criminal behavior to leave the state.[34] We return to this concern below in Section IVC, providing evidence that program availability and funding do not appear to influence migration rates.[35,36]

[34] Roughly 70 percent of individuals born in North Carolina during this period reside there between the ages of 18 and 35. This share is even higher (roughly 80 percent) for those with the highest rates of criminal behavior (between ages 18 and 24, non-White, or with less than a high school degree). Authors' calculations using public census and American Community Survey (ACS) data.

[35] Specifically, for Head Start, we explore the relationship between measures of childhood Head Start availability (at the state-of-birth-by-birth-cohort level) and the likelihood of living in one's state of birth. Across a variety of approaches and subsamples, our estimates indicate a small and nonsignificant relationship between childhood Head Start availability and the likelihood of living in one's state of birth. Similar analyses support Ladd, Muschkin, and Dodge's (2014) conclusion that Smart Start funding did not result in differential migration out of North Carolina. We address concerns about differential migration further in Section IVC.

[36] We may also be missing individuals with one-time nonviolent convictions (at any age) or one-time drug convictions (under age 22) who hired a lawyer and had the record expunged. However, these possible missing observations are a threat to identification only if they occur differentially based on early childhood educational status.

B. *Head Start Availability Measure*

We follow Barr and Gibbs's (forthcoming) definition of Head Start availability, which relies on county–year data from the Community Action Programs and Federal Outlay System files obtained from the National Archives and Records Administration.[37] We construct an availability indicator equal to one if a county had Head Start expenditures per four-year-old above the national tenth percentile.[38,39] We do not otherwise leverage data on appropriated dollar amounts due to concerns about the accuracy of the recorded funding amounts in the early years of the Head Start program as well as the potential endogeneity of funding levels.

Smart Start Funding Exposure Measure.—We follow Ladd, Muschkin, and Dodge's (2014) definition of Smart Start funding penetration, which relies on county–year data on Smart Start funding obtained from the budget and grants compliance officer at Smart Start. For each county–year, we compute the annual funding amount per child aged zero to five. We then sum this county–year funding penetration by cohort, generating the total amount of funding exposure that a given cohort received from ages zero to five. Thus, our Smart Start funding measure tracks the cumulative exposure to Smart Start funding that a given county birth cohort received.

III. Estimation of Program Availability Effects

To estimate the effect of early childhood education on adult crime, we exploit within-county variation in exposure generated by the rollout of each program. For Head Start, we leverage within-county variation in the availability of Head Start generated by the initial rollout of the program in the 1960s. For example, we utilize the fact that eligible 4-year-olds in 55 out of North Carolina's 100 counties had access to Head Start in 1968, while no 4-year-olds had access to Head Start prior to 1965 (panel A of Figure 1). For Smart Start, we leverage the funding expansion from 0 counties in 1992 to all 100 counties by 2000, with substantial differences in per capita funding levels (panel B of Figure 1).

For both interventions, we estimate the following specification:

$$C_{ct} = \beta EC_{ct} + \gamma(X_c \times t) + \alpha_c + \sigma_t + \epsilon_{ct}, \tag{1}$$

[37] See Barr and Gibbs's (forthcoming) online Data Appendix for details.

[38] The measure for a cohort is based on funding four and five years after birth, when the cohort was three-to-five years old, following Barr and Gibbs (forthcoming). The modal age of participation in the summer program was five, but the modal age of participation in the full-year program was four between 1966 and 1968. Throughout this period, 86 to 92 percent of participants were ages 3 to 5. We use the national tenth percentile threshold for consistency with Barr and Gibbs (forthcoming), who find that using this threshold better predicts Head Start take-up, but neither the values of the availability indicator (in North Carolina) nor the main results are sensitive to moving this threshold. Indeed, in North Carolina nearly all positive cohort-by-county Head Start spending exceeds the tenth percentile. As a result, there is little concern regarding misclassification, and the point estimates are insensitive to using thresholds between the first and thirtieth percentiles.

[39] Kindergarten was not widely available in North Carolina in the 1960s. Even by 1973, only 16 percent of children were enrolled in kindergarten, and it wasn't until 1978 that public kindergarten became available for all age-eligible children.

where C_{ct} is the conviction rate for those born in county c in year t, EC_{ct} is a measure of county–year birth cohort exposure to the early childhood policy, and α_c and σ_t are birth county and birth year fixed effects. In robustness checks, we include $X_c \times t$, which are controls for baseline birth county characteristics interacted with a time trend, to account for any differential trends by county characteristics.[40] For example, this would account for differences in crime trends between more and less rural counties.[41] Standard errors are clustered at the county-of-birth level.

The key identifying assumption is that exposure to each early childhood education policy is, conditional on birth county and birth year fixed effects, unrelated to the propensity to be convicted of a serious crime for some reason other than the policy. For example, if Head Start were rolled out at the same time that other programs targeted at children were adopted, we could be mistakenly attributing the effects of this other program to Head Start.

The exogeneity of the national-level Head Start rollout has been supported at length in related work, with multiple studies referencing the quick and haphazard nature of initial Head Start grant making as well as demonstrating a limited association between baseline county observables and the presence and timing of Head Start adoption once a county's poverty level is accounted for (see, for example, Barr and Gibbs forthcoming; Johnson and Jackson 2018; and Thompson 2017).[42] We further bolster the exogeneity of the Head Start rollout in North Carolina with an event study (Figure 2), discussed in detail in Section IVA, that finds no "effects" on the likelihood of conviction in the years of Head Start's initial adoption in a county and a sharp jump immediately following the program's introduction.

Just as in the case of Head Start, the nature of Smart Start's introduction suggests that program exposure was, conditional on birth county and birth year fixed effects, unrelated to an individual's propensity for criminal behavior. The pilot partnerships were chosen to be representative of North Carolina's diversity and geography, with one county selected from each congressional district. As noted by Ladd, Muschkin, and Dodge (2014), experts used county-level data on poverty rates, tax base, and number of young children in need to rank North Carolina's 100 counties into 1 of 4 resource bands: high, medium high, medium low, and low. As required by the Smart Start legislation, the experts selected three counties within each resource band to receive initial funding. The program quickly expanded, reaching all counties by the 1998–1999 school year. Consistent with the initial intent to provide balanced exposure to the program, there is little relationship between baseline county observables and the timing or magnitude of Smart Start funding (online Appendix Figures A11

[40] For Head Start, the baseline characteristics are from 1960, whereas for Smart Start they are from 1980.

[41] Following Hoynes, Schanzenbach, and Almond (2016), the 1960 county characteristics include the percent of people living in families with less than $(1960)3,000, the percent living in urban areas, the percent Black, the percent under 5 years old, the percent over 65 years old, the percent of land in farming, and the percent of employment in agriculture. Following Ladd, Muschkin, and Dodge (2014), the 1980 county characteristics include the share of births to Black mothers, to Hispanic mothers, and to low-education mothers; the share of the population using food stamps; the total number of births; the total population; and the median family income.

[42] We present analogous analyses specific to North Carolina in online Appendix Tables A1 and A2. We find no statistically significant relationship between county characteristics in 1960 and the timing of Head Start availability within North Carolina. Online Appendix Figure A10 presents this relationship graphically and similarly suggests little relationship between baseline county characteristics and the timing of introduction. Online Appendix Table A1 indicates that more populous counties were more likely to get the program at all during this time period.

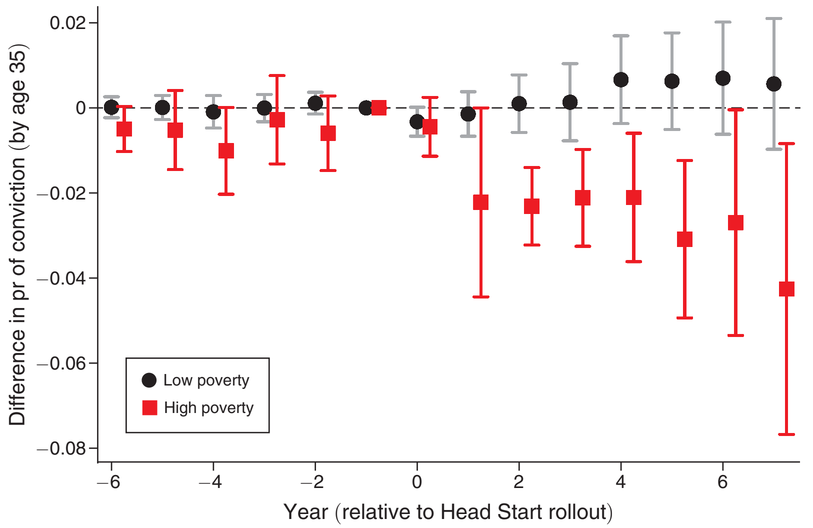

Figure 2. Event Study of Head Start's Impact on Criminal Conviction

Notes: The figure shows the coefficient estimates and 95 percent confidence interval from estimating equation (2) separately for high- and low-poverty counties. Observations are at the birth-county-by-birth-year level and are weighted by the number of births in each county in 1955. The dependent variable is the fraction of individuals in a given birth-county-and-birth-year cohort that are later convicted of a UCR Part 1 crime in North Carolina by age 35. UCR Part 1 crimes include violent crimes (those in which the description of the offense contains the words "murder," "assault," or "robbery" (rape not being included)) and property crimes (those in which the description of the offense contains the words "burglary" or "larceny"). All specifications include birth county and birth cohort fixed effects as well as 1960 county characteristics interacted with a time trend in birth cohort. The 1960 county characteristics include percent of land in farming, percent of people living in families with less than $3,000, percent of population in urban area, percent Black, percent less than age 5, percent greater than age 65, and percent of employment in agriculture. Those counties whose poverty rate in 1960 was above the median in North Carolina (40.2 percent poverty) are called "High poverty," while those below the median are called "Low poverty." The sample is restricted to counties that ever received Head Start between 1965 and 1976. The sample is further restricted to cohorts who were born between 1955 and 1968.

and A12, online Appendix Table A1). We provide additional evidence for the exogeneity of the Smart Start rollout in North Carolina with another event study (Figure 3), discussed in detail in Section IVA, that finds no "effects" on the likelihood of conviction prior to the initial funding of Smart Start in a county and a phasing in of effects following initial funding that is consistent with the targeting of funds to a broader age range.[43]

While both programs experienced a similarly rapid rollout across the state, they differed substantially in how funding was distributed to participating counties. As online Appendix Figure A2 shows, Head Start funding was concentrated in higher-poverty counties, while Smart Start funding was distributed relatively evenly across counties. This difference reflects the different mandates of the two programs. As part of the "war on poverty," the Head Start program focused on children from families at or below the federal poverty line; indeed, at least 90 percent of Head Start participants at each site were required to be from families below the poverty line. Consistent with this, Head Start funding per four-year-old is three-to-four times as high in high-poverty counties

[43] We discuss the expected and observed pattern of results further below.

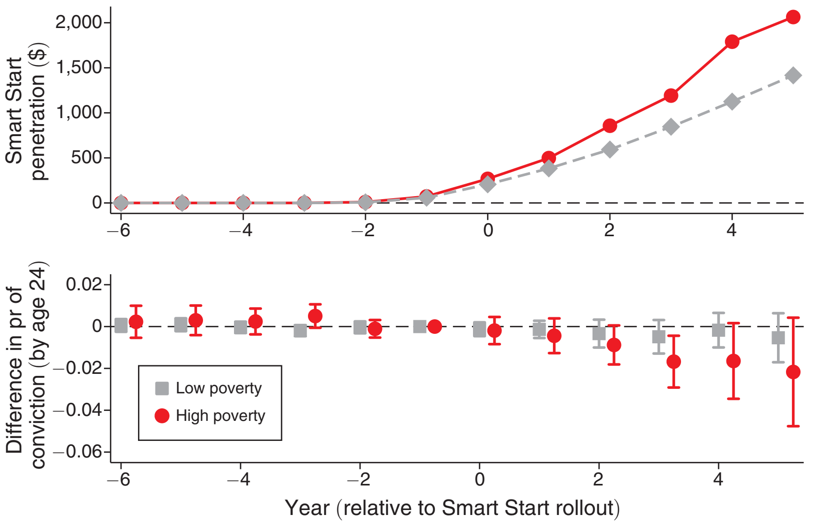

Figure 3. Event Study of Smart Start's Impact on Criminal Conviction

Notes: In the top panel, figure shows trends in Smart Start funding penetration in nominal dollars defined following Ladd, Muschkin, and Dodge (2014), separately for high- and low-poverty counties. In the bottom panel, figure shows the coefficient estimates and 95 percent confidence interval from estimating equation (2) separately for high- and low-poverty counties. Observations are at the birth-county-by-birth-year level and are weighted by the number of births in each county in 1980. The dependent variable is the fraction of individuals in a given birth-county-and-birth-year cohort that are later convicted of a UCR Part 1 crime in North Carolina by age 24. UCR Part 1 crimes include violent crimes (those in which the description of the offense contains the words "murder," "assault," or "robbery" (rape not being included)) and property crimes (those in which the description of the offense contains the words "burglary" or "larceny"). All specifications include birth county and birth cohort fixed effects. Binary Smart Start availability, the independent variable of interest, is defined as Smart Start penetration above the twenty-fifth percentile of penetration. Those counties whose poverty rate in 1980 was above the median in North Carolina (17.3 percent poverty) are called "High poverty," while those below the median are called "Low poverty." The sample is restricted to cohorts who were born between 1980 and 1994.

(online Appendix Figure A3). By contrast, poverty was not an explicit focus of the Smart Start program. This yielded a more equal distribution of funds across high- and low-poverty counties. Nevertheless, differential effects of Smart Start by poverty status may still occur if families in poverty have fewer outside options for high-quality early childhood education. Accordingly, for both policies we conduct much of our analyses separately for high- and low-poverty counties, splitting counties at the median county poverty level for North Carolina in the baseline year for each program (i.e., 1960 for Head Start and 1980 for Smart Start).[44]

IV. Main Results

Our primary interest is in the coefficient β in equation (1), which represents the effect of early childhood education availability or funding on adult crime.[45] We find

[44] It is worth noting, however, that poverty rates improved dramatically across the state between the 1960s and 1970s and between the 1980s and 1990s.

[45] We observe county of birth and not county of residence throughout early childhood. If individuals move from their county of birth before age five, this will introduce some classification error into our exposure measures. While prior evidence presented here and elsewhere suggests that this natural migration is unlikely to bias our estimates, the

Table 2—Effect of Early Childhood Education on Criminal Conviction

	All (1)	High poverty (2)	Low poverty (3)
Panel A. Head Start			
Head Start availability	−0.0017	−0.0128	0.0026
	(0.0031)	(0.0058)	(0.0033)
Observations	882	308	574
Mean	0.0469	0.0462	0.0471
Panel B. Smart Start			
Smart Start ($1,000s)	−0.0064	−0.0118	−0.0030
	(0.0029)	(0.0051)	(0.0035)
Observations	1,500	750	750
Mean	0.0516	0.0512	0.0517

Note: Each cell reports a separate OLS regression with standard errors clustered at the birth county level and reported in parentheses. All specifications include birth county and birth cohort fixed effects. Panel A reports results using the Head Start sample. Observations are at the birth-county-by-birth-year level and are weighted by the number of births in each county in 1955. The dependent variable is the fraction of individuals in a given birth-county-and-birth-year cohort that are later convicted of a UCR Part 1 crime in North Carolina by age 35. The reported variable of interest is an indicator for whether Head Start was available to a given county birth cohort. Panel B reports results using the Smart Start sample. Observations are at the birth-county-by-birth-year level and are weighted by the number of births in each county in 1980. The dependent variable is the fraction of individuals in a given birth-county-and-birth-year cohort that are later convicted of a UCR Part 1 crime in North Carolina by age 24. The reported variable of interest is a measure of Smart Start funding penetration for a given county birth cohort, constructed following Ladd, Muschkin, and Dodge (2014). See the notes to Table 1 for additional sample restrictions and definitions.

large effects on the likelihood of an adult criminal conviction for individuals born in high-poverty counties, with effect sizes around 20 percent (Table 2).[46]

Our main Head Start estimates indicate that program availability generated a 1.3 percentage point reduction in the likelihood of a criminal conviction by age 35 for high-poverty counties but no measurable effect for low-poverty counties (top row of Table 2).[47] While our main estimates are identified using the set of counties that ever received Head Start (by 1976), the results are similar when including the counties that did not receive Head Start during this period (online Appendix Table A3).[48] Our reliance on later-treated counties to serve as controls for the earlier-treated counties and vice versa has the potential to generate bias in the resulting estimate if the treatment causes a change in trends (for example, if the effect of treatment grows substantially over time) (Goodman-Bacon 2018). Our dynamic estimates (presented

estimated effect of availability in one's county of birth will be attenuated somewhat from the effect of availability in one's county of residence. Given rates of moving during the late 1960s between birth and age 5 of around 0.15, the estimates should be scaled up by 17 to 18 percent for Head Start. Overall lower rates of moving in the 1990s suggest that this is an upper bound for the scaling of the Smart Start estimates.

[46] For Smart Start, the effect size is for average funding exposure, which is slightly less than $1,000 for cohorts born after Smart Start enters a county (these cohorts are affected throughout ages 0 to 5).

[47] Figure A4 presents coefficient estimates for the same specification by poverty quintiles. The most dramatic effect occurs in counties in the highest poverty quintile. We have also estimated specifications that interact the continuous poverty rate with an indicator for Head Start availability (online Appendix Table A4). Using this approach, we estimate that the reduction in crime rate due to Head Start availability is 0.2 percentage points larger for each 10 percentage point increase in the poverty rate. Consistent with our prior estimates, these estimates suggest that the effect of Head Start ranged from 0.0 percentage points in the county with the lowest poverty rate (23 percent) to roughly 1.4 percentage points in the county with the highest (74 percent).

[48] Online Appendix Figure A7 presents the corresponding event study, which similarly suggests a reduction in criminal behavior following Head Start availability.

in Figure 2) suggest at most a modest increase in treatment effects over time, implying that the magnitude of our overall DD estimate may be slightly attenuated. There may still be bias in the identification of our treatment effects if there is heterogeneity in the dynamic effects of treatment across counties that adopt Head Start at different times. In online Appendix E, we use techniques in Sun (2020) and de Chaisemartin, D'Haultfoeuille, and Guyonvarch (2018) to diagnose and directly address the potential for this type of "contamination bias" to influence our estimates. We find little evidence to suggest that contamination bias is a concern in our setting. Consistent with this, estimates that allow for dynamic heterogeneity are similar to our main estimates; if anything, they suggest that our standard estimates are modestly attenuated.

Because Smart Start funding is not targeted at a particular age group, we follow Ladd, Muschkin, and Dodge (2014) in estimating the effect of the average funding provided per zero-to-five-year-old over the first five years of life. In the bottom row of Table 2, we report estimates per $1,000, which is roughly equal to the average funding exposure for cohorts born after Smart Start enters a county during our sample period.[49] Our main estimates suggest a 0.6 percentage point (12 percent) reduction in the likelihood of a serious criminal conviction by age 24 from an additional $1,000 in funding exposure (Table 2).[50] Due to the timing of the Smart Start variation, we cannot present effects on conviction by age 35. As with Head Start, the estimated effects are larger in high-poverty counties, a 1.2 percentage point (23 percent) reduction in the likelihood of a serious criminal conviction.[51] While Smart Start did not explicitly target children in poverty, this pattern of results is perhaps explained by the lower-quality counterfactual care options available to low-income families or in high-poverty areas. This would conform with recent evidence that suggests larger short-term effects of early childhood programming for those who would otherwise have relied on parental or relative care (Kline and Walters 2016).

For both programs, the estimated effects in high-poverty counties appear to increase with age by which conviction is measured. Online Appendix Table A9 shows our Head Start estimates for conviction by ages 24, 30, and 35, while online Appendix Table A10 shows Smart Start results for conviction by ages 24, 25, and 26 (older ages are not feasible given the timing of the Smart Start variation). While the estimate of Head Start's effect on convictions by age 24 is not statistically significant, the percentage reduction in convictions remains similar to Smart Start for the same age. Furthermore, the corresponding event study (online Appendix Figure A5) strongly suggests a reduction in the likelihood of conviction by age 24 following Head Start availability in a county.

[49] These cohorts are fully treated, unlike those born one-to-four years before Smart Start entry, which experience partial treatment.

[50] For completeness, we have included online Appendix Tables A5 and A6, which contain estimates from a binary specification (the analogs of Tables 2 and 3). Because (i) this specification averages effects across age of exposure and (ii) the resulting DD estimate is downward biased in the presence of treatment effects that strengthen over time, the estimates are attenuated somewhat from those implied by the baseline exposure measure.

[51] As with Head Start, we have estimated specifications that interact the continuous poverty rate with our Smart Start funding measure (online Appendix Table A7). Using this approach, we estimate that the reduction in crime rate due to $1,000 of Smart Start funding is 0.5 percentage points larger for each 10 percentage point increase in the poverty rate.

Table 3—Effect of Smart Start Funding on Criminal Conviction, by Presence of Head Start

	All			High poverty		
	(1)	(2)	(3)	(4)	(5)	(6)
Smart Start ($1,000s)	−0.0064	−0.0146	−0.0040	−0.0118	−0.0156	−0.0015
	(0.0029)	(0.0050)	(0.0032)	(0.0051)	(0.0055)	(0.0033)
Observations	1,500	555	945	750	435	315
Mean	0.0516	0.0528	0.0513	0.0512	0.0532	0.0494
Head Start	All	No Head Start	Head Start	All	No Head Start	Head Start

Notes: Each column reports a separate OLS regression with standard errors clustered at the birth county level and reported in parentheses. All specifications include birth county and birth cohort fixed effects. Observations are at the birth-county-by-birth-year level and are weighted by the number of births in each county in 1980. The dependent variable is the fraction of individuals in a given birth-county-and-birth-year cohort that are later convicted of either UCR Part 1 property crimes or Part 1 violent crimes in North Carolina by age 24. The reported variable of interest is a measure of Smart Start funding penetration for a given county birth cohort, constructed following Ladd, Muschkin, and Dodge (2014). See the notes to Table 1 for additional sample restrictions and definitions. Columns 2 and 5 further restrict the sample to counties without a Head Start program by 1980, while columns 3 and 6 restrict to counties with a Head Start program by 1980.

While the rollout of the two programs occurred two-to-three decades apart, Head Start continued to operate in many North Carolina counties during the rollout of Smart Start. This provides an opportunity to investigate whether early childhood education funding yields increasing or diminishing returns. In Table 3, we examine whether Smart Start funding was more or less effective in the presence of Head Start. The estimates indicate that the effects of Smart Start are largest in counties that were not previously served by Head Start (columns 2 and 5). In contrast, in counties previously served by Head Start, there were small effects of Smart Start that are not significantly different from zero (columns 3 and 6), underscoring the importance of targeting funds to areas with the greatest need. Online Appendix Table A8 shows estimates of the full interaction of Smart Start funding and the presence of Head Start, illustrating that these differences are statistically different and underscoring the diminishing returns to early childhood education investments. In addition to confirming the results in Table 3, the full interaction results demonstrate that among both high- and low-poverty counties, counties without Head Start experienced larger effects of Smart Start funding than those with Head Start.

Our results appear consistent with the general pattern of effects in the literature, though it is difficult to draw strong conclusions from a small set of studies that identify heterogeneous parameters of different programs. For example, the fact that the treatment effects on criminal behavior are larger for Perry than Abecedarian could be explained by differences in the affluence of the local communities and the availability of counterfactual care options. The participants in Perry were drawn from Ypsilanti, a town with high levels of poverty and limited access to alternative care options. In contrast, the participants in Abecedarian were from Chapel Hill, a more affluent town with extensive alternative access to care, including Head Start. Indeed, 75 percent of control group members in Abecedarian attended alternative center-based care. While it is more difficult to compare across the heterogeneous parameters estimated in the Head Start literature, the general pattern of larger effects in earlier years (when there were greater levels of disadvantage and fewer alternative care options) is consistent with this explanation.

A. *Dynamics and Robustness*

To understand the dynamics of how the program may have affected adult criminal outcomes and to test for pretrends that may confound our baseline specification, we also present estimates from event study specifications. For example, for Head Start, we center counties around the first year that the program is available and estimate the effect of leads and lags of program availability.[52]

Figure 2 indicates a flat trend in cohort conviction rates before Head Start rollout for both high- and low-poverty counties.[53] This provides evidence that our difference-in-difference (DD) estimates are not capturing differential preexisting trends in the years prior to a county's rollout of Head Start.[54] For cohorts exposed to Head Start, we see significant decreases in the conviction rate for the high-poverty counties but continue to see no evidence of changes in the low-poverty counties.[55] In the high-poverty counties, the estimates of crime reduction appear to grow somewhat as the program persists in a county. In particular, the impact of Head Start availability in the first year of the program is substantially smaller than in subsequent years. This may be due to centers improving (or increasing the size of) their Head Start programs during the first years of operation or as a result of peer effects.[56] Funding does appear to increase somewhat during the early years of program operation, consistent with program growth (online Appendix Figure A3).

The targeting of the Smart Start intervention to a broad age range does not lend itself to a standard event study, because cohorts experience different intensities of treatment based on their age relative to the maximum targeted age (i.e., five) when the funds become available in a county. This is illustrated in the top panel of Figure 3, which demonstrates that our treatment measure rises for five-to-six years

[52] We estimate the following specification separately for counties above and below the median poverty rate:

$$C_{ct} = \sum_{\tau=-6}^{7+} \beta_\tau \mathbf{1}\left[t = T_c + \tau\right] + \alpha_c + \alpha_t + \gamma\left(X_{c,60} \times t\right) + \epsilon_{ct}. \tag{2}$$

We are primarily interested in the coefficients on the indicators, $\mathbf{1}\left[t = T_c + \tau\right]$, each of which indicates how many years cohort t in county c is removed from the first cohort in county c exposed to Head Start, T_c. The distant event times are binned (i.e., 7+), and the corresponding estimates are not presented in the figure. We use the first year in which the treatment measure exceeds the national tenth percentile (i.e., $22 per child per year) as the year of adoption within a county. Nearly all of the positive Head Start funding amounts in North Carolina exceed the national tenth percentile, so our results would be almost identical regardless of whether we set our Head Start availability variable equal to one for county–birth year cohorts that exceeded the first, fifth, tenth, …, thirtieth national percentile.

[53] While the estimates in Figure 2 includes linear trends interacted with 1960 county characteristics, online Appendix Figure A6 shows that the results are similar without the inclusion of these linear trends.

[54] Two of the preperiod estimates for high-poverty counties suggest small but borderline significant negative effects before the availability of Head Start in a county. This is driven by the choice of omitted time period ($t = -1$), which has slightly higher conviction rates than the other preperiod cohorts. If anything, the pattern of preperiod effects suggests that there may be a very slight upward trend in criminality across cohorts prior to the first cohort exposed to Head Start (suggesting an attenuation of our estimated effect), although we cannot distinguish this trend from zero.

[55] This figure also addresses concerns that there were subsequent changes in a county that affected crime rates, such as changes to its criminal justice system, that are correlated with but not caused by the timing of a county's Head Start adoption. For such a correlation to produce our event study results, the policy change would have to precisely target only cohorts exposed to Head Start availability and have no effect on cohorts born just a couple years earlier.

[56] If peer effects are an important factor in criminal behavior, we would expect smaller effects of the program in the first year as compared to subsequent years when older peers would have also experienced the program.

following adoption. The bottom panel of Figure 3 provides event study estimates of the effect of Smart Start availability on the likelihood of a criminal conviction. As with Head Start, our baseline dynamic estimates (bottom panel of Figure 3) indicate a flat trend in cohort conviction rates before Smart Start funding arrived in a county. This provides evidence that our estimates are not capturing differential preexisting trends in the years prior to a county's adoption of Smart Start. Consistent with the phase in of our treatment measure, we observe a phase-in of effects on the likelihood of a criminal conviction. The figure also shows that higher-poverty counties enjoyed larger decreases in cohort conviction rates than low-poverty counties despite Smart Start funding being similar across county poverty levels.

The main estimates of the effects of both programs using equation (1) are robust to the inclusion of pretreatment county characteristics interacted with time trends (see column 2 of online Appendix Table A11). These controls limit concerns that counties with different characteristics have differing trends in the likelihood of criminal behavior across cohorts that are correlated with the timing of program rollout. In the case of Smart Start, the results are also robust to the inclusion of time-varying controls at the county-by-birth-cohort level, which addresses concerns about changes in the composition or wealth of birth cohorts over time that could be driving the observed result (online Appendix Table A12). Similarly, the Head Start estimates are robust to the inclusion of covariates indicating availability of other "war on poverty" programs, such as Food Stamps, Medicaid, Community Health Centers, etc. (online Appendix Table A13).[57,58] Finally, in online Appendix Table A2, we explore whether the timing of adoption is correlated with either the predicted level of or growth in crime for cohorts surrounding the rollout of each program. For both programs and samples, we construct predicted crime measures by regressing each county's baseline crime rate and the sample-period growth in crime on the county's characteristics in the base year. We then regress the predicted crime measures against our timing and funding measures. For Smart Start, we find no evidence of earlier program availability or greater funding in North Carolina counties that were predisposed to lower crime (or crime growth) based on their characteristics in the base year. In the case of Head Start, while the coefficients on funding suggest that program rollout was not endogenously determined by conviction levels and trends, the coefficients on timing suggest that early-adopting counties experienced somewhat higher baseline conviction levels and predicted growth in convictions over our sample cohorts. While these associations are small, if anything, they suggest that our estimates may be *slightly* biased downward, since the baseline characteristics of early-adopting counties predicted a growth in convictions over this time period, while we still estimate that program availability decreased convictions.

[57] We also test directly for relationships between these potential confounders and our measure of Head Start availability. Consistent with the limited effect of the "war on poverty" controls on our estimates, we find no significant relationships between funding for various "war on poverty" programs and Head Start availability (online Appendix Table A14)

[58] While the same time-varying county characteristics that we employ for our Smart Start robustness checks are not available for Head Start cohorts, we find no relationship between Head Start availability and measures of infant mortality. This suggests that the relationship between Head Start availability and later criminal behavior is not driven by broader improvements in infant health or medical treatment unrelated to Head Start.

While our baseline inference relies on standard errors clustered at the county of birth, we have also explored the robustness of our *p*-values to an even more conservative approach: randomization inference. Under this procedure, we randomly assign the rollout year or funding level of Head Start and Smart Start in each county and estimate our baseline specification. The distribution of these estimates over 1,000 iterations is contained in online Appendix Figures A8 and A9. As can be seen in the figures, the estimates we observe in our baseline results are quite unlikely under random assignment. The implied two-tailed "*p*-values" that we obtain from this randomization inference approach are similar to those obtained using our baseline approach with standard errors clustered on the county of birth.[59]

B. *Heterogeneity of Effects*

The most natural channel through which early childhood education may affect crime is by raising the return to work and thereby increasing the opportunity cost of participating in criminal behavior. This channel is supported by the observed increases in human capital demonstrated in recent evaluations of modern early childhood programs (for example, Fitzpatrick 2008 or Ladd, Muschkin, and Dodge 2014), although the existing evidence suggests that improvements in test scores may not translate into crime reduction in the absence of improvements in noncognitive skills. Early childhood education may also affect the financial or psychological benefits of criminal behavior, or it may alter preferences in other ways, either by directly influencing child development or by influencing parenting, perhaps via effects on parental labor supply (Fitzpatrick 2010).

Heterogeneity in effects across crime types may provide a hint at how these early childhood programs are affecting an individual's adult propensity to commit crime. For example, changes in the opportunity cost of crime may be more likely to affect property offenses, whereas effects on child development may be more likely to affect violent offenses. In online Appendix Table A15, we explore whether the effects of early childhood education on serious convictions differ by crime type. The coefficients for violent crimes are not statistically distinguishable from those for property crimes for Head Start or Smart Start. The point estimates indicate a 0.25 percentage point reduction in the likelihood of a property conviction and a 0.39 percentage point reduction in the likelihood of a violent conviction by age 24 (column 1, panel B); as before, these effects are larger in high-poverty counties. The lack of clear differences in effects across crime types limits any insights into the channels through which early childhood education influences later criminal behavior.

Racial differences in the likelihood of being eligible for or affected by each program, along with racial differences in the counterfactual options for early childhood education, suggest that the effects of Head Start and Smart Start may differ

[59] Implied "*p*-values" presented are the two-tailed statistics calculated as the share of coefficient estimates obtained under random assignment of Head Start timing or Smart Start funding that are larger in absolute magnitude than the estimate produced using the true timing of assignment and funding. We preserve the patterns of availability and funding and randomly reassign to counties to better approximate the structure of rollout. This results in more conservative *p*-values than a strictly random assignment of timing or funding and better mimics our assumption of conditionally random assignment.

Table 4—Effect of Early Childhood Education on Criminal Conviction, by Race

	All (1)	High poverty (2)	Low poverty (3)
Panel A. Head Start			
White			
Head Start availability	−0.0038	−0.0105	0.0001
	(0.0039)	(0.0069)	(0.0040)
Observations	667	252	415
Mean	0.0267	0.0254	0.0271
Non-White			
Head Start availability	0.0014	−0.0119	0.0051
	(0.0067)	(0.0062)	(0.0106)
Observations	667	252	415
Mean	0.1008	0.0737	0.1091
Panel B. Smart Start			
White			
Smart Start ($1,000s)	−0.0026	−0.0029	−0.0012
	(0.0020)	(0.0053)	(0.0022)
Observations	1,329	674	655
Mean	0.0315	0.0272	0.0328
Non-White			
Smart Start ($1,000s)	−0.0191	−0.0216	−0.0128
	(0.0064)	(0.0057)	(0.0122)
Observations	1,329	674	655
Mean	0.0948	0.0805	0.1061

Notes: Each cell reports a separate OLS regression with standard errors clustered at the birth county level and reported in parentheses. All specifications include birth county and birth cohort fixed effects. Panel A reports results using the Head Start sample for White cohorts and non-White cohorts separately. Observations are at the birth-county-by-birth-year level and are weighted by the number of births in each county in 1955. The dependent variable is the fraction of individuals of a given race in a given birth-county-and-birth-year cohort that are later convicted of a UCR Part 1 crime in North Carolina by age 35. (Sample sizes are smaller for these specifications because the natality files for 25 percent of counties in North Carolina do not have race breakdowns before 1969 and we do not know the race of approximately 13 percent of births in our sample.) The reported variable of interest is an indicator for whether Head Start was available to a given county birth cohort. Panel B reports results using the Smart Start sample for White cohorts and non-White cohorts separately. Observations are at the birth-county-by-birth-year level and are weighted by the number of births in each county in 1980. The dependent variable is the fraction of individuals of a given race in a given birth-county-and-birth-year cohort that are later convicted of a UCR Part 1 crime in North Carolina by age 24. The reported variable of interest is a measure of Smart Start funding penetration for a given county birth cohort, constructed following Ladd, Muschkin, and Dodge (2014). (Sample sizes are smaller for these specifications because from 1989 to 1993 the natality files for 25 percent of counties in North Carolina do not have race breakdowns. For these years, race is available only for counties in which 1980 populations for the non-White group formed at least 10 percent of the total population or numbered at least 10,000.) See the notes to Table 1 for additional sample restrictions and definitions.

by race. In Table 4, we explore these heterogeneous effects by presenting our main estimates from Table 2 separately by race.[60] We find similar effects of Head Start availability for Whites and non-Whites, though only the reductions in adult convictions for non-Whites in high-poverty counties are different from zero at a 10 percent

[60] In online Appendix Table A17, we explore heterogeneous effects by sex by presenting our main estimates from Table 2 separately by sex. We cannot rule out that in both programs, early childhood availability in high-poverty counties resulted in similar percentage decreases in criminal behavior for both sexes.

significance level.[61] The lack of significant differences across races is somewhat surprising given that Black children were more likely to be eligible for and enroll in Head Start. In the case of Smart Start, we find that funding exposure had significantly larger effects for non-Whites, and again, we find that these effects are concentrated in high-poverty counties. Smart Start reduced non-White convictions by 20 percent overall and 27 percent in high-poverty counties (Table 4). This result is even more dramatic in counties without Head Start by 1980: in high-poverty counties without prior Head Start availability, Smart Start reduced non-White convictions by 35 percent (online Appendix Table A16). These results are consistent with a substantial racial gap in counterfactual early childhood education opportunities in high-poverty counties in North Carolina in the 1990s.

Ladd, Muschkin, and Dodge's (2014) estimates similarly suggest larger effects of Smart Start funding on the test scores of Black children, although the implied differences are modest and the effects are entirely driven by differences in maternal education.[62] This lesser degree of heterogeneity in test scores is consistent with differential effects of early childhood education on the accumulation of cognitive versus noncognitive skills. For example, evaluations of Perry Preschool, which enrolled disadvantaged Black children, did not produce long-run gains in IQ (particularly for males) but did produce persistent improvements in behavior.

C. *Concerns about Migration out of North Carolina*

One potential threat to the validity of our estimates relates to the data's coverage of convictions. While the data contain the universe of individuals convicted of a crime in North Carolina during this time period and allow us to link these individuals to their counties of birth, they are limited in that they do not allow us to observe convictions for individuals who are born in North Carolina and then leave the state. Fortunately, most individuals born in North Carolina remain there during adulthood; roughly 70 percent of individuals born in North Carolina during our sample periods reside there between the ages of 18 and 35. This share is even higher (roughly 80 percent) for those with the highest rates of criminal behavior (between ages 18 and 24, non-White, or with less than a high school degree).

If Head Start availability or Smart Start funding has differential effects on migration out of the state, it will not affect our estimates of convictions in North Carolina, but it may limit our ability to interpret them as representing an overall reduction in criminal behavior. A specific concern is that Head Start availability or Smart Start funding did not reduce criminal behavior but simply increased the likelihood of leaving the state, leading us to misattribute the estimated reduction in the likelihood of criminal convictions.

[61] During this period in North Carolina, the Black population comprised more than 95 percent of the non-White population (1970 census).

[62] Ladd, Muschkin, and Dodge (2014) find no evidence of differential effects of Smart Start funding on the test scores of Black children once they condition on maternal education and its interaction with Smart Start funding. However, they also show that the effects on test scores are larger for mothers with less than a high school degree. The lower average education levels among Black mothers in North Carolina imply larger effects for Black children when not separately conditioning on maternal education and its interaction with Smart Start funding levels, but the implied difference in the effect of Smart Start on test score outcomes is modest.

In online Appendix Table A18, we explore the relationship between measures of childhood Head Start availability (at the state-of-birth-by-birth-cohort level) and the likelihood of living in one's state of birth. Across a variety of approaches and subsamples, our estimates indicate a small and nonsignificant relationship between childhood Head Start availability and the likelihood of living in one's state of birth. Assuming similar patterns of criminality among North Carolina leavers and stayers, our upper-bound estimate of additional migration can explain at most 5 percent of our estimated effect.[63] Even this upper bound is likely an overestimate, as the mean rate of criminal conviction for movers to North Carolina (i.e., the equivalent of state-of-birth leavers) is lower than the rate for those born in North Carolina in our data. We have conducted similar analyses at the county–year birth cohort level using restricted-use ACS data.[64] While these results are undergoing the disclosure process, they are consistent with the estimates in online Appendix Table A18.

Smart Start funding similarly appears to have no effect on migration out of North Carolina, at least as measured by administrative schooling data through grade 3 (Ladd, Muschkin, and Dodge 2014). While we have been unable to determine a way to use publicly available data to test for differential migration at older ages (18 to 24), we have conducted similar analyses (as discussed above) with restricted-use ACS data that contain county of birth. While the results are undergoing the disclosure process, they are consistent with the conclusion of Ladd, Muschkin, and Dodge (2014) regarding the effect of Smart Starting funding on migration.[65]

Given the modest level of migration out of one's state of birth (and the lack of evidence of any differential out-migration), we view our estimates as providing a lower bound for the overall effect on criminal behavior.

D. *Magnitude of Effect on Criminal Behavior*

Our estimates suggest that the availability and funding of early childhood education causes substantial reductions in adult criminal behavior. Our preferred estimates from studying both Head Start and Smart Start indicate reductions in the likelihood of any serious conviction of approximately 1 percentage point (approximately 20 percent) for the average program in high-poverty counties. While both programs generate similar reductions in serious criminal convictions in high-poverty counties, Head Start is somewhat more expensive in real terms.[66] Scaling by the average cost of Head Start in high-poverty counties and adjusting everything to 2015 dollars, we find that Head Start availability reduced the likelihood of a serious criminal conviction by 0.621 percentage points per $(2015)1,000, while Smart Start reduced the

[63] Even assuming the largest estimated effect on migration, it would have to be the case that 65 percent of the marginal migrants were criminals to account for our estimates.

[64] These restricted data are available in the Federal Statistical Research Data Center and allow us to directly observe county of birth for a large sample.

[65] If anything, exposure to additional Smart Start funding in early childhood appears to make individuals less likely to leave North Carolina, suggesting that the magnitude of our estimated effects on overall crime may be slightly attenuated.

[66] Around $(2015)2,000 per individual in the cohort for Head Start and around $(2015)1,600 per fully treated individual for Smart Start in high-poverty counties.

likelihood by roughly 0.720 percentage points per \$(2015)1,000.[67] Alternatively put, Head Start and Smart Start cost about \$1,500 per percentage point reduction in serious criminal convictions.

While the specific criminal outcome measures differ, the implied cost per percentage point of crime reduction for Head Start and Smart Start spending are similar to but somewhat smaller than those implied by evaluations of the Perry Preschool program: \$1,800 per percentage point arrest reduction (authors' calculations from Heckman et al. 2010 estimates). When we scale arrests to convictions, we find that this cost-effectiveness advantage for Head Start and Smart Start increases.[68]

E. *Quantifying the Benefits*

How do the future benefits of crime reduction compare to the costs of these programs? To enable this comparison, we use the detailed conviction histories of offenders in North Carolina to construct back-of-the-envelope measures of the discounted social cost of their crimes and then estimate the effect of Head Start and Smart Start on these measures. We construct this measure by applying the social cost estimates for the given crime from McCollister, French, and Fang (2010) and Miller, Cohen, and Wiersema (1996) to each conviction by age 24 (Smart Start) or 35 (Head Start) and then discounting that cost from the age at conviction back to age 4 (for ease of comparison with the program costs).[69] We then aggregate this social cost measure to the birth cohort level. This measure serves as a lower bound for the social cost of a cohort's crimes, as it includes only crimes with a social cost estimate available and does not account for the fact that multiple offenses are often associated with a single conviction. To address the latter issue, we construct a second measure using a similar approach, but where we weight each conviction by the average number of offenses per conviction for each crime type.[70] For this measure, we are limited to the set of crimes for which we can observe the number of offenses in the FBI's Uniform Crime Reports (i.e., Part 1 crimes). In all cases, we limit our estimates to convictions observed within our data (i.e., by age 35 for Head Start cohorts and age 24 for Smart Start cohorts); expected benefits are likely substantially larger if we included forecasted commission of crime across ages.

Table 5 shows the equation (1) estimates of the effects of early childhood education programs on these discounted social cost of crime measures, where each cell represents the result from a different regression. The estimates are imprecise and based on imperfect measures, but the magnitudes suggest that, particularly in high-poverty counties, the crime benefits of these programs make up a large proportion of the cost

[67] The Smart Start treatment is denoted in nominal dollars, so we adjust to \$(2015)1,000.

[68] Another approach to comparing magnitudes is to calculate implied treatment-on-the-treated (TOT) effects of each intervention. While this isn't possible in the case of Smart Start, our Head Start estimates imply TOT effects of 6 to 9 percentage points, somewhat smaller than those reported in evaluations of the Perry Preschool program for somewhat similar measures (See online Appendix D for additional discussion.)

[69] When cost estimates for the same crime are available in McCollister, French, and Fang (2010) and Miller, Cohen, and Wiersema (1996), we use the more recent estimates from McCollister, French, and Fang (2010). Following the baseline scenario in Hendren and Sprung-Keyser (2020), we use a 3 percent discount rate.

[70] Weights are calculated as the number of offenses reported in North Carolina from 1980–2010 (FBI Uniform Crime Reports) divided by the number of convictions in North Carolina from 1980–2010, for a given type of crime.

Table 5—Social Benefits of Crime Reduction from Early Childhood Education

	HS availability		SS ($1,000)	
	All crimes (1)	Part 1 (offense-weighted) (2)	All crimes (3)	Part 1 (offense-weighted) (4)
All counties	−129.836	−226.994	−170.869	−785.579
	(174.046)	(993.404)	(80.533)	(550.538)
Mean	3,128.42	1,3295.42	2,050.44	1,1521.31
High-poverty counties	−537.455	−2,348.500	−133.408	−930.291
	(365.998)	(2,311.682)	(103.431)	(920.457)
Mean	3,243.10	14,787.51	1,899.64	11,135.20
Low-poverty counties	55.537	790.803	−141.017	−609.216
	(170.637)	(874.318)	(103.420)	(707.764)
Mean	3,096.28	12,877.19	2,100.52	11,649.52

Notes: Each cell reports a separate OLS regression with standard errors clustered at the birth county level and reported in parentheses. All specifications include birth county and birth cohort fixed effects. In columns 1 and 3, the dependent variable is a measure of discounted social cost of crime constructed by applying the social cost estimates for each crime from McCollister, French, and Fang (2010) and Miller, Cohen, and Wiersema (1996) to each conviction by age 24 (Smart Start) or 35 (Head Start) and then discounting that cost from the age at conviction back to age 4 using a 3 percent discount rate. In columns 2 and 4, the dependent variable is a similar measure constructed for UCR Part 1 crimes only that adjusts for the average number of offenses per conviction (1980–2010) using counts from North Carolina UCR known offenses and North Carolina Department of Public Safety data (for ease of comparison with the program costs). Observations are at the birth-county-by-birth-year level and are weighted by the number of births in each county in 1955 (Head Start) or 1980 (Smart Start). All values are converted to 2015 dollars. For Head Start, the reported variable of interest is an indicator for whether the program was available to a given county birth cohort. For Smart Start, the reported variable of interest is a measure of Smart Start funding penetration for a given county birth cohort, constructed following Ladd, Muschkin, and Dodge (2014).

of the programs when accounting for multiple offenses per conviction.[71] For example, column 2, row 2 of Table 5 shows that Head Start availability caused a $2,348 (16 percent) reduction in discounted social crime costs in high-poverty counties. Comparing the magnitudes of the Head Start estimates in column 2 to the average cost of the program per child (both participants and nonparticipants) of $886 in all counties and $2,044 in high-poverty counties suggests that the crime benefits of the program account for 26 percent of the costs in all counties and 115 percent of the costs in high-poverty counties.[72] Performing a similar comparison between the magnitudes of the Smart Start estimates in column 4 with the $1,000 cost per child suggests that crime benefits through age 24 account for 79 percent of the costs of the program in all counties and 93 percent in high-poverty counties.[73] These results support the conclusion that individual investments in early childhood education are inefficiently low without subsidies, particularly given the large impacts of these programs on other outcomes with substantial social benefits (e.g., education).

[71] Effects are substantially larger when including costs to victims rather than only supervision costs. Supervision only costs are smaller for the Smart Start sample than the Head Start sample, reflecting both the differences in age ranges over which these costs are calculated and the shift from incarceration to probation for these cohorts.

[72] 2015 dollars are used in all estimates/calculations.

[73] These estimates are based on convictions through age 24. If we scale by the ratio of Head Start benefits at age 35 versus 24 (1.5 to 2.5), crime benefits generally exceed the costs.

V. Conclusion

We contribute to the sparse literature on the developmental factors that influence an individual's likelihood of becoming a criminal by exploring the effect of early childhood education on criminal behavior. This relationship has become increasingly relevant given recent expansions in the share of children attending public preschools and widespread efforts to improve the quality of early childhood education. These expansions have been driven in large part by policymakers' belief that early childhood education has large impacts later in life. Crime reduction in particular is central to the widely publicized benefit-cost analyses of these programs (e.g., crime reduction accounts for 40–65 percent of the benefits estimated in the context of Perry Preschool). However, inconsistent findings across similar studies, limited statistical power with correspondingly high implied false discovery rates, and differences in contexts across decades challenge whether we can generalize the crime effect estimates from earlier studies to the present day.

We bring new evidence to this literature by using administrative crime data to investigate the effects of two large-scale early childhood education programs operating in two different time periods. This approach yields substantial external validity improvements over prior long-run crime effect estimates by relying on a much broader treated population and enabling a direct comparison of similar adult crime reductions from two different interventions implemented in very different contexts. Our administrative crime data grant us significant improvements in sample size (hundreds of thousands versus a hundred or thousands) and corresponding increases in precision. This precision reduces the likelihood that any statistically significant estimated effect is a false positive (false discovery rate), reduces the expected upward bias in reported effects, and analogously increases precision. This precision also allows us to explore heterogeneity in effects across areas with different levels of disadvantage and access to existing care, potentially informing the debate between universal and more targeted early education policies. While the precision and five-decade-long time span of these data are substantial advantages, they come at the cost of limiting our analysis to the state of North Carolina.[74]

We show that early childhood education reduces adult criminal behavior across two different programs and time periods. We find that Head Start availability in the 1960s and 1970s reduces the likelihood of a serious conviction by age 35 by 1.3 percentage points in high-poverty counties but has no measurable effect in low-poverty counties. Implemented two-to-three decades later, Smart Start generates similar reductions in adult criminal behavior, with effects similarly concentrated in high-poverty counties and among Black children. Taken together, these results suggest the general capacity of early childhood education to reduce the propensity for criminal behavior in those areas and among those individuals with the fewest resources. We also find that Smart Start's effects are largest in counties

[74] While North Carolina is the ninth most populous state and has a similar age distribution (and percentage of young children) as the rest of the country, the state has a much higher percentage of Black individuals. This is advantageous for studying the effects of early childhood education on the criminality of Black individuals (who are both more likely to be affected by these programs and more likely to be convicted of a crime) but may limit the generalizability of the resulting estimates to states with smaller minority populations.

without Head Start access, which suggests that there are diminishing returns to early childhood education funding. Both patterns of results are consistent with the general pattern of effects in the literature, though we cannot draw strong conclusions from a small set of studies that identify heterogeneous parameters of different programs.

Also consistent with prior literature on earlier cohorts, we find that the effects of a more modern program on cognitive skills are not predictive of effects on longer-run behavioral outcomes. Indeed, our estimated effects of Smart Start on criminal behavior are more than seven times the effect implied by earlier estimates of the program's test score effects applied to the conditional correlation between test scores and crime. Future work might explore whether effects on noncognitive measures may be more predictive of effects on this margin.

Finally, we take advantage of the greater sample sizes and detail contained in the data to show that the discounted benefits generated by early childhood education's later crime reduction are substantial relative to the costs of the education provided. This is especially noteworthy considering that later crime reduction was not the stated objective of either program and that these benefits likely accrue in large part to those who did not experience the program themselves. The magnitude of these external benefits implies that the social benefits of early childhood education vastly exceed the private benefits. That said, these benefits accrue disproportionately to high-poverty areas and to those lacking access to other early childhood education subsidies. Taken together, our results provide evidence in support of recent efforts to expand and improve early childhood education but point to large potential gains from targeting these efforts toward areas with the greatest need.

REFERENCES

Anders, John, Andrew C. Barr, and Alexander A. Smith. 2023. "Replication data for: The Effect of Early Childhood Education on Adult Criminality: Evidence from the 1960s through 1990s." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.38886/E140241V1.

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484): 1481–95.

Bailey, Martha J., and Andrew Goodman-Bacon. 2015. "The War on Poverty's Experiment in Public Medicine: Community Health Centers and the Mortality of Older Americans." *American Economic Review* 105 (3): 1067–1104.

Barr, Andrew, and Chloe Gibbs. Forthcoming. "Breaking the Cycle? Intergenerational Effects of An Anti-poverty Program in Early Childhood. *Journal of Political Economy*.

Barr, Andrew, and Alexander A. Smith. 2021. "Fighting Crime in the Cradle: The Effects of Early Childhood Access to Nutritional Assistance." *Journal of Human Resources* 0619-10276R2.

Bayer, Patrick, Randi Hjalmarsson, and David Pozen. 2009. "Building Criminal Capital behind Bars: Peer Effects in Juvenile Corrections." *Quarterly Journal of Economics* 124 (1): 105–47.

Campbell, Frances A., Elizabeth P. Pungello, Margaret Burchinal, Kirsten Kainz, Yi Pan, Barbara H. Wasik, Oscar A. Barbarin, Joseph J. Sparling, and Craig T. Ramey. 2012. "Adult Outcomes as a Function of an Early Childhood Educational Program: An Abecedarian Project Follow-up." *Developmental Psychology* 48 (4): 1033–43.

Campbell, Frances A., Craig T. Ramey, Elizabeth Pungello, Joseph Sparling, and Shari Miller-Johnson. 2002. "Early Childhood Education: Young Adult Outcomes from the Abecedarian Project." *Applied Developmental Science* 6 (1): 42–57.

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

Chaiken, Jan M., and Marcia R. Chaiken. 1982. *Varieties of Criminal Behavior*. Santa Monica, CA: Rand Corporation.

De Chaisemartin, Clément, Xavier D'Haultfoeuille, and Yannick Guyonvarch. 2018. "DID_MULTIPLEGT: Stata Module to Estimate Sharp Difference-in-Difference Designs with Multiple Groups and Periods." https://ideas.repec.org/c/boc/bocode/s458643.html.

Deming, David. 2009. "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start." *American Economic Journal: Applied Economics* 1 (3): 111–34.

Deming, David J. 2011. "Better Schools, Less Crime?" *Quarterly Journal of Economics* 126 (4): 2063–2115.

Doyle, Joseph J., Jr. 2007. "Child Protection and Child Outcomes: Measuring the Effects of Foster Care." *American Economic Review* 97 (5): 1583–1610.

Doyle, Joseph J., Jr. 2008. "Child Protection and Adult Crime: Using Investigator Assignment to Estimate Causal Effects of Foster Care." *Journal of Political Economy* 116 (4): 746–70.

Farrington, David P. 2006. "Methodological Quality and the Evaluation of Anti-crime Programs." *Journal of Experimental Criminology* 2 (3): 329–37.

Fitzpatrick, Maria D. 2008. "Starting School at Four: The Effect of Universal Pre-kindergarten on Children's Academic Achievement." *BE Journal of Economic Analysis and Policy* 8 (1). https://doi.org/10.2202/1935-1682.1897.

Fitzpatrick, Maria Donovan. 2010. "Preschoolers Enrolled and Mothers at Work? The Effects of Universal Prekindergarten." *Journal of Labor Economics* 28 (1): 51–85.

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

García, Jorge Luis, James J. Heckman, and Anna L. Ziff. 2019. "Early Childhood Education and Crime." *Infant Mental Health Journal* 40 (1): 141–51.

Gibbs, Chloe, Jens Ludwig, and Douglas L. Miller. 2011. "Does Head Start Do Any Lasting Good?" NBER Working Paper 17452.

Goodman-Bacon, Andrew. 2018. "Difference-in-Differences with Variation in Treatment Timing." NBER Working Paper 25018.

Heckman, James, Rodrigo Pinto, and Peter Savelyev. 2013. "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes." *American Economic Review* 103 (6): 2052–86.

Heckman, James J., Seong Hyeok Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz. 2010. "The Rate of Return to the HighScope Perry Preschool Program." *Journal of Public Economics* 94 (1–2): 114–28.

Heckman, James J., Rodrigo Pinto, Azeem M. Shaikh, and Adam Yavitz. 2011. "Inference with Imperfect Randomization: The Case of the Perry Preschool Program." NBER Working Paper 16935.

Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24 (3): 411–82.

Hendren, Nathaniel, and Ben Sprung-Keyser. 2020. "A Unified Welfare Analysis of Government Policies." *Quarterly Journal of Economics* 135 (3): 1209–1318.

Hill, Patrick L., Brent W. Roberts, Jeffrey T. Grogger, Jonathan Guryan, and Karen Sixkiller. 2011. "Decreasing Delinquency, Criminal Behavior, and Recidivism by Intervening on Psychological Factors Other than Cognitive Ability: A Review of the Intervention Literature." NBER Working Paper 16698.

Hindelang, M.J., T. Hirschi, and J.G. Weis. 1981. *Measuring Delinquency*. Thousand Oaks, CA: Sage Publications.

Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond. 2016. "Long-Run Impacts of Childhood Access to the Safety Net." *American Economic Review* 106 (4): 903–34.

Jackson, C. Kirabo. 2018. "What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes." *Journal of Political Economy* 126 (5): 2072–2107.

Johnson, Rucker C., and C. Kirabo Jackson. 2018. "Reducing Inequality through Dynamic Complementarity: Evidence from Head Start and Public School Spending." NBER Working Paper 23489.

Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman. 2001. "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment." *Quarterly Journal of Economics* 116 (2): 607–54.

Kids Counts Data Center. 2000–2021. "Kids Count Data Indicators." Annie E. Casey Foundation. https://datacenter.kidscount.org/data (accessed June 1, 2019).

Kline, Patrick, and Christopher R. Walters. 2016. "Evaluating Public Programs with Close Substitutes: The Case of Head Start." *Quarterly Journal of Economics* 131 (4): 1795–1848.

Kling, Jeffrey R., Jens Ludwig, and Lawrence F. Katz. 2005. "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment." *Quarterly Journal of Economics* 120 (1): 87–130.

Ladd, Helen F., Clara G. Muschkin, and Kenneth A. Dodge. 2014. "From Birth to School: Early Childhood Initiatives and Third-Grade Outcomes in North Carolina." *Journal of Policy Analysis and Management* 33 (1): 162–87.

List, John A., Fatemeh Momeni, and Yves Zenou. 2019. "Are Estimates of Early Education Programs Too Pessimistic? Evidence from a Large-Scale Field Experiment That Causally Measures Neighbor Effects." Unpublished.

Lochner, Lance, and Enrico Moretti. 2004. "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports." *American Economic Review* 94 (1): 155–89.

Ludwig, Jens, and Jeffrey R. Kling. 2007. "Is Crime Contagious?" *Journal of Law and Economics* 50 (3): 491–518.

Ludwig, Jens, and Douglas L. Miller. 2007. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design." *Quarterly Journal of Economics* 122 (1): 159–208.

McCollister, Kathryn E., Michael T. French, and Hai Fang. 2010. "The Cost of Crime to Society: New Crime-Specific Estimates for Policy and Program Evaluation." *Drug and Alcohol Dependence* 108 (1): 98–109.

Miller, Douglas L., Na'ama Shenhav, and Michel Z. Grosz. 2019. "Selection into Identification in Fixed Effects Models, with Application to Head Start." NBER Working Paper 26174.

Miller, Ted R., Mark A. Cohen, and Brian Wiersema. 1996. *Victim Costs and Consequences: A New Look*. Washington, DC: National Institute of Justice.

North Carolina Department of Public Safety. 1972–2018. "Offender Public Information." North Carolina Department of Public Safety. https://webapps.doc.state.nc.us/opi/downloads.do?method=view (accessed June 1, 2020).

North Carolina Office of State Budget and Management. 2010–2019. "County Total Age Groups." North Carolina Office of State Budget and Management. https://www.osbm.nc.gov/content/july-1-2010-county-total-age-groups-standard (accessed June 1, 2019).

North Carolina Partnership for Children. 2019. "Smart Start Funding." North Carolina Partnership for Children (accessed April 1, 2019).

Office of Child Development. 1968. *Project Head Start 1965–1967: A Descriptive Report of Programs and Participants.* Washington, DC: Department of Health, Education, and Welfare.

Office of Child Development. 1970. *Project Head Start 1968: A Descriptive Report of Programs and Participants.* Washington, DC: Department of Health, Education, and Welfare.

Olds, D., C.R. Henderson, Jr., R. Cole, J. Eckenrode, H. Kitzman, D. Luckey, L. Pettitt, K. Sidora, P. Morris, and J. Powers. 1998. "Long-Term Effects of Nurse Home Visitation on Children's Criminal and Antisocial Behavior: 15-Year Follow-up of a Randomized Controlled Trial." *JAMA* 280 (14): 1238–44.

Olds, David L., Lois Sadler, and Harriet Kitzman. 2007. "Programs for Parents of Infants and Toddlers: Recent Evidence from Randomized Trials." *Journal of Child Psychology and Psychiatry* 48 (3–4): 355–91.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Jose Pacas, Megan Schouweiler, and Matthew Sobek. 2021. "Integrated Public Use Microdata Series, Decennial Census." https://doi.org/10.18128/D010.V11.0 (accessed June 1, 2019).

Sanbonmatsu, Lisa, Jens Ludwig, Lawrence F. Katz, Lisa A. Gennetian, Greg J. Duncan, Ronald C. Kessler, Emma K. Adam, Thomas McDade, and Stacy T. Lindau. 2011. *Moving to Opportunity for Fair Housing Demonstration Program: Final Impacts Evaluation*. Washington, DC: U.S. Department of Housing and Urban Development.

Stuart, Elizabeth A. 2010. "Matching Methods for Causal Inference: A Review and a Look Forward." *Statistical Science* 25 (1): 1–21.

Sun, Liyang. 2020. "EVENTSTUDYWEIGHTS: Stata Module to Estimate the Implied Weights on the Cohort-Specific Average Treatment Effects on the Treated (CATTs) (Event Study Specifications)." https://ideas.repec.org/c/boc/bocode/s458833.html.

Sun, Liyang, and Sarah Abraham. 2020. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics* 225 (2): 175–99.

Thompson, Owen. 2017. "Head Start's Long-Run Impact: Evidence from the Program's Introduction." *Journal of Human Resources*: 0216–7735r1.

United States Census Bureau. 1990–2018. "Cartographic Boundary Files—Shapefile." United States Census Bureau. https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html (accessed May 1, 2016).

United States Department of Agriculture Food and Nutrition Service. 1969–2021. "SNAP Data Tables." United States Department of Agriculture. https://www.fns.usda.gov/pd/supplemental-nutrition-assistance-program-snap (accessed June 1, 2019).