Title
Does Head Start Improve Long-Term Outcomes? Evidence from a Regression Discontinuity Design

Journal
UCSC Economics Department Seminars, 50(3)

ISSN
0942-5616

Authors
LUDWIG, JENS O
Miller, Doug

Publication Date
2004-11-30

Very Preliminary
Please Do Not Cite or Quote

Does Head Start Improve Long-Term Outcomes? Evidence from a Regression Discontinuity Design

draft date: November 30, 2004

Jens Ludwig
Georgetown University

Douglas L. Miller
University of California, Davis

Contact information:
Jens Ludwig
Georgetown Public Policy Institute
Georgetown University
3520 Prospect Street, NW
Washington, DC 20007
(202) 687-4997
fax (202) 687-5544
ludwigj@georgetown.edu

This paper substantially extends an earlier paper written with Nate Balis presented at the Fall 2001 APPAM meetings, and was supported in part by the Georgetown University Graduate School of Arts and Sciences as well as a grant from the Foundation for Child Development to the Georgetown Center for Research on Children in the U.S. Thanks to Bradley Hardy, Zac Hudson, Sinead Keegan, Robert Malme, Meghan McNally, Julie Morse, Berkeley Smith and Eric Younger for excellent research assistance, to Jule Sugarman, Craig Turner and Edward Zigler for information about the history of Head Start, to Rob Baller, Eliana Garces and Michael Maltz for sharing their data, and to Mark Cohen, Philip Cook, William Dickens, Greg Duncan, Ted Gayer, William Gormley, Alison Hagy, Brian Jacob, Deborah Phillips, Peter Reuter, Dave Ribar and seminar participants at UC-Davis and Cornell for helpful comments. Any errors and all opinions are of course ours alone.

Abstract

This paper exploits a new source of variation in Head Start funding to identify the program's long-term effects. In 1965 the Office of Economic Opportunity (OEO) provided technical assistance to the 300 poorest counties in the U.S. to develop Head Start funding proposals, but did not provide similar assistance to other counties. We show that the result is a substantial difference in Head Start funding and participation rates in those counties just above and below OEO's poverty-rate cutoff for technical assistance, differences that seem to have persisted through at least the 1970's. This discontinuity in Head Start funding and participation are associated with discontinuities in educational attainment.

I. Introduction

Head Start was established in 1964 as part of the War on Poverty to provide educational, health and other services to poor children. Economists have been interested in Head Start because of the program's potential to improve the human capital of poor children and thereby reduce the inter-generational correlation of income. Early childhood may provide a particularly promising target for human capital interventions given the cumulative nature of learning, the possibility that learning may occur at an unusually rapid rate during the early years, and the fact that much of the gap in achievement test scores across race or class lines arises before children even start school (Bloom, 1964; Entwisle, Alexander and Olsen, 1997; Phillips, Crouse and Ralph, 1998; Phillips *et al*., 1998; Shonkoff and Phillips, 2000). Head Start is also of interest to economists because the program may reduce the external costs associated with a variety of anti-social behaviors that may have their antecedents in childhood poverty and early educational failure (Currie, 2001).

Initially implemented as an eight-week summer intervention, Head Start has grown considerably over time and now serves more than 800,000 children for ninth months each year at a total annual cost of $6.7 billion (Haskins, 2004).[1] Yet controversy about whether Head Start produces lasting benefits in practice has surrounded the program since its inception. The first claim about "fade out" of Head Start's benefits was made in 1966, only one year after the program's launch.[2] The general principle that early childhood intervention can produce lasting benefits is suggested by a number of randomized experimental programs such as Perry Preschool and Abacedarian (Barnett, 1992, 1995; Donohue and Siegelman, 1998; Karoly *et al*., 1998; Campbell *et al*., 2002; Currie, 2001). However these programs are much smaller and more intensive than Head Start, and as a result are not very informative about what we might expect from the less-costly, larger-scale Head Start program.[3]

To date the best available research on the longer-term effects of Head Start comes from within-family comparisons of siblings who have and have not participated in the program (Currie and Thomas, 1995, Garces, Thomas and Currie, 2002, hereafter CT and GTC). These sibling differences suggest that the program may have long-term effects on educational outcomes for whites, while reducing criminal involvement among African-Americans.[4] These papers

---

1 The number of total participants (summer # in parentheses) by year (Jones, 1979): 1965 – 561,000 (561,000); 1966 – 733,000 (573,000); 1967 – 681,000 (466,000); 1968 – 693,825 (476,825); 1969 – 635,121 (421,665); 1970 – 434,880 (195,328); 1971 – 419,971 (123,485); 1972 – 379,000 (86,400); 1978 – 389,500 (26,000).

2 The 1966 study was conducted in New York City and found that children who had enrolled in the program during its first summer of operation did not have significantly different cognitive test scores six to eight months later compared to non-participants (Wolff and Stein, 1966, Zigler and Muenchow, 1992). However such comparisons may confound program effects those of unmeasured child or family attributes associated with program participation.

3 One exception is the Chicago Child-Parent Center (CPC) Program, which, like Head Start, is a federally funded large-scale public program. Reynolds et al. (2001) suggest that CPC participants are more likely to complete high school and less likely to be arrested compared to non-participants. However these conclusions stem from a non-experimental regression that conditions on only a fairly basic set of socio-demographic characteristics, and so may confound the CPC program's impact with unobserved individual attributes associated with program participation.

4 On the other hand, using cross-section data from the National Longitudinal Survey of Youth 1997, Aughinbaugh

substantially improve upon previous non-experimental evaluations by controlling for unmeasured family fixed-effects that may be associated with Head Start participation. Yet there necessarily remains some uncertainty about what drives variation across siblings in Head Start participation. Of particular concern is the possibility that participation is related to unmeasured child or time-varying family characteristics that also affect children's outcomes. As one recent review notes, with respect to long-term impacts: "The jury is still out on Head Start" (Currie, 2001, p. 213). A more pessimistic assessment of Head Start's long-term effects contributed to the Bush Administration's push for, among other things, a shift in the program's focus from "comprehensive services to intellectual development" (Haskins, 2004).

The present paper provides new evidence on the long-term effects of the Head Start program as it was implemented in its original form, with a focus on comprehensive service delivery to poor children. The paper complements the excellent papers by CT and GTC by identifying the effects of Head Start using a very different source of variation in program participation. Specifically, we exploit a discontinuity in program funding across counties resulting from how the Office of Economic Opportunity (OEO) launched the program during the spring of 1965. Unlike many other federal social programs, Head Start provides funding directly to local service providers. Out of concern that the most disadvantaged communities would be unable to develop proposals for Head Start funding, OEO sent 100 Presidential Management Interns (PMIs) to the 300 poorest counties in the country to identify potential service providers and help them develop proposals (Jones, 1979, pp. 6-7). The result is substantially higher Head Start funding in "treatment" counties with 1960 poverty rates that place them among the 300 poorest in the country – what we will call the "OEO cutoff" – compared to "control" counties with poverty rates just below this cutoff.

We use the partially linear regression approach from Porter (2003) to look for discontinuous changes at the OEO cutoff in long-term outcomes such as educational attainment, achievement test scores, labor market outcomes and criminal activity. We assume that absent the Head Start program, differences in long-term outcomes across counties with respect to their 1960 poverty rates would be smooth near the OEO cutoff. This assumption seems plausible in our application given that 1960 county poverty rates were determined well before the War on Poverty, and so could not have been manipulated in response to possible funding advantages. Moreover as we demonstrate below, it appears that the arbitrary cutoff that OEO used for Head Start grant-writing assistance was not used to allocate funds for other federal programs.

Our main finding is that Head Start appears to increase educational attainment, a conclusion that comes from two separate data sources – county-level data from the 1990 Census and student-level micro-data from the National Education Longitudinal Study of 1988 (NELS). In the Census data the discontinuity at the OEO cutoff in educational attainment is concentrated among those adults that are young enough to have been affected by Head Start either directly or

---

(2002) does not find much evidence for long-term impacts of Head Start participation. Currie and Thomas (2000) find little long-term effect on Head Start in the NELS data for blacks, although this may be contingent on the quality of the public schools that children go on to attend.

indirectly. Everyone in the NELS sample of 8th graders in 1988 was of Head Start age in 1977 or 1978, after the program was in operation. The fact that we observe a pronounced discontinuity in Head Start participation rates at the OEO cutoff for NELS respondents suggests that the cross-sectional variation in funding generated by the 1965 program rollout persisted for many years.

Arguably the main challenge to our findings is that while both datasets provide information on people's outcomes when they are adolescents or adults, we first identify county of residence for people several years after they were of Head Start age. With the 1990 Census data we can identify where people are living as adults in 1990, which will be at least 15 years after they would have been age-eligible to participate in Head Start. With the NELS we first identify county of residence for respondents in 8th grade, about 10 years after they would have been of Head Start age. This data limitation raises a concern about selective migration into or out of those counties with 1960 poverty rates near the OEO cutoff.

To address the problem of selective migration with the 1990 Census data we exploit the fact that within each county, some birth cohorts but not others could have participated in Head Start by virtue of their year of birth and when the program was launched. If selective migration and other unmeasured county factors have similar effects on cohorts born close together in time but differ in whether they could have participated in Head Start, then we can net out the influence of these confounders by focusing on the discontinuity in the difference in outcomes across cohorts. This "discontinuity-in-differences" estimator yields qualitatively similar findings to our main findings, consistent with the fact that most of the effects on schooling with our 1990 Census data are concentrated among those cohorts young enough to have been affected by Head Start. We cannot implement this estimator directly with the NELS, given that the survey's primary outcome data are available for only one birth cohort (8th graders in 1988). But in the same spirit as our discontinuity-in-difference estimator, we show that there is no evidence of a discontinuity in the NELS for educational attainment or income for the parents of NELS respondents, an arguably untreated, or at least less treated, cohort.

As an additional specification check we look for discontinuities in outcomes at other cutoffs where there are no discontinuities in Head Start funding. We find no discontinuities in outcomes at our pseudo-cutoff, which enhances our confidence in the research design.

As best we can tell the discontinuity in educational outcomes observed at the OEO cutoff appears to be due primarily to differences in enrollment rates rather than to higher spending per enrollee. This conclusion is offered tentatively given the large standard error around our estimate for the discontinuity in funding. In principle an alternative mechanism for our results could be a discontinuity in the technology of Head Start provision if OEO's grant-writing assistance led to changes in the program's production function. But we do not find it very plausible that a weekend of grant-writing help could have produced significant changes in the program's production technology other than through program funding.

Two other questions of interpretation are more difficult to resolve. First, our estimates identify the effects of expanding Head Start in counties around the OEO cutoff. This treatment parameter is of some policy interest, since we might expect future expansions on Head Start (the program still does not cover all income-eligible children) to focus on the most disadvantaged communities. Nevertheless our estimates are not directly informative about expanding or contracting Head Start funding and enrollments in other types of communities. Second, our estimates provide information about the long-term effects of the Head Start program in its original incarnation, when the program provided a full menu of services to low-income children. Whether the revised Head Start program with its increased emphasis on cognitive skills and intellectual development produces similar effects is not known.

The remainder of the paper is organized as follows. The next section discusses the data sources used in our analyses, while the third section provides more background on Head Start, with a particular focus on how Head Start might affect long-term outcomes and the features of the program's rollout that generates the natural experiment underlying our research design. The fourth section discusses our empirical approach for exploiting the discontinuity in county-level Head Start funding that resulted from the way the program was launched in 1965. The fifth section of the paper presents our findings for Head Start's impacts on long-term outcomes and the sixth section discusses both the limitations and policy implications of our results.

II. Data

Before discussing our research design below, we review the different sources of county- and individual-level data used in our analysis. While the latter provide us with better data about where adults or adolescents were living when they were of (or within at least 10 years of) Head Start age, the county-level data provide us with more information "near" the OEO cutoff.

A. County-Level Data

Perhaps the most important county-level characteristic for our study is the variable used by OEO to identify the 300 poorest counties that were to receive technical assistance for writing Head Start grants. One complication is that only in 1970 did the decennial census begin to collect information on the number or proportion of county residents officially living "in poverty," which was a concept defined for the first time in 1964. The 1960 Census instead reports the proportion of families in the county with incomes below $3,000.[5] We have also obtained from the National Archives and Records Administration (NARA) a special 1964 re-analysis of the 1960 conducted by the Census Bureau for OEO using the then-newly-defined federal poverty rate.[6] While the Head Start histories do not specify which of these two possible measures OEO used to identify the 300 counties to get grant-writing assistance – percent families under $3,000, or percent in poverty – our empirical analysis below suggests OEO used the 1960 poverty rate.

5 Since the official poverty threshold for a family of four in 1960 dollars is $3,002 (Citro and Michael, 1995, p. 35), it is not surprising that the percent of a county's families with incomes less than $3,000 is highly correlated with the official poverty rate for 1960 (+.95). The correlation between the two measures is not perfect because the income level used to define whether a family is in poverty varies with the family's composition, and the two measures of disadvantage produce slightly different rankings of which counties were the "poorest" in 1960.

6 NARA, Records of the Community Services Administration, Record Group 381: Putnam Print File, 1960.

From NARA we also obtained a series of OEO data files on federal expenditures per county.[7] These data are intended to provide information on spending and spending beneficiaries in each county for each year from 1967 through 1980 for each federal program, including Head Start. Unfortunately the data are poorly documented and riddled with problems,[8] and for obvious reasons OEO's staff is not helpful in resolving these puzzles. In the end only data from 1968 and 1972 were usable, in the sense that only in these years did the electronic file match published figures at the national level for total spending on all programs and for Head Start specifically. In these years we were also able to match the electronic and published figures for Head Start spending at the state level. In addition to generating a measure of Head Start spending, we also create variables for spending on other social programs and all other federal spending.[9]

Our primary data source on long-term outcomes comes from county-level files from the 1990 Census. Of particular interest for our "difference discontinuity" design are the measures that are publicly reported by the Census separately for different age groups in the STF4 file: educational attainment (high school completion or more; attendance of some college; college completion or more); unemployment; and poverty status. Unfortunately not all outcomes are reported separately by age group in the STF4, and for the outcomes that are reported in this way the age groupings sometimes differ across variables.

The next version of this paper will also examine criminal activity by using county-level data from the FBI's Uniform Crime Report (UCR) system, which compiles crimes reported by victims to police and then submitted to the FBI. While analysis of these data is not included in the present draft, we are working to assemble UCR data for 1989-1991, including data on arrest rates by age to facilitate the same sort of difference-discontinuity design as implemented with the Census data. Of course the problems with the UCR are well known, and include differences in how crimes are defined and recorded across areas and over time, as well as variation in victim

---

[7] Federal Outlays, County and State File [Machine-readable data file], 1967-1980 / conducted by the Office of Economic Opportunity for the Executive Office of the President. – Washington: OEO [producer], 1968: Washington: National Archives and Records Service [distributor]. Record Group 381. File Number: 3-381-73-157(A).

8 The records for 1968 and 1972 still contain a number of glitches such as alphabetic characters or brackets in the last columns of the program expenditure and beneficiary files, which we infer should be zeros based on comparisons with published expenditure and beneficiary data.

[9] For 1968 the data on program beneficiaries for Head Start matches up with published figures for the U.S. as a whole, although the beneficiary variable seems to be largely missing for most other federal programs. Interpretation of Head Start beneficiary data is complicated in 1968 because the program at that time was mostly summer-only, although some areas had shifted towards a year-round program. For 1972 even the beneficiary variable for Head Start records are generally missing. For 1972 we define total Head Start expenditures as federal spending dedicated to three OEO programs with activity codes listed Head Start ($328.0 million), OEO's Follow Thru program ($25.3 million), and OEO Community Services spending devoted to early childhood education ($11.7 million). The sum of these three programs is approximately equal to published figures for total Head Start spending for this year (see notes 8/18/03). Spending on other social programs is defined as expenditures through the Department of Health, Education and Welfare, excluding those made through Head Start.

reporting of crimes to the police. Other problems come from incomplete reporting of local law enforcement agencies of crime in a given year, and the limitations of the imputation methods used by the FBI and other national agencies in adjusting for incomplete reporting. These imputation problems are particularly severe when using UCR data measured at the county level (Maltz, 1999, Maltz and Targonski, 2002).

B. Individual-Level Data

Our main source of individual-level data is a restricted-use geo-coded version of the NELS, sponsored by the U.S. Department of Education to survey a nationally representative sample of 8th graders in 1988 with follow-up interviews in 1990, 1992, 1994 and 2000. These individual-level data enable us to identify the long-term outcomes of Head Start participants directly, rather than compare county-wide Head Start funding and average outcomes. These micro-data also enable us to link the behavior of people as young adults to where they were living at age 13, which is at least somewhat closer to when they would have been of Head Start age compared to when we first measure addresses for Census respondents (in 1990 when everyone is already an adult). The disadvantage is that the NELS is intended to provide a nationally representative sample and so the number of respondents who live in counties with 1960 poverty rates "close" to the OEO cutoff is fairly limited.

The original sample employed a two-stage sampling design, with 1,052 schools selected in the first stage and 26 students per school selected in the second.[10] Base year participants were selected to participate in follow-up surveys in part on the basis of the number of other base-year NELS participants in the student's school at the time; dropouts were also retained in the sampling frame (U.S. Department of Education, 1994). The Department of Education provides weighting variables that account for the probability of participation in the base-year and follow-up surveys, as well as school administrator and student survey non-response (U.S. Department of Education, 1994). Our descriptive and main findings below are all calculated using these sampling weights.

The key explanatory variable of interest is whether the respondent has participated in Head Start, which is reported at baseline by the child's parent rather than taken from administrative records. The problem of recall errors with the NELS may be exacerbated by the fact that parents of eighth graders are asked to report on their child's involvement in Head Start or other preschool programs nearly 10 years earlier (1977-1979). Nevertheless the Head Start participation rate suggested by the NELS data (13 percent) is generally consistent with that implied by other data.[11]

[10] Excluded from the NELS sample in 1988 were students with mental handicaps, physical or emotional problems, and inadequate command of the English language. In most cases, 24 of the 26 students per school included in NELS were randomly sampled, while the other two students were selected from among the Hispanic and Asian Islander students (U.S. Department of Education, 1994).

[11] This figure is similar to that reported by parents in the 1979 National Longitudinal Survey of Youth Child-Mother file (NLSCM), in which 14 percent of white and 32 percent of African-American children participated in Head Start

The other key explanatory variable for our analysis comes from the NELS respondent's county of residence, which we identify using information on the location of the school that each respondent attended in 8th grade in 1988. For students in public schools we identified counties by matching NELS school identifiers with information from the Common Core of Data, while for private-school students we identified the counties of their schools from the 1988 Private School Survey. Through this procedure we were able to identify the 1988 county of residence for 96% of base-year NELS respondents. The fact that we can only observe where NELS respondents lived in 8th grade, but not when they were actually of Head Start age (3 or 4 years old), is a limitation of the NELS data and is likely to introduce some noise into our identification strategy.

Our main measures of educational attainment and labor market outcomes come from responses to the 2000 follow-up survey, by which time respondents were around 25 years of age. Our measures of academic achievement come from standardized tests administered in 1988.[12] We also focus on self-reported arrests collected by self-administered pencil-and-paper questionnaires in the 1990 and 1992 interviews (U.S. Department of Education, 1994).[13] Students in school are asked about arrests during the past academic term, while dropouts are asked about the last academic term spent in school.[14] This raises the possibility that students and dropouts may be reporting on arrests at a different point in calendar time, which is of some concern given that crime rates were changing quite dramatically during the 1990's (Levitt, 2004). This is likely to be more of a problem with the 1992 than the 1990 NELS survey.[15]

(Currie and Thomas, 1995), and to figures reported in the PSID suggesting participation rates of between 10 and 12 percent for children born in the 1970's (Garces, Thomas and Currie, 2000). Head Start participation in the NELS is also consistent with the figures implied by administrative data collected by the Federal government: If we assume each Head Start participate is in the program for only one year, then around 12 percent of children four years old in 1978 were enrolled in Head Start. In 1978, the year in which the average NELS child would have been four years of age, a total of 337,531 children participated in Head Start (GAO, 1981). Since each cohort under the age of 5 in 1978 averaged around 3 million children (U.S. Census Bureau, 1979), the ratio of program participants to children age four was on the order of 0.11. Put differently, if children were only allowed to participate in Head Start at age four, the available administrative data would suggest that 11 percent of the cohort of children enrolled in eighth grade in 1988 (the NELS cohort) participated in Head Start.

12 We only use achievement tests for the base year because follow-up achievement test results are missing for an unusually large share of dropouts in later waves (U.S. Department of Education, 1994, Grogger and Neal, 2000).

13 Self-administered questionnaires seem to yield somewhat lower estimates for the prevalence of sensitive behaviors than computer-assisted methods (Turner *et al.*, 1998).

14 The arrest rates reported by NELS teens are quite similar to those implied by national arrest data. For example, in the first NELS follow-up in 1990 (when most students were 15 or 16), 6 percent of male students had been arrested during the previous term. By comparison, data from the Federal Bureau of Investigation's Uniform Crime Report system suggest that 10 percent of teens age 15 and 12 percent of teens age 16 were arrested during 1990 (FBI, 1991). Since the NELS question covers half a school year, and a fair proportion of juvenile criminal activity may occur over the summer, the NELS results seem reasonable.

15 This is for two reasons. First, the fraction of NELS respondents who have dropped out is much lower in 1990 than 1992. Second, those who have dropped out are likely to have dropped out more recently prior to the interview for the 1990 than the 1992 waves, suggesting that in the 1990 interview a larger fraction of dropouts will be

In principle an alternative micro-data source for our project would be the Panel Study of Income Dynamics (PSID), which in 1995 asked all respondents ages 18 to 30 about their participation in Head Start and other preschool programs and serves as the data source for GTC. One advantage of the PSID relative to the NELS is the ability to identify where respondents live when they are actually of Head Start age rather than at some later point in time. In practice the PSID, which like the NELS is intended to be representative at the national but not state or local levels, appears to provide an unrepresentative draw of people in the treatment counties just above the OEO cutoff. The result is that among PSID sample members who answered the 1995 Head Start question, we do not see the discontinuity in Head Start participation that we observe in the NELS and the county-level federal spending data. For this reason we do not use the PSID to directly estimate the effects of Head Start on outcomes using the OEO discontinuity. However we do exploit the availability of geo-coded data for the larger PSID sample in all poor counties to explore the problem of selective across-county migration with our NELS and county-level data.[16]

III. Head Start

In what follows we discuss how Head Start might affect long-term outcomes, and then review the features of the program's rollout in 1965 that is the key to our research design.

A. Program Objectives

While Head Start is widely perceived to be an educational intervention, the program as originally conceived and implemented was more than that: Head Start is (or at least was) also a health program, a nutritional program, a social services program, a parenting program, and even a jobs program. Despite Head Start's broad objectives, until recently most evaluations focused only on academic outcomes.

Perhaps the most obvious way in which Head Start may improve children's outcomes in the short- and long-term is by improving the ability of children to take advantage of the educational opportunities provided to them during their K-12 schooling careers. When children show up for first grade there is already a gap in academic achievement test scores between rich and poor children and between minorities and whites (Phillips *et al*., 1998). Since learning is a cumulative process, there is the possibility that initial improvements during early childhood

---

reporting on the same calendar period as are enrolled students.

16 An alternative explanation for the difference between the PSID and NELS in documenting a Head Start discontinuity at the OEO cutoff is that the Head Start variable with the former may suffer from relatively greater measurement error. The reason is that while the NELS asks parents of potential Head Start participants to report on program involvement 10 years after their children would have been age-eligible to participate, the PSID asks people to self-report on whether they were in Head Start from 15 to 25 years after they would have been of Head Start age. We believe that sampling variability rather than measurement error is more likely to explain the PSID pattern of Head Start participation around the OEO cutoff because we do not see any difference in outcomes for PSID respondents at the cutoff.

caused by Head Start's educational or other services could translate into long-term gains. Consistent with this intuition, previous research finds children's academic outcomes are most strongly correlated with family income during early childhood (Duncan *et al*., 1998).

Head Start could also improve children's educational outcomes by improving parenting practices and family income. Head Start was originally part of OEO's Community Assistance Program (CAP), which was intended in part to provide jobs to poor families and involve them in the administration of local anti-poverty programs. As a result, from the first year of Head Start's existence many of the parents of children were employed in Head Start centers (Zigler and Valentine, 1979). Even parents who were not employed by the program might have experienced an increase in employment or disposable income if Head Start serves as a form of subsidized child care. All parents of participating children may have been exposed to new parenting styles and practices, or may have experienced reductions in stress due to the program's social services.

From the beginning, Head Start administrators argued that the non-academic gains from the program may be at least as important as any changes in educational outcomes. Particularly relevant from a societal perspective may be the program's impacts on criminal behavior, given the enormous costs of crime to society – estimated to be on the order of $1 trillion per year (Anderson, 1999).[17] Head Start's effects on delinquency and crime may be non-trivial, given the common finding that about 6 percent of each birth cohort – presumably drawn disproportionately from the set of low-income families eligible for Head Start – commits about half of that cohort's criminal activity (Tracy, Wolfgang and Figlio, 1990).

Despite the numerous plausible mechanisms through which Head Start could produce lasting benefits, previous studies often find evidence of "fade out" in program gains, particularly for African-American children (Currie and Thomas, 1995). However most of the evidence on the effects of Head Start comes from non-experimental comparisons of program participants with non-participants. These comparisons are of course susceptible to bias from unmeasured variables associated with both program participation and children's outcomes (see Currie, 2001). Even the best studies that rely on within-family across-sibling comparisons may be susceptible to such bias to some degree.

To address the limitations of existing Head Start research, the federal government recently funded a randomized experimental evaluation of the program (Haskins, 2004). However a study of children currently participating in Head Start will obviously not yield information about the program's long-term outcomes for many years.

In sum, there are theoretical reasons to suspect that Head Start could produce lasting

[17] Since aggressive children are more likely to become violent teens and adults (Reiss and Roth, 1993), any effects of Head Start on children's behavior may translate into long-term reductions in crime. Improvements in parenting practices that result from Head Start may also have some desirable effect on later criminal behavior (Loeber and Southamer-Loeber, 1986, Buka and Earls, 1993). Any improvements in long-term educational and economic outcomes may indirectly reduce anti-social behaviors by increasing the opportunity costs of crime.

effects on the outcomes of low-income children, but a somewhat limited body of empirical research that raises questions about whether Head Start accomplishes this goal in practice. The key to our own study is the "natural experiment" generated by the program's rollout.

B. Program Rollout

Planning for Project Head Start began in the fall of 1964 as part of CAP. The challenge for OEO administrators in the spring of 1965 was to publicize Head Start, encourage local organizations to submit proposals, review the proposals and fund enough local programs to launch Head Start in the summer of 1965 on the grand scale desired by President Lyndon Johnson – all within the span of several months.

Despite OEO's efforts to publicize the new Head Start program among local school principals, welfare administrators and public health officials, federal officials were concerned that in a nationwide grant competition many poor counties would be unable to develop acceptable proposals. Julius Richmond, national director of Project Head Start in 1965, noted that OEO administrations were "making a very determined effort to get the communities with greatest need in" (Gillette, 1996, p. 231). In response to this concern, Head Start associate director Jule Sugarman initiated an effort to generate applications from the 300 poorest counties in the U.S., which as noted above appear to have been identified using the 1960 poverty-rate calculations conducted by the Census Bureau for OEO in 1964.[18] Volunteers from the federal PMI program were provided with funding to travel to the selected counties on weekends during the spring of 1965, locate local actors who would be able to implement a Head Start program, work with them to develop a suitable proposal, fly the completed application back to Washington and then defend the proposal to OEO reviewers (see also Jones, 1979, p. 6-7).

In the end, Head Start grants were awarded to 240 of the 300 poorest counties, most of which were located in the South (GAO, 1981). Also noteworthy is the heavy concentration of Head Start funding nationwide in the most populous counties: While only 43% of all counties in the U.S. received some Head Start funding in 1968, fully 83% of the American population in 1968 was living in counties that received some Head Start funding.

To provide some additional sense of the geography of our "treatment" and "control" groups, one-third of the 300 poorest counties in 1960 were in Mississippi, Kentucky or Georgia. Almost all of the 300 poorest counties were in just ten states (Alabama, Arkansas, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, and Texas). These ten states also account for more than two-thirds of the 300 "control" counties (with 1960 poverty rates that rank from 301st to 600th in the U.S.), with most of the rest located in Florida, Oklahoma, Virginia or West Virginia. Put differently, most of the variation that we use to identify the effects of Head Start comes from differences in Head Start funding across very poor

18 As noted above, while the Head Start histories do not specify which of the two possible 1960 census measures were used to identify the 300 poorest counties – the proportion of a county's families with incomes under $3,000, or the poverty rate calculated by the Census Bureau in 1964 – the latter measure reveals a much sharper discontinuity in Head Start funding than the former.

counties within the South.

The resulting discontinuity in Head Start funding across counties with different 1960 poverty rates is at the heart of our research design. The other key to our study is that these across-county funding differences appeared to persist for many years, as a result of slower-than-expected growth in funding for Head Start and a "hold harmless" rule that prevented states from experiencing declines in Head Start funding levels from one year to the next (Jones, 1979). The persistence of the discontinuity in Head Start funding means that many cohorts of disadvantaged children were exposed to the "natural experiment" that we use to evaluate the program's effects.

Figure 1 provides an initial look at the Head Start funding discontinuity that resulted from the program's launch. In the top panel the solid line presents average Head Start funding per capita from the National Archives for 1968, calculated using a bin width of 4, while the bottom panel uses Head Start funding data for 1972 instead. In both panels the dashed line uses the same bin width and shows Head Start participation rates from the NELS, re-scaled to fit on the same graph as the spending data.

The key point of Figure 1 is the drop-off observed in the raw data at the OEO cutoff in 1968 and 1972 Head Start funding as well as NELS Head Start participation rates. Also remarkable is the overall similarity between the three series across counties, even though they come from two entirely different data sources and are measured at different points in time.[19] These similarities speak to the stability of the cross-sectional differences in county Head Start resources, particularly at the OEO cutoff, that we use to identify the program's impact.

IV. Empirical Approach

Our empirical strategy is to examine whether the discontinuity in Head Start funding and participation rates described above translate into discontinuities in long-term outcomes. We use the partially linear regression approach from Porter (2003), which allows us to non-parametrically control for other factors associated with funding that affect outcomes and vary smoothly across counties. We begin by discussing our main regression discontinuity estimation strategy and then discuss how we modify this approach to exploit within-county variation in Head Start exposure across different birth cohorts to account for other possible confounding factors such as selective migration.

A. Regression Discontinuity Estimates

Our analyses are conducted using county-level data, which is the geographic unit for which our Census and UCR crime data are reported. We aggregate the NELS data up to the

---

19 Of the 1,346 counties that received any Head Start funding in 1968, 72% received Head Start funding in 1972. Of the 1,084 counties that received Head Start funding in 1972, 90% had received Head Start funding in 1968. The set of counties funded in 1968 and 1972 may not overlap perfectly because of noise in the NARA data on federal spending (noted above), termination of OEO funding for some of the original Head Start programs, and the addition of new Head Start programs in some areas.

county level as well using the sampling weights, which is the simplest way within the partially linear regression framework to account for the non-independence of NELS observations living within the same county. That is, for each county (c) and NELS respondent (i) we calculate the average outcome within the county as $Y_c = \sum_c(Y_{ic}w_{ic}) / \sum_c(w_{ic})$, where w_{ic} represents the sampling weight for the survey wave from which we draw the accompanying outcome measure Y_{ic}, and (c) indexes the county in which each NELS respondent lives in 8th grade (about 10 years after they would have been of Head Start age). In the Census and UCR data, (c) indexes county of residence 1990, when we observe adult long-term outcomes. Below we return to the problem of unmeasured county of residence at Head Start age (3-5) for Census and NELS respondents.

Let P_c represent each county's poverty rate in 1960, and let the index (c) be defined over counties sorted in descending order by their 1960 poverty rate (so that c=1 is the poorest county and the OEO cutoff for Head Start grant-writing assistance occurs at c=300). Each county's 1960 poverty rate is a function of a set of "fundamental" factors p_c as well as a random component ε_c, as in equation (1). The provision of grant-writing assistance is a deterministic function of the county's 1960 poverty rate, as in equation (2), where P_{300}=59.2.

(1) $P_c = p_c + \varepsilon_c$

(2) $G_c = 1(P_c \geq P_{300})$

We can use the "sharp" regression discontinuity implied by (2) to estimate discontinuities in outcomes at the OEO cutoff (Trochim, 1984), which is in some sense like an "intent to treat" effect (ITT) – the effect of offering local service providers assistance in securing Head Start funding. If the offer of grant-writing assistance has no effect beyond increasing the amount of funding, then we can calculate the effect on long-term outcomes per dollar of additional Head Start funding by dividing the ITT effect for some educational or other outcome by the ITT effect on Head Start funding.

Less clear is whether we can estimate the effects of attending Head Start. One problem is that (as discussed below) across-county variation in funding may influence spending per participant as well as overall participation rates. Estimating the effects of Head Start enrollment itself also requires the assumption that program effects on participants and non-participants alike are not related to the county's overall Head Start funding or participation rates. This "stable unit treatment value assumption" (SUTVA) may be violated if social interactions among children or parents affect children's long-term outcomes, in which case Head Start's impacts may be amplified by "social multipliers" (Glaeser, Sacerdote and Scheinkman, 2003).[20] For these reasons, we focus our analysis primarily on the "reduced form" ITT-style estimates for the overall discontinuity in outcomes at the OEO cutoff.

Given that the unit of treatment in this setup is the county, our empirical analysis focuses

20 For example, Head Start funding might affect the probability that a given classroom contains a "rotten apple" that disrupts everyone's learning (Lazear, 2001), or that parents of participating children learn about new parenting skills that they then share with others within their social networks.

on estimating the effect for the average county rather than the average child. In the next section we show that the results estimated for the average child (that is, weighting by county population) are qualitatively similar to our preferred (un-weighted) estimates.

Our main estimating equation is given by (3), where Y_c is the outcome for county c, $m(P_c)$ is an unknown smooth function of 1960 poverty levels, and α is the impact of grant writing assistance. We note that the effect that we seek to identify is the one relevant for the poorest counties with 1960 poverty rates near the OEO cutoff.

(3) $Y_c = m(P_c) + G_c\alpha + \upsilon_c$

Identification of the causal effects of Head Start grant-writing assistance – the ITT corresponding to a treatment of increased Head Start funding in the county – comes from assuming smoothness in potential outcomes near the OEO cutoff (Porter, 2003). It strikes us as plausible that in the absence of OEO's grant-writing assistance to the poorest 300 counties, outcomes would vary smoothly around the cutoff, particularly because this cutoff does not seem to have been used to distribute funding for other federal programs. Below we present empirical evidence on funding patterns for other federal programs that is consistent with this conclusion.

An alternative way to think about identification in this RD model comes from Lee (2003). If the probability density of ε_c (the stochastic component of each county's 1960 poverty rate) is continuous at the OEO cutoff, then the allocation of technical grant-writing assistance for Head Start, G, can in the limit be thought of as randomized in the neighborhood of P_{300}. This assumption strikes us as plausible given that each county's 1960 poverty rate was determined before the War on Poverty was launched, and OEO's decision to target grant-writing assistance on the poorest 300 counties seems to have been an unannounced, *ad hoc* decision made in the rush to launch a nationwide Head Start program within the span of a few months. There would appear to be little room for strategic behavior on the part of local officials, and little incentive for strategic behavior on the part of OEO officials (given that the concern in Spring 1965 was one of excess supply of federal funding rather than excess demand). So long as the mapping between P and Y is also smooth in the neighborhood of the OEO cutoff then potential outcomes will be independent of G given P, the necessary condition for identification (Hahn, Todd and van der Klaauw, 2001).

We estimate the parameter α from equation (3) using the partially linear regression model from Porter (2003). The approach is "partially linear" in the sense that we use a kernel estimator to non-parametrically model $m(P_c)$, which gives more weight to data points that are closer to the OEO cutoff and allows us to control for the variety of factors that vary across counties and affect long-term outcomes without having to impose strong functional-form assumptions. Following Porter we re-write equation (3) as follows:

(4) $(Y_c - G_c\alpha) = m(P_c) + \upsilon_c$

Our estimate for α comes from finding the value that minimizes the average squared deviation between the new dependent variable in equation (4) and the nonparametric estimate of $m(P_c)$. That is, our estimate for the change in the conditional expectation of Yc at P300 comes from choosing the value of α that minimizes the following value, where the summations for c, j and k are all from 1 to N:

(5) $\min \Sigma_c [Y_c - G_c\alpha - \Sigma_j w_j^c(Y_j - G_j\alpha)]^2$

where $w_j^c = K_h(P_c - P_j) / \Sigma_k K_h(P_c - P_k)$

To estimate (5) we use the Epanechnikov kernel, $K(z) = (.75)(1-.2z^2)/\surd 5$ for $|z| < /\surd 5$. For the estimates below that use county-level Census data we use a bandwidth of 2. Given that we have less information near the OEO cutoff with the nationally-representative NELS sample, we use a larger bandwidth (equal to 6) with those data. We chose these bandwidths in part by examining whether they produced balance on our background covariates, before we looked at results for our outcome measures of interest. We explore the sensitivity of our estimates to the choice of bandwidth in the next section.

B. Discontinuity in differences estimator

Application of the approach outlined above with our 1990 Census data raises a few additional concerns. First, most county-level outcomes in the standard Census files, such as the county's poverty rate, will reflect the outcomes of birth cohorts that were exposed to Head Start with older cohorts that were of Head Start age before the program was in existence. Second, there is also the potential problem of selective migration in or out of counties between the time people were of the ages to be Head Start eligible (3 or 4) and when we measure their outcomes during adulthood at least 15 years later.

To address both problems we exploit age-disaggregated variables available with the county-level data from the 1990 Census STF4 files. The problem of averaging together outcomes for treated and untreated cohorts within a county is the easiest to solve, since we can just conduct our analysis using data on average outcomes for just treated age groups. However, the fact that we also have data available for untreated (or at least only partially treated) cohorts means that we can also address the problem of variation across counties in migration patterns or economic or educational development. We do this by examining whether the difference in average outcomes within a given county between treated and untreated cohorts changes discontinuously at the OEO cutoff. To the extent to which differently aged cohorts experience similar migration patterns and other unmeasured county-level conditions, focusing on the discontinuity in differences helps account for other factors that affect outcomes and could in principle vary discontinuously at the cutoff.

More formally, let Y_{ac} represent the average outcome of cohort (a) within county (c). Let $\Delta Y_c = Y_{ec} - Y_{nc}$ represent the difference in average outcomes between a cohort that was exposed to Head Start, that is, of Head Start age after the program was in operation, Y_{ec}, and the average outcome of a non-exposed cohort who were 3-5 years of age before Head Start began in 1965,

Y_{nc}. In this case our regression discontinuity comes from choosing the value of β that minimizes the sum of squared deviations between the dependent variable in equation (6) and our non-parametric estimate of $m(P_c)$.

$$(6) \qquad (\Delta Y_c - G_c\beta) = m(P_c) + \eta_c$$

The problem of averaging over treated and untreated cohorts is not a problem in the NELS, given that by design everyone in the NELS sampling frame was of Head Start age after the program was in effect. However the problem of selective migration is still of some potential concern with the NELS, given that we first measure respondents' county of residence in 8th grade, about 8-10 years after they could have participated in Head Start. Unfortunately we cannot implement our discontinuity-in-differences research design with the NELS because we do not have comparable outcome data for untreated cohorts. However we can informally generate an estimate that is similar in spirit by estimating equation (5) for the educational attainment and earnings of the parents of NELS respondents, most of whom would have been of Head Start age before the program was in operation.

V. Findings

We begin this section by using the regression discontinuity method discussed above to estimate the discontinuity in Head Start funding and participation rates. We then present our main findings, which include evidence for Head Start effects on educational attainment in both our county-level and NELS datasets. We do not find statistically significant discontinuities at the OEO cutoff in labor market outcomes in the Census or NELS, or (in the NELS) in 8th grade reading or math achievement test scores. Finally, we show that both the NELS and county-level data are well balanced with respect to other background characteristics at the OEO cutoff, and that other forms of federal spending also appear to be similar. Furthermore, we do not find evidence for discontinuities in outcomes at an alternative "pseudo-cutoff" point at which there is no discontinuity in Head Start funding.

A. Results for Head Start Funding and Participation

Figure 2 presents estimates for the discontinuity in Head Start around the OEO cutoff. The top panel presents results for 1972 per capita Head Start funding across counties with different 1960 poverty rates.[21] The bottom two panels are participation rates in Head Start and other preschool programs by county using our sample of NELS respondents. All three of these panels and all of the results shown below employ the partially linear regression approach developed by Porter (2003) and discussed in the previous section. The vertical line in each graph at 1960 county poverty of 59.2 represents the OEO cutoff, which distinguishes the 300 poorest counties that received grant-writing assistance (to the right of the cutoff) from other counties. The smooth horizontal lines in each graph represent our estimates for m(P) from

21 In what follows we focus on the 1972 rather than 1968 funding figures from the National Archives because Head Start was closer to a year-round program by the early 1970s.

above. We also show the means and 95% confidence intervals for the raw data around our estimates for m(P) to provide readers with some sense for the underlying data, and at the top of each graph we report our estimate for the discontinuity (the estimate for α from equation 3 above and its accompanying t-statistic).

The top panel of Figure 2 shows that the discontinuity in Head Start funding across counties at the OEO cutoff is proportionately large, even though the standard error around our point estimate for alpha is also quite large.[22] The limit from the left at the OEO cutoff shows average 1972 Head Start spending of about $3.50 per capita, so that the discontinuity represents an increase in funding of about one-third.

The second panel of Figure 2 presents the estimated discontinuity in Head Start participation rates across counties for NELS respondents. Recall that we identify program participation from parent reports, and relate this to the 1960 poverty rate for the counties in which children were living in 8th grade in 1988. At the OEO cutoff the left limit indicates a Head Start participation rate of about 15 percent, so that the estimated discontinuity of .099 represents about a two-thirds increase in participation rates. The bottom panel of Figure 2 shows that there is no detectable discontinuity in participation in other forms of preschool.

Taken together Figure 2 provides suggestive evidence that the discontinuity in Head Start program funding across counties translates into both expanded enrollment and increased spending per program enrollee, given that the proportional increase in funding is not larger than the proportional increase in participation rates. This tentative conclusion should be tempered by the observation that the standard error around our estimated discontinuity in funding is very large and so we cannot rule out the possibility of some increase in spending per pupil at the cutoff. (The fact that the funding data are for 1972 and the NELS data capture participation rates in 1977 or 1978 does not seem like a big problem given that funding and enrollment rates overall were fairly flat during this period. See Haskins, 2004, Figure 1). Comparisons between panels B and C also suggest that our estimates below may be identifying the effects of increasing enrollments in Head Start compared to the counterfactual of time with parents or other forms of informal care, rather than center-based child care or other more formal preschool education programs.

B. Discontinuity in Long-Term Outcomes

Figure 3 presents our estimates from the 1990 Census for discontinuities in educational attainment (completion of high school or more, attendance of at least some college, or

---

22 One reason that the standard error for the estimated discontinuity may be so large is that the underlying data series might be quite noisy due to random measurement error. As noted above, we have tried to carefully "clean" the data on federal spending per county in the presence of a large number of undocumented data irregularities. Of the dozen years or so of federal spending data obtained from the National Archives, only the cleaned data for 1968 and 1972 could be made to match published figures at the state and national level for Head Start and overall federal spending. The fact that the data match at the aggregate level does not mean that there is not still a significant amount of measurement error.

completion of college or more) for four different age groups. The cohort of primary interest consists of people ages 18-24 in 1990, who were born 1966-1972 and so all came of Head Start age (typically between 3 and 5 years old) while the program was in existence. We call this the "directly treated" group. Also of some interest are people 25-34 in 1990, born 1956-1965. Because about one-third to one-half of this cohort might have been of Head Start age after the program was in operation, we call this the "partially directly treated" group. The set of people ages 35-54 in 1990 might include some older siblings as well as parents of Head Start participants. Head Start might produce positive spillover effects for siblings by improving parenting practices or household resources (Currie and Thomas, 1995), and has also been argued to have positive effects on parents as well, including the possibility of employment with the program. For both of these reasons we term people 35-54 the "potentially indirectly treated" group. People ages 55 and older in 1990 are unlikely to have been parents (much less older siblings) of Head Start participants,[23] and so we label this the "untreated" group.

Figure 3 shows very large discontinuities in 1990 in educational attainment for the directly treated group (18-24), equal to an increase of 4 percentage points in high school graduation rates (panel A) and 5.3 percentage points in college attendance (panel B). We do not find a discontinuity in college completion rates for this group, which makes sense given that most people in this group will be too young to have graduated from college.

We also find some evidence of statistically significant effects on high school completion and college attendance for the partially directly treated group (25-34) and the potentially indirectly treated group (35-54), which are typically about half the magnitude of the point estimates for our directly treated group. For the partially and indirectly treated groups we also see signs of some effect on college completion. On the other hand for our untreated group we find no evidence of a discontinuity in any of our educational attainment measures.

The pattern of estimates across age groups in Figure 3 is consistent with what we would predict if Head Start had a positive causal effect on educational attainment and produced positive spillover effects for older siblings and parents. A counter-explanation is that the jump in educational attainment at the OEO cutoff reflects the influence of other factors, such as a discontinuity in selective migration out of these counties. One argument against this alternative explanation for our findings is that we do not see discontinuities in educational attainment for our untreated group. Nevertheless a skeptical interpretation of Figure 3 might lead some readers to conclude that the effects on at least the partially directly treated and indirectly treated groups might reflect some spurious association with other factors.

Figure 4 presents estimates that try to account for the possibility of a discontinuity in omitted variables at the OEO cutoff by estimating the discontinuity in the difference in

---

23 In 1990 the oldest person who could have participated in Head Start would have been about 30 years of age (born 1960, and so five years old when Head Start went into operation in 1965). Data from the Vital Statistics for 1960 suggests that about one-half of all births that year were to parents 25 or older (55 or older in 1990), and only about one-quarter were to parents 30 or older (HEW, 1960). This means that most of the parents of Head Start participants born in 1960 or later were younger than 25 in 1960 and so of course younger than 55 in 1990.

educational attainment across age groups. Suppose that the discontinuity in educational attainment observed for the partially directly treated and indirectly treated groups reflect some spurious association rather than the causal effect of Head Start. We might expect that the effects of any such confounders might be similar for the directly treated group and the other groups, particularly the partially treated group given the similarity in ages (18-24 versus 25-34 in 1990). The difference across age groups within the same county may help take out the effect of county-specific omitted variables that vary discontinuously at the OEO cutoff, so that the discontinuity across counties in the difference across age groups within counties may provide a cleaner estimate of Head Start's impact.

The results in Figure 4 show that if we compare the directly to the partially treated group (the one closest in age), the estimated Head Start effect on high school completion and college attendance is about half as large as the main effect in levels. Given that the main effect for our oldest untreated group (55 plus) was around zero in all cases, contrasting the directly treated group to the untreated group yields estimates for the former that are of about the same magnitude as the directly treated main level effects.

Figure 5 provides qualitatively similar findings for NELS respondents to the 2000 wave of that survey, when most respondents would have been around 25 years of age. The top row of Figure 2 shows positive discontinuities in high school completion and college completion (in favor of the treatment counties) equal to 15 and 13 percentage points, respectively, which are statistically significant at the 10 percent cutoff.

The NELS and Census data are also consistent with one another in not yielding evidence of discontinuities in labor market outcomes. The bottom row of Figure 5 shows the results from the NELS for whether the respondent was working full time at the time of the 2000 survey, as well as log annual earnings for 1999. In the 1990 Census we find almost no statistically significant discontinuities in either unemployment rates or poverty status for any of our age groups. We return to the apparent discrepancy in findings for education versus labor market outcomes in the final section of the paper.

The top row of Figure 6 shows that there is no statistically significant discontinuity at the OEO cutoff in 8th grade reading or math achievement scores in the NELS, and in fact the point estimates are of the opposite sign of what we would expect if Head Start increased test scores. In contrast the point estimates for self-reported arrests in the 1990 and 1992 NELS surveys are in the direction of a Head Start effect to reduce long-term criminal activity. Given that schooling and criminal activity are negatively related (Lochner and Moretti, 2004), we would expect Head Start to reduce arrests if the program also increases educational attainment. Interestingly the NELS point estimates for arrest probabilities are consistent with what we would expect given our findings for schooling in the NELS and Lochner and Moretti's (2004) estimates for the effects of schooling on crime, although our NELS arrest findings are not statistically significant.[24]

24 Lochner and Moretti (2004, p. 175) find that each additional year of schooling reduces arrest rates by 11 to 16 percent (Table 10). In 1990 NELS respondents were about 15 years of age. In 1997 (the earliest year for which data were conveniently available) this age group had an arrest rate of about 0.1 per capita. If we interpret the top panel of

C. Specification and Sensitivity Checks

One natural concern is the possibility that the federal government may have disproportionately directed to the 300 poorest counties funding for a variety of other programs. In this case we will not be able to distinguish between the effects of Head Start funding and the effects of funding for other programs.

The top left panel of Figure 7 shows that the estimated discontinuity in other forms of federal social spending in 1972 at the OEO cutoff is very small (equal to \$3.60 in 1972 dollars, or less than 1% of the value of the left limit) and not statistically significant. The top right panel initially suggests that there might be some proportionately large albeit not statistically significant difference in other forms of (non-social) federal spending, even after we delete the four outlier counties with non-social federal spending in excess of \$10,000 per capita. The two bottom panels of Figure 7 show that a histogram of the raw data (with or without outliers included) the trend in other federal spending is quite flat, suggesting that the apparent discontinuity in such funding from our partially linear estimator may be an artifact of how this procedure gives influence to counties with unusually high levels of per capita spending. In the next version of the paper we will explore whether counties with such high funding levels for other federal programs are due to things like military bases located in low-population counties.

An alternative concern is that the estimated discontinuities in educational attainment and arrests reflect some imbalance in other county characteristics at the OEO cutoff. In Table 1 we show that for most background characteristics the NELS is quite balanced, as are county-level Census data from 1990, 1980, 1970, 1960 and 1950. Since a particular concern is the possibility of selective mobility over time out of the treatment counties, Table 1 also shows results from the 1980 and 1990 Censuses for the percent of county residents who were living in the same county 5 years ago. (The Census did not ask these questions before 1980). We find a small difference in population mobility in both years although only the difference for 1990 is statistically significant.

Another way to address the possibility of selective migration is to draw on the geo-coded version of the PSID. As discussed above, the PSID is not useful for estimating the effects of Head Start because the sample appears to include a "bad draw" with respect to Head Start enrollment rates in counties with 1960 poverty rates just above the OEO cutoff. However we can aggregate the PSID data up further and exploit the longitudinal address information to explore mobility patterns. Of those PSID respondents for whom we have address data at both age 3 and 18 between 1968 and 1992, 71% were in the same county at ages 3 and 13 (the first

---

Figure 5 as suggesting an increase in educational attainment of around four years at the OEO cutoff for about 15% of the sample, then the implied reduction in arrest probabilities of about 3 percentage points in the bottom left panel of Figure 6 would seem to be consistent with the Lochner and Moretti estimate, given that at least part of their estimate comes from reductions in the intensive rather than extensive margin of criminal offending and that Head Start is likely to enroll many of the 6% of each birth cohort that accounts for about 50% of each cohort's total criminal activity (Tracy, Wolfgang and Figlio, 1990).

age at which we capture addresses in the NELS) and 66% were in the same county at 3 and 18. These figures are only slightly higher for those people living in the poorest counties. About 60 to 65% of people who were living in the poorest 600 counties at age 13 or 18 were living in the same counties when they were of Head Start age (3). More generally movers and stayers appear to have quite similar outcomes on average, at least in the national sample; larger differences are observed in the poorest counties, although the sample sizes here are quite small. Within our national PSID sample of 10 birth cohorts, age is not a significant predictor of mobility, suggesting that our discontinuity-in-differences estimator that focuses on within-county across-age group differences may take out most or all of any confounding effect from selective migration. Appendix A discusses these results in more detail.

A final way to check to see whether our results are spurious or instead may reflect the effects of Head Start is to examine whether we see discontinuities in long-term outcomes at other "pseudo-cutoffs." We must be careful in conducting this sort of specification check because as shown in Figures 1 and 2, there are significant differences in Head Start funding and participation rates across counties. If we arbitrarily choose a cutoff where there is an actual program funding difference we might detect in part the actual effects of Head Start funding. Put differently, we believe that the discontinuity in Head Start funding around the OEO cutoff provides us with a useful source of variation given the institutional details surrounding the program's launch in 1965. But we cannot rule out the possibility that differences in outcomes in other places where there are jumps in Head Start funding might not also reflect Head Start effects.

Table 2 shows the results of generating new estimates at a pseudo-cutoff equal to 1960 county poverty rate of 40%. This cutoff was chosen because we do not see evidence of a discontinuity in 1972 Head Start spending per capita here. We should note that we chose this cutoff before looking at our county-level Census outcome data or any of our NELS data. We find that only 1 out of 21 point estimates is statistically significant at the 5% cutoff, while 3 out of 21 are significant at the 10% cutoff. Moreover the pattern of these significant point estimates are quite different from those derived at the actual OEO cutoff. The only significant point estimate in the NELS is for high school completion and is equal to negative 8 percentage points ($p<.10$), the opposite of what we find at the OEO cutoff. In the Census data the few statistically significant differences that do arise are, unlike at the OEO cutoff, not concentrated among the directly treated group or among our high school completion or college attendance measures.

When we re-calculate all of our estimates weighting by county population, which provides us with information about the effect on the average person rather than the average county, the results for the directly treated cohort in the Census and in the NELS data are at least as strong as those shown above (in terms of the absolute magnitude of the point estimates and their size in relation to the standard errors). However the weighted estimates show somewhat more pronounced discontinuities in educational outcomes for the directly treated group at the pseudo-cutoff used in Table 2. To the extent to which this serves as a diagnostic test on our model specification, this finding provides further empirical justification for preferring the un-weighted to the weighted estimates.

VI. Conclusions

In this paper we use data from the 1990 census and the NELS to examine the long term impact of the Head Start program. In particular, we exploit a discontinuity with regard to the 1960 poverty rate in the federal government's provision of Head Start grant writing assistance. We document that counties just above the threshold had larger levels of Head Start funding and head start enrolment rates. We also document that there is a corresponding discontinuity in some important long run outcomes, mainly educational attainment. However, the increased Head Start funding does not appear to lead to improved test scores or to improved employment outcomes.

Several considerations suggest that these results are reasonable. First, we find a striking similarity in the qualitative story emerging from two distinct data sets (NELS and the 1990 Census): an improvement in educational outcomes, but not for labor market outcomes. Second, we find stronger evidence for effects on behavioral outcomes such as educational persistence than for academic achievement test scores, which seems consistent with studies of other early childhood interventions (see Donohue and Siegelman, 1998, p. 21). Third, we find the strongest effects for the groups that should be most affected (younger adults in 1990), and no effects when we look for groups that should not be affected (the elderly in 1990). In the NELS data this corresponds to effects on educational attainment of students exposed to Head Start but not for their parents. Fourth, there is little evidence of a discontinuity at the OEO cutoff in other federal spending, particularly for other social programs. Finally, we do not find a similar pattern of discontinuities in educational outcomes for "treated" cohorts at a pseudo-cutoff where there is no significant discontinuity in Head Start funding.

Arguably the most important concern with our findings comes from the possibility of selective migration, given that we first identify county of residence at age 13 for NELS respondents and at adulthood with our Census data. While we have no way of directly assessing the degree to which selective migration is a problem, several pieces of indirect evidence suggest that this may not impart a large bias to our estimates. In the Census data there is only a small discontinuity in the fraction living in the same county five years ago in the 1980 and 1990 Censuses (negative 2 percentage points for the treatment compared to control counties). Almost all of the extra in-migrants in the treatment counties would have to be high school dropouts to explain away our findings.[25] The PSID also suggests that there are only small differences in out-migration across counties above and below the OEO cutoff, although the sample is quite small in these poorest counties (one reason we do not use the PSID to directly estimate Head Start's effects at the OEO cutoff). The national sample of PSID respondents suggests there is little difference in adult outcomes between those who stay in the same county from age 3 to 13 or 18 versus those who leave, and that migration patterns do not appear to be systematically related to

25 Figure 3, panels A and B show that the effect on high school completion is equal to 4 percentage points for those 18-24, 1.7 points for those 25-34, and 1.5 points for those 35-54. Assume that these three groups account for, say, half of the population, and that the difference in migration rates lead to 4 percentage points more in-migrants in the treatment than control counties. If the in-migrants are evenly distributed across age groups then they would all need to be dropouts in order to explain away the 18-24 effect.

age. This last finding suggests that our discontinuity-in-differences estimator that focuses on the difference in outcomes across age groups within counties may help account for much of any selective-migration bias that may arise.

The increase in Head Start enrollment seems to come at the expense of informal care rather than private or other sources of formal preschool education or child care. Put differently, we believe that we are measuring the impact of attending Head Start compared with parental (or other informal) care – we measure no impact on other preschool attendance near the poverty cutoff in the NELS.

Are the magnitudes of the effects reasonable? On the one hand, the discontinuities in outcomes are large relative to the discontinuities in at least HS participation in the NELS. For example, counties above the poverty threshold have a 10 percentage point higher participation rate; they also have a 15 percentage point higher high school graduation rate and a 13 percentage point advantage in college completion. This may suggest either that there is a social multiplier at work, consistent with the results found in GTC, or that there may be meaningful differences in Head Start spending per participant that we are unable to measure.

What do the results imply about the relative benefits and costs of the Head Start program? Figure 2 suggests that the difference in Head Start funding at the OEO cutoff is equal to about \$1 per capita in 1972 dollars. If we assume that about 1.7% of the of the population is age-eligible for Head Start at any point in time (assuming a single birth cohort is age-eligible at a point in time, that children participate for only one year, and a uniform age distribution up to 60), then the funding difference per age-eligible person (in 2004 dollars) is about \$150.[26]

On the benefit side, one of the puzzles is why the discontinuities in educational attainment at the OEO cutoff estimated in the Census and NELS data do not translate into improved labor market outcomes. One possibility may be that the cohorts directly treated by Head Start are still too early in their working careers for the increases in schooling to yield detectable differences in labor market outcomes.[27] An alternative possibility could be weak labor demand in these high-poverty counties combined with some friction that prevents people from seeking out better job opportunities elsewhere, as in the spatial mismatch hypothesis.[28]

---

26 To put this figure in perspective: If we assume participation rates in the treatment counties at the cutoff are around 15% (as in the NELS) and that average spending per participant is around \$1,500 (Haskins, 2004 cites a figure of \$1,380 in 2002 dollars for the program in 1966), then this funding difference could be used to increase participation rates by about 10 percentage points (as the NELS suggests) or increase funding per pupil by about two-thirds.

27 For example, Card (1999, p. 1805) shows that log hourly wages for college graduates are much closer to those of high school graduates at age 25 than at later ages. Moreover some people in our directly treated Census group (18-24) or even the NELS sample may still be in school at the time we measure labor market outcomes.

28 The spatial mismatch hypothesis originates with Kain (1968), and suggests that minority workers in the inner-city have difficulty following job movement to the suburbs because of racial discrimination in the housing market. Limited access to informal social networks that can provide information about suburban jobs is also sometimes offered as a mechanism. A recent review suggests that most studies in this literature find evidence for spatial mismatch (Ihlanfdfeldt and Sjoquist, 1998), although the hypothesis has its skeptics (for example Ellwood, 1986 and

Suppose that the estimated reduction in arrests from the NELS data (which are not statistically significant) are "real," consistent with the finding from Lochner and Moretti (2004) that increases in educational attainment lead to reductions in criminal behavior. The point estimates shown in the bottom row of Figure 6 if taken literally suggest that the discontinuity in Head Start funding could reduce arrest rates by perhaps half. The magnitude of this finding is not implausible if we remember that cohort studies suggest that 6% of each birth cohort account for about 50% of the cohort's total criminal activity (Tracy, Wolfgang and Figlio, 1990), and that Head Start presumably disproportionately draws participants from this high-risk sub-group.

The next version of this paper will complete this benefit-cost analysis assuming different lengths of people's "criminal careers," as well as that the reduction in arrests is distributed across crime categories in proportion to juvenile arrests across groups and under different interest rate assumptions. We also hope to complement the NELS arrest rate findings with analysis of county-level age-specific arrest data from the FBI's UCR system. But in the meantime we simply note that to the degree to which any increase in educational attainment from Head Start translates into fewer arrests (even if not into improved labor market outcomes), the program could plausibly pass a benefit-cost test given the substantial costs of crime to society each year. This tentative conclusion seems even more likely given Currie's (2001, p. 231) observation that the child-care benefits that Head Start provides to low-income mothers may be worth perhaps one-fifth to one-third of the program's cost.

Jencks and Mayer, 1990) and the most recent study of the Moving to Opportunity (MTO) residential mobility experiment finds little evidence for labor market effects (Kling *et al*., 2004).

References

Anderson, David A. (1999) "The Aggregate Burden of Crime." *Journal of Law and Economics*. 42(2): 611-642.

Angrist, Joshua D. and Victor Lavy (1999) "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics*. 114: 533-575.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin (1996) "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association*. 91(434): 444-455.

Aughinbaugh, Alison (200?) "Does Head Start Yield Long-Term Benefits?" Working Paper, Washington, DC: U.S. Bureau of Labor Statistics.

Barnett, W. Steven (1992) "Benefits of Compensatory Preschool Education." *Journal of Human Resources*. 27(2): 279-312.

Bloom, Benjamin S. (1964) *Stability and Change in Human Characteristics*. New York: John Wiley and Sons.

Buka, Stephen and Felton Earls (1993) "Early Determinants of Delinquency and Violence." *Health Affairs*. Winter. 46-63.

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.

Card, David (1999) "The Causal Effect of Education on Earnings." In the *Handbook of Labor Economics, Volume 3A*. Edited by Orley Ashenfelter and David Card. Amsterdam: Elsevier. pp. 1801-1864.

Citro, Constance F. and Robert T. Michael (1995) *Measuring Poverty: A New Approach*. Washington, DC: National Academy Press.

Cohen, Mark (1998) "The Monetary Value of Saving a High Risk Youth," *Journal of Quantitative Criminology*. 14(1): 5-33.

Coleman, James S. (1975) "Comment on David K. Cohen, 'The Value of Social Experiments.'" In *Planned Variation in Education: Should We Give Up or Try Harder?* Edited by Alice M. Rivlin and P. Michael Timpane. Washington, DC: Brookings Institution Press. pp. 173-175.

Coles, Robert (1969) "Rural Upheaval: Confrontation and Accomodation." In *On Fighting Poverty: Perspectives from Experience*. James L. Sundquist (Ed.) New York: Basic Books. pp.

103-126.

Currie, Janet (2001) “Early Childhood Education Programs.” *Journal of Economic Perspectives*. 15(2): 213-238.

Currie, Janet and Duncan Thomas (1995) “Does Head Start Make a Difference?” *American Economic Review*. 85(3): 341-364.

Currie, Janet and Duncan Thomas (2000) “School Quality and the Longer-Term Effects of Head Start.” *Journal of Human Resources*. 35(4): 755-774.

Donohue, John J. and Peter Siegelman (1998) “Allocating Resources Among Prisons and Social Programs in the Battle Against Crime.” *Journal of Legal Studies*. 27: 1-43.

Duncan, Greg J., Jeanne Brooks-Gunn, J. Yeung, and J. Smith (1998) “How much does childhood poverty affect the life chances of children?” American Sociological Review. 63: 406-423.

Ehrlich, Isaac (1996) “Crime, Punishment, and the Market for Offenses.” *Journal of Economic Perspectives*. 10(1): 43-68.

Entwisle, Doris R., Karl L. Alexander, and Linda Steffel Olson (1997) *Children, Schools, and Inequality*. Boulder, CO: Westview Press.

Freeman, Richard B. (1996) “Why Do So Many Young American Men Commit Crimes and What Might We Do About It?” *Journal of Economic Perspectives*. 10(1): 43-68.

Freeman, Richard B. and William M. Rodgers (1999) “Area Economic Conditions and the Labor Market Outcomes of Young Men in the 1990's Expansion.” National Bureau of Economic Research Working Paper # 7073.

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

General Accounting Office (1981) *Head Start: An Effective Program But the Fund Distribution Formula Needs Revision And Management Controls Need Improvement*. Washington, DC: General Accounting Office Report HRD-81-83.

Gillette, Michael L. (1996) *Launching the War on Poverty: An Oral History*. New York: Twayne Publishers.

Glaeser, Edward L., Bruce Sacerdote and Jose Scheinkman (2003) “The Social Multiplier.” *Journal of the European Economic Association*. 1(2-3).

Guryan, Jonathan (2001) "Does Money Matter? Regression-Discontinuity Estimates from Education Finance Reform in Massachusetts." National Bureau of Economic Research Working Paper 8269.

Hahn, Jinyong, Petra Todd and Wilbert Van der Klaauw (1999) "Evaluating the Effect of An Antidiscrimination Law Using a Regression-Discontinuity Design." National Bureau of Economic Research Working Paper 7131.

Hahn, Jinyong, Petra Todd and Wilbert Van der Klaauw (2001) "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica*. 69(1): 201-209.

Harmon, Carolyn and Edward J. Hanley (1979) "Administrative Aspects of the Head Start Program." In *Project Head Start: A Legacy of the War on Poverty*. Edited by Edward Zigler and Jeanette Valentine. New York: Free Press. pp. 379-398.

Haskins, Ron (2004) "Competing Visions." *Education Next*. 4(1): 26-33.

Heckman, James J. (1999) "Doing it right: Job training and education." *The Public Interest*. Spring 1999. 86-107.

HEW (1960) *Vital Statistics of the United States, 1960. Volume 1: Natality*. Washington, DC: United States Department of Health, Education and Welfare.

Ihlanfeldt, Keith R., and David J. Sjoquist (1998) "The Spatial Mismatch Hypothesis: A Review of Recent Studies and their Implications for Welfare Reform." *Housing Policy Debate*. 9(4): 849-892.

Imbens, Guido and Joshua Angrist (1994) "Identification of Local Average Treatment Effects." *Econometrica*. 62: 467-475.

Jacob, Brian A. and Lars Lefgren (2001a) "The Impact of Teacher Training on Student Achievement: Quasi-Experimental Evidence from School Reform Efforts in Chicago." Working Paper, John F. Kennedy School of Government, Harvard University.

Jacob, Brian A. and Lars Lefgren (2001b) "Remedial Education and Student Achievement: A Regression-Discontinuity Analysis." Working Paper, John F. Kennedy School of Government, Harvard University.

Jones, Jean Yavis (1979) *The Head Start Program – History, Legislation, Issues and Funding, 1964-1978.* Washington, DC: Congressional Research Service Report 79-14 EPW.

Kain, John F. (1968) "Housing Segregation, Negro Employment and Metropolitan Decentralization." *Quarterly Journal of Economics*. 82: 175-197.

Kling, Jeffrey R., Jeffrey B. Liebman, Lawrence F. Katz and Lisa Sanbonmatsu (2004) "Moving to Opportunity and Tranquility: Neighborhood Effects on Adult Economic Self-Sufficiency and Health from a Randomized Housing Voucher Experiment." Princeton University Industrial Relations Working Paper.

Lazear, Edward P. (2001) "Educational Production." *Quarterly Journal of Economics*. 116(3): 777-803.

Lee, David S. (2003) "Randomized Experiments from Non-Random Selection in U.S. House Elections." Working Paper, Department of Economics, University of California at Berkeley.

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): 155-189.

Loeber, Rolf and Magda Stouthamer-Loeber (1986) "Family Factors as Correlates and Predictors of Juvenile Conduct Problems and Delinquency." In *Crime and Justice: An Annual Review of Research*, M. Tonry and N. Morris (Eds.). Chicago: University of Chicago Press. pp. 29-149.

Maltz, Michael (1999) *Bridging Gaps in Police Crime Data (NCJ 1176365)*. Washington, DC: Bureau of Justice Statistics.

Maltz, Michael and Joseph Targonski (2002) "A Note on the Use of County-Level UCR Data." *Journal of Quantitative Criminology*. (September). 297-318.

Messner, Steven F., Luc Anselin, Darnell F. Hawkins, Glenn Deane, Stewart E. Tolnay, and Robert D. Baller (1998) Codebook for National Data Set (1960-1990). National Consortium on Violence Research Working Paper, presented at the November, 1998 meetings of the American Society of Criminology, Washington, DC.

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

Moffitt, Robert A. (2001) "Policy Interventions, Low-Level Equilibria, and Social Interactions." In *Social Dynamics*, Edited by Steven N. Durlauf and H. Peyton Young. Washington, DC: Brookings Institution Press. pp. 45-82.

Nagin, Daniel S. and Richard E. Tremblay (1999) "Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency." *Child Development*. 79(5): 1181-1196.

Phillips, Meredith, Jeanne Brooks-Gunn, Greg J. Duncan, Pamela Klebanov, and Jonathan Crane (1998) "Family Background, Parenting Practices, and the Black-White Test Score Gap." In *The Black-White Test Score Gap*, edited by Christopher Jencks and Meredith Phillips. Washington, DC: Brookings Institution Press. pp. 103-145.

Phillips, Meredith, James Crouse and John Ralph (1998) "Does the Black-White Test Score Gap Widen After Children Enter School?" In The Black-White Test Score Gap, edited by Christopher Jencks and Meredith Phillips. Washington, DC: Brookings Institution Press. pp. 229-272.

Porter, Jack (2003) "Estimation in the Regression Discontinuity Model." Working Paper, Harvard University Department of Economics, draft date September 25, 2003.

Raphael, Steve and Rudolf Winter-Ebmer (2001) "Identifying the Effect of Unemployment on Crime," *Journal of Law & Economics,* 44(1): 259-284.

Reiss, Albert J. and Jeffrey A. Roth (1993) *Understanding and Preventing Violence*. Washington, DC: National Academy Press.

Reynolds, Arthur J., Judy A. Temple, Dylan L. Robertson, and Emily A. Mann (2001) "Long-term Effects of an Early Childhood Intervention on Educational Achievement and Juvenile Arrest." *Journal of the American Medical Association*. 285(18): 2339-2346.

Schonfeld, I., *et al.* (1988) "Conduct Disorder and Cognitive Functioning: Testing Three Causal Hypotheses." *Child Development*. 59: 993-1007.

Shonkoff, J.P. and D.A. Phillips (2000) *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Washington, DC: National Academy Press.

Solon, Gary (1992) "Intergenerational Income Mobility in the United States." *American Economic Review*. 82(3): 393-408.

Thistlewaite, D. and D. Campbell (1960) "Regression-Discontinuity Analysis: An Alternative to the Ex-Post Facto Experiment." *Journal of Educational Psychology*. 51: 309-317.

Tracey, Paul E., Marvin E. Wolfgang, and Robert M. Figlio (1990) *Delinquency Careers in Two Birth Cohorts*. New York: Plenum Press.

Trochim, W. (1984) *Research Design for Program Evaluation: The Regression Discontinuity Approach*. Beverly Hills, CA: Sage Publications.

Wiersema, Brian, Colin Loftin and David McDowall (2000) "A Comparison of Supplementary Homicide Reports and National Vital Statistics System Homicide Estimates for U.S. Counties." *Homicide Studies*. 4(4): 317-340.

Wolff, Max and Annie Stein (1966) *Study I: Six Months Later, A Comparison of Children Who Had Head Start, Summer 1965, with Their Classmates in Kindergarten (A Case Study of Kindergartens in Four Public Elementary Schools, New York City)*. Washington, DC: Research and Evaluation Office, Project Head Start, Office of Economic Opportunity.

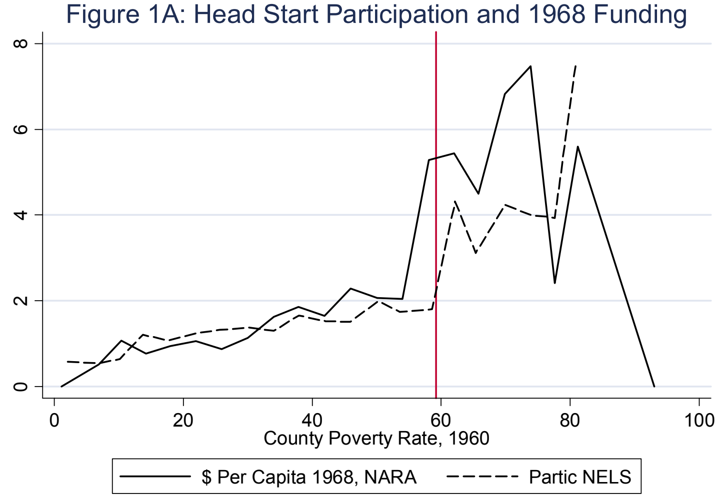
Yarmolinsky, Adam (1969) "The Beginnings of OEO." In *On Fighting Poverty: Perspectives from Experience*. James L. Sundquist (Ed.) New York: Basic Books. pp. 34-51.

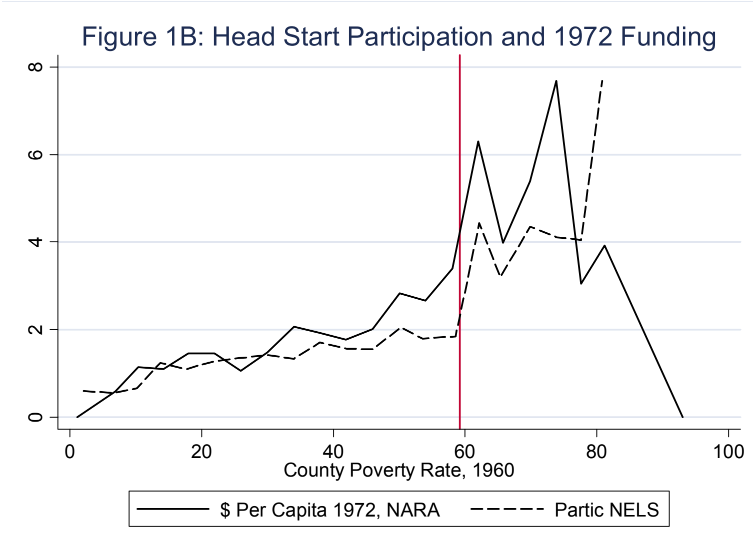
Zigler, Edward and Jeanette Valentine (1979) *Project Head Start: A Legacy of the War on Poverty*. New York: Free Press.

Zigler, Edward and Susan Muenchow (1992) *Head Start: The Inside Story of America's Most Successful Educational Experiment*. New York: Basic Books.

Zimmerman, David J. (1992) "Regression Toward Mediocrity in Economic Stature." *American Economic Review*. 82(3): 409-429.

Figure 1A: Head Start Participation and 1968 Funding

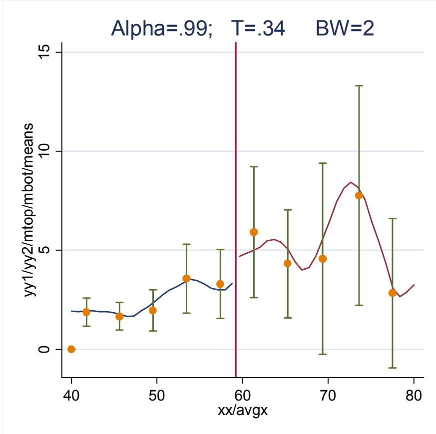




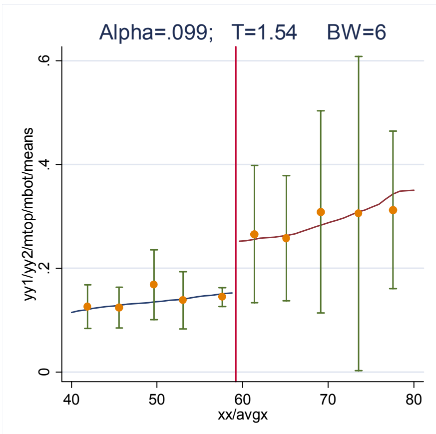
J:\julie2\u_graph_doug4.do & u_graph_doug68.do

Figure 2: Estimated Discontinuity in Head Start Funding & Participation

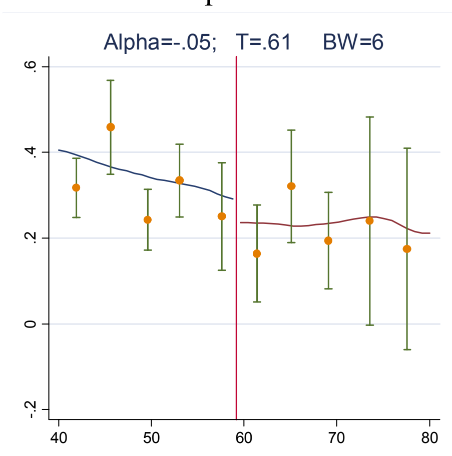
Panel A: 1972 spending per capita, National Archives data



Panel B: Head Start Participation, NELS Base-Year Respondents



Panel C: Participation in Other Preschool Programs, NELS Base-Year Respondents



Top panel from julie2\u_figure2_funding.do (uses hsspend_per_cap72_c)
Bottom two panels from julie2\u_figure2_nels.do

Figure 3: Discontinuity in Educational Attainment by Age, 1990 Census
Panel A: High School or more

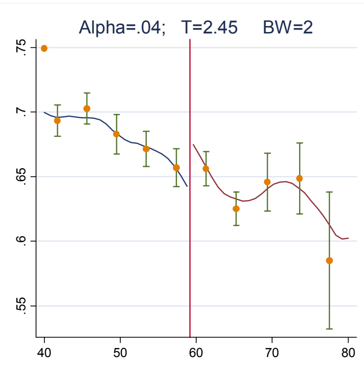
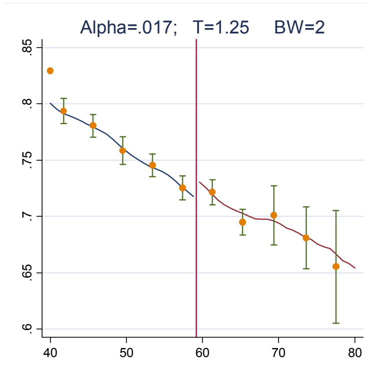
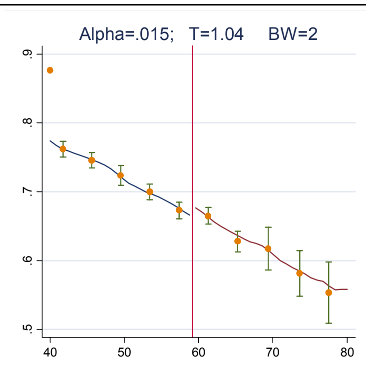
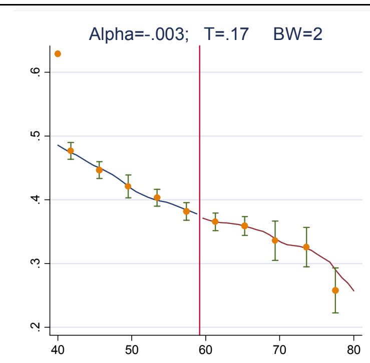
Age	High school or more	Age	
18-24		25-34	
35-54		55+	

Figure 3: Discontinuity in Educational Attainment by Age, 1990 Census
Panel B: Some college

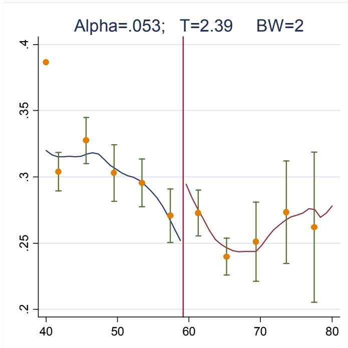
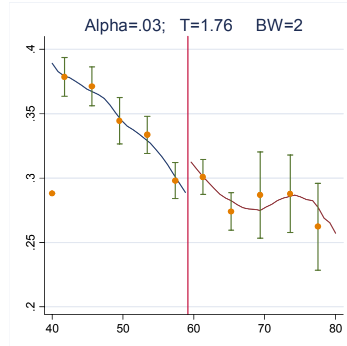
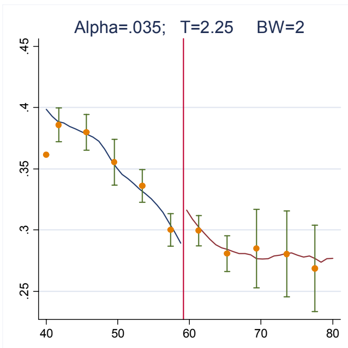
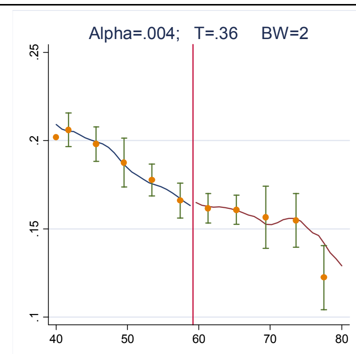
<table>
<tr><td>Age</td><td>Some college</td><td>Age</td><td></td></tr>
<tr><td>18-24</td><td>
</td><td>25-34</td><td>
</td></tr>
<tr><td>35-54</td><td>
</td><td>55+</td><td>
</td></tr>
</table>

Figure 3: Discontinuity in Educational Attainment by Age, 1990 Census
Panel C: Completed college

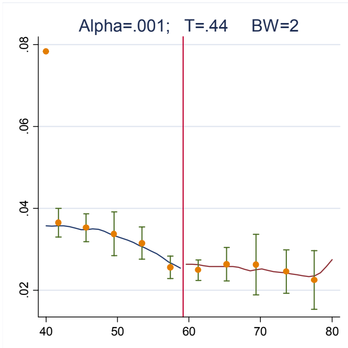
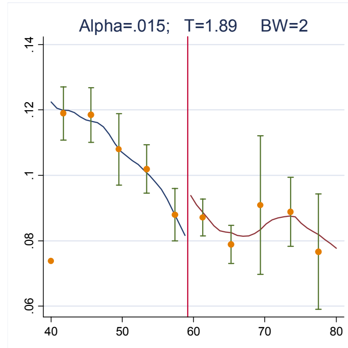
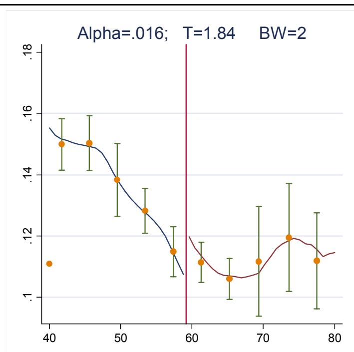
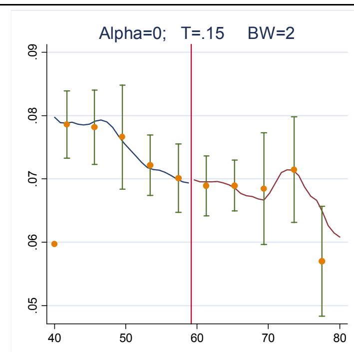
Age	Completed college or more	Age	
18-24		25-34	
35-54		55+	

Figure 4: Discontinuity-in-Differences Estimate for Educational Attainment, Directly Treated (18-24) versus Other Cohorts, 1990 Census

Panel A: High school or more

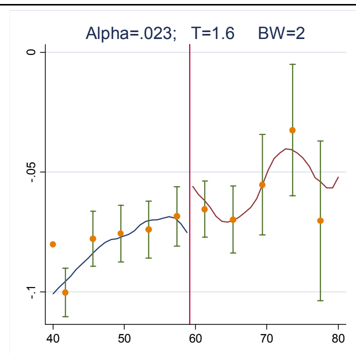
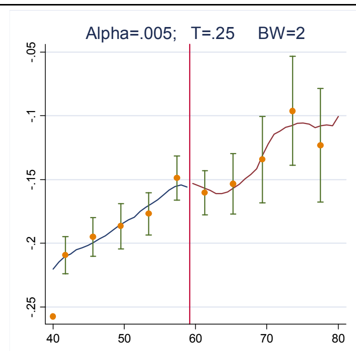
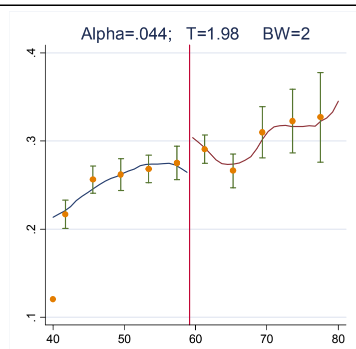
Age contrast	High school or more
(18-24) vs (25-34)	
(18-24) vs (35-54)	
(18-24) vs (55+)	

Figure 4: Discontinuity-in-Differences Estimate for Educational Attainment, Directly Treated (18-24) versus Other Cohorts, 1990 Census

Panel B: Some college

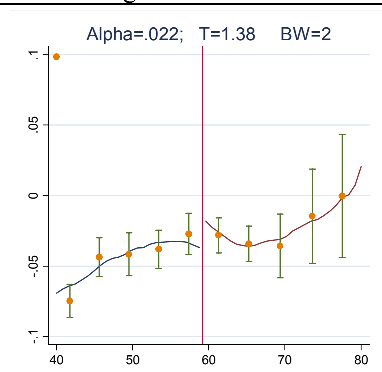
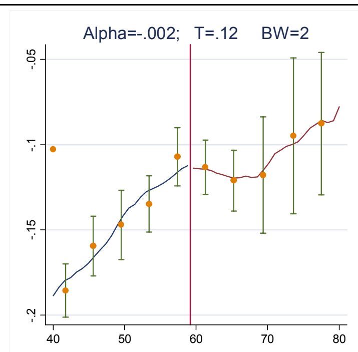
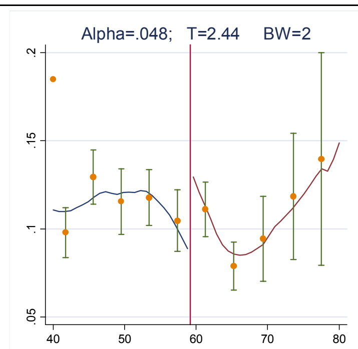
Age contrast	Some college
(18-24) vs (25-34)	
(18-24) vs (35-54)	
(18-24) vs (55+)	

Figure 4: Discontinuity-in-Differences Estimate for Educational Attainment, Directly Treated (18-24) versus Other Cohorts, 1990 Census

Panel C: college or more

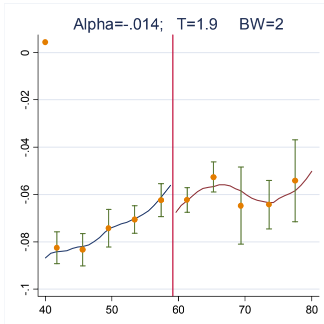
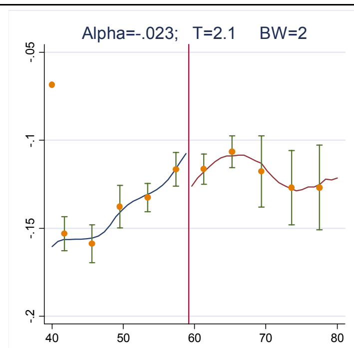
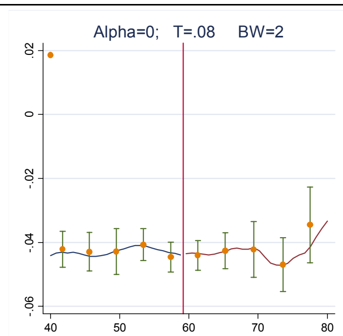
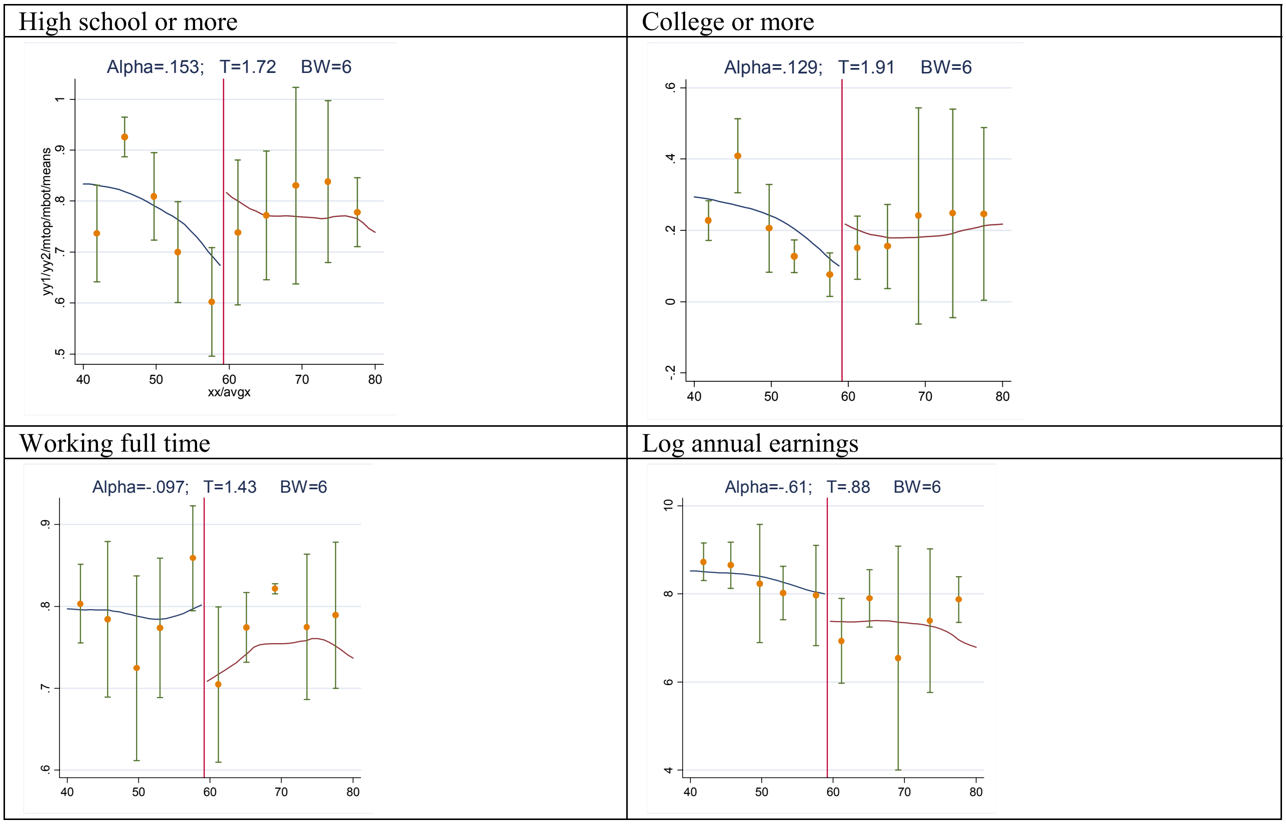
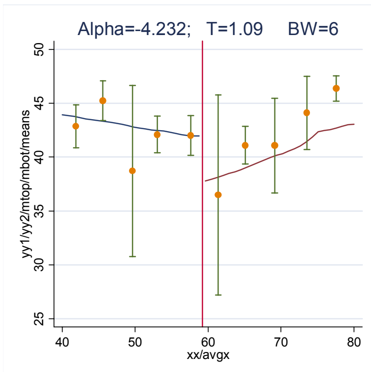
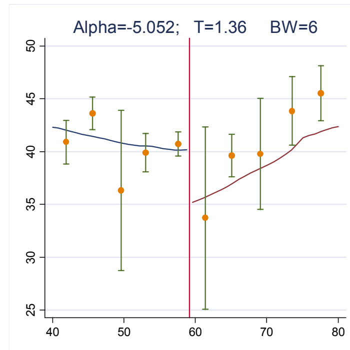
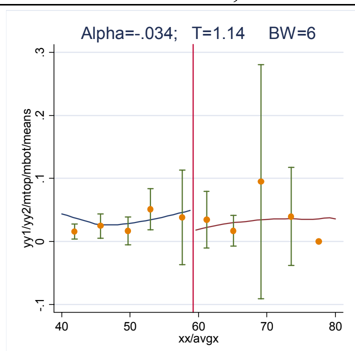
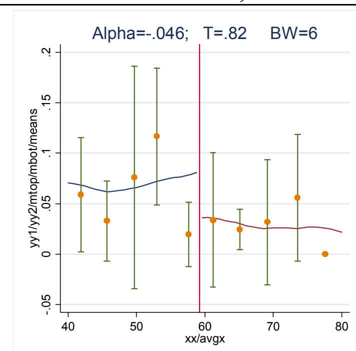
Age contrast	College or more
(18-24) vs (25-34)	
(18-24) vs (35-54)	
(18-24) vs (55+)	

Figure 5
Discontinuity in Educational Attainment and Labor Market Outcomes,
NELS 2000 Follow-Up Survey



Julie2\nels_porter2000.do. Above results calculated using f4pnlfl=1 / f4pnlwt sample and weight combination for 2000 survey; using the f4bypnwt=1 / f4bypnwt sample and weight produces stronger evidence here for effects on educational attainment, but yields more imbalance in background covariates, particularly mom’s educational attainment and 1987 family income.

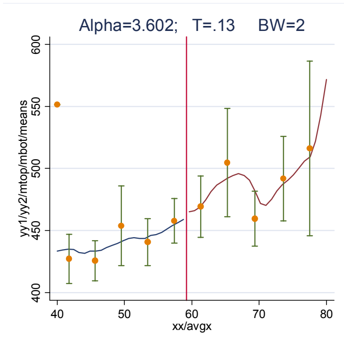
Figure 6
Discontinuity in Achievement Test Scores and Arrests,
NELS Base-Year and First Follow-Up Surveys

8th grade reading test scores	8th grade math test scores
	
Number of arrests, first follow-up survey	Number of arrests, second follow-up survey
	

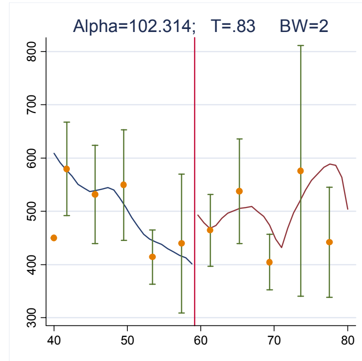
Julie2\nels_porter2000.do

Figure 7
Discontinuity in Other Federal Spending, 1972 National Archives Data

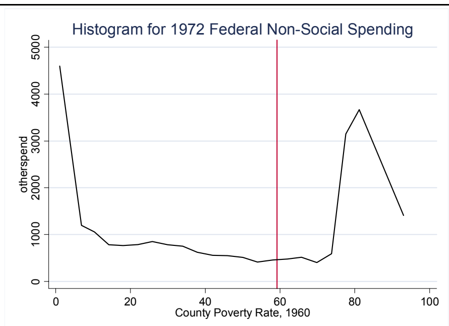
Other Federal Social Spending



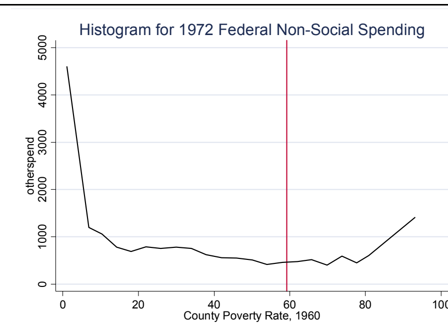
Other Federal Non-Social Spending (drop outliers with spending > $10,000 per capita)



Histogram for Non-Social Spending, with Outliers



Histogram for Non-Social Spending, without Outliers



julie2\u_other_spending1.do, u_graph_otherfed.do, u_graph_otherfed2.do
Note that Porter non-parametric estimates for discontinuity in other (non social) spending drops 4 counties w/ avg other fed spending per capita >$10,000 (2 of those obs had >$30K/capita, 1 had >$20K/capita, and another had >$10K / capita).

Table 1: Test of Balance on Covariates for Census and NELS Data at OEO Cutoff

	BW=6 NELS, BW=2 Census Alpha (t-statistic)
NELS Base Year Sample	
Black	-.19 (1.59)
Hispanic	-.00 (0.04)
Log family income 1987	.31 (0.69)
Mother's education (years)	-.05 (0.05)
Urban	-.02 (0.29)
NELS 2000 Follow-Up	
Black	-.12 (1.03)
Hispanic	-.02 (0.17)
Log family income 1987	-.05 (0.13)
Mother's education (years)	.29 (0.22)
Urban	-.03 (0.81)
1990 Census	
Black	-.01 (0.13)
Urban	.02 (0.73)
Same county, 1985	-.02 (1.82)*
1980 Census	
Black	.00 (0)
Urban	.00 (0.20)
Same county, 1975	-.02 (1.51)
1970 Census	
Black	.02 (0.49)
Urban	.04 (0.71)
Pop change (fraction)	.06 (2.54)**
1960 Census	
Black	.02 (0.48)
Urban	.04 (0.71)
1950 Census	
Non-white	.03 (0.71)
Urban	.01 (0.32)

NELS results from julie2\nels_porter2000.do. Note that the 2000 NELS results are calculated using the f4pnlfl==1 sample (those who responded in every wave). If we use this sample and the attendant weight (f4pnlfl) we get results that are reasonably balanced with respect to these background characteristics, but if we instead use the f4bypnfl=1 / f4bypnwt sample and weight combination we get terrible balance w/ respect to Hispanic, 1987 income, mom education (although urban is a touch better than above). County results from julie2\county_balance1.do

Table 2
NELS and 1990 Census Results for Pseudo-Cutoff (1960 Poverty Rate = 40%)

Dataset / variable	Alpha (t-statistic)
NELS	
Head Start participation	-.00 (0.13)
Reading scores, 1988	-0.83 (0.48)
Math scores, 1988	-0.76 (0.43)
Arrests, 1990	-.02 (0.56)
Arrests, 1992	-.02 (0.53)
High school completion, 2000	-.08 (1.66)*
College completion, 2000	.02 (0.38)
Work full time, 2000	-.03 (0.72)
Log earnings, 1999	.55 (1.50)
1990 Census	
High school completion, 18-24	.018 (1.11)
High school completion, 25-34	.026 (1.79)*
High school completion, 35-54	.015 (1.07)
High school completion, 55+	.027 (1.73)*
Some college, 18-24	-.002 (0.11)
Some college, 25-34	.028 (1.62)
Some college, 35-54	.010 (0.62)
Some college, 55+	.010 (0.91)
College completion, 18-24	.010 (2.56)**
College completion, 25-34	.012 (1.24)
College completion, 35-54	.002 (0.20)
College completion, 55+	-.001 (0.10)

Note: Bandwidth = 6 for NELS, bandwidth = 2 for 1990 Census
* = Statistically significant at 10 percent; ** = Statistically significant at 5 percent

Nels_arbitrary40.do

Appendix A
Mobility Patterns in the PSID

To explore mobility patterns over time we use a special restricted-use geo-coded version of the Panel Study of Income Dynamics (PSID) that provides address information for PSID respondents for the years 1968 to 1992. Our analytic sample in what follows consists of respondents born from 1965 to 1974, for whom we have address data at both age 3 and age 18. All estimates presented below are calculated using the 1995 individual-survey sampling weights.

Appendix Table A1 shows that of all PSID respondents born 1965-74, fully 71% are living in the same county at age 13 (the age at which we first capture addresses in the NELS) as at age 3. Two-thirds of this sample are living in the same county at age 18 as at age 3. These figures are slightly higher for those PSID respondents who were living at age 3 in counties with 1960 poverty rates that put them 301-600 in the national ranking, and slightly higher still for those in counties with 1960 poverty rankings in the top 300. This suggests that there might be a very modest positive relationship between the 1960 poverty rate of one's county of residence at age 3 and the probability to stay within that county through age 13 or age 18, although the samples of PSID respondents in these poor counties is quite small.

Appendix Table A1

	All PSID respondents Born 1965-1974	At age 3, in county w/ 1960 pov rank 1-300	At age 3, in county w/ 1960 pov rank 301-600
(N)	(2,079)	(98)	(123)
Same county age 3 and 13	71%	84%	75%
Same county age 3 and 18	66%	76%	66%

For our "discontinuity-in-differences" estimator, bias would come only from within-county across-cohort differences in mobility across county lines. Appendix Table A2 shows that when we focus on the national sample of PSID respondents there is very little difference in across-county mobility between those born 1965-69 and those born 1970-74. Another way to see this is to regress an indicator for whether the respondent lives in the same county at age 3 and age 13 against either a continuous year-of-birth variable or a series of dummy variables for year of birth for those born 1965-74. In neither case are these year-of-birth variables statistically significant predictors of staying within the county. Similar results hold for the probability of being in the same county at age 3 and age 18.

Appendix Table A2

	All PSID respondents born 1965-69	All PSID respondents born 1970-74
(N)	(1,000)	(1,079)
Same county age 3 and 13	71%	71%
Same county age 3 and 18	66%	65%

Because our Census and NELS data organize people by either their age when the outcome of interest is measured (Census) or at age 13 (NELS), perhaps more relevant is the fraction of people in the poorest counties at age 13 or age 18 who were "treated" in those counties (that is, were of Head Start age, 3, in those counties). Appendix Table A3 shows that just under two-thirds of people living in the 300 counties with the highest 1960 poverty rates, or counties with 1960 poverty rankings of 301-600, were living in the same county at age 3.

Appendix Table A3

Panel A: By age 13 county	Pov rank 1-300	Pov rank 301-600
(N)	(100)	(128)
Same county age 3 and 13	63%	69%
Same county age 3 and 18	56%	60%
Panel B: By age 18 county		
(N)	(98)	(N)
Same county age 3 and 13	62%	65%
Same county age 3 and 18	60%	58%

Finally, we can ask whether adult outcomes are systematically different for those who are and are not living in the same county at age 3 and 13 or age 3 and 18. The top panel of Appendix Table A4 shows that those who do and do not stay within the same county are very similar with respect to employment outcomes, educational attainment or criminal activity as measured by the 1995 PSID survey. Panels B and C suggest that there might be some differences between stayers and leavers when we focus on those living in the poorest counties, but the sample sizes are so small when we focus on these sub-groups of respondents that it is not clear how much we can learn from those comparisons.

Appendix Table A4

Panel A: National PSID sample born 1965-74				
	Age 3 vs 13 county		Age 3 vs 18 county	
1995 outcome	Same	Diff't	Same	Diff't
(N)	(1,515)	(564)	(1,405)	(674)
Employed	67%	68%	66%	71%
High school	95%	96%	96%	95%
High school or GED	99%	100%	99%	100%
Crime	10%	12%	10%	11%
Panel B: PSID respondents at age 3 in counties w/ 1960 poverty rank 1-300				
(N)	(80)	(18)	(76)	(22)
Employed	50%	44%	50%	44%
High school	98%	84%	98%	85%
High school or GED	98%	97%	98%	98%
Crime	20%	3%	13%	30%
Panel C: PSID Respondents at age 3 in counties w/ 1960 poverty rank 301-600				
(N)	(94)	(29)	(87)	(36)
Employed	83%	51%	82%	58%
High school	95%	100%	97%	92%
High school or GED	100%	100%	100%	100%
Crime	13%	8%	9%	17%