

**Does Head Start Improve Children's Life Chances?
Evidence from a Regression Discontinuity Design**

Jens Ludwig
Georgetown University and NBER
E-mail: ludwigj@georgetown.edu

Douglas L. Miller
University of California, Davis
E-mail: dlmiller@ucdavis.edu

November 2005

This paper substantially extends an earlier version written with Nate Balis presented at the Fall 2001 APPAM meetings. This research was supported in part by the Georgetown University Graduate School of Arts and Sciences, UC-Davis, and a grant from the Foundation for Child Development to the Georgetown Center for Research on Children in the U.S., and was conducted in part while Ludwig was the Andrew W. Mellon Fellow in Economic Studies at the Brookings Institution and a visiting scholar at the Rockefeller Foundation's Bellagio Conference and Study Center. Miller gratefully acknowledges funding from the National Institute on Aging, through Grant Number T32-AG00186 to the NBER. Thanks to Zehra Aftab, Bradley Hardy, Zac Hudson, Sinead Keegan, Robert Malme, Meghan McNally, Julie Morse, Joe Peters, 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 Eliana Garces and Michael Maltz for sharing data and programs, and to Doug Almond, Colin Cameron, Mark Cohen, Philip Cook, William Dickens, Sue Dynarski, Greg Duncan, Ted Gayer, William Gormley, Jon Gruber, Hilary Hoynes, Brian Jacob, Leigh Linden, Deborah Phillips, Sarah Reber, Peter Reuter, Romi Webster and seminar participants at BLS, UC-Berkeley, UC-Davis, UC-Santa Cruz, Columbia, Cornell, NBER, Rutgers, Stanford and the Institute for Research on Poverty at the University of Wisconsin for helpful comments. Any errors and all opinions are of course ours alone.

IRP Publications (discussion papers, special reports, and the newsletter *Focus*) are available on the Internet. The IRP Web site can be accessed at the following address: <http://www.irp.wisc.edu>

Abstract

This paper exploits a new source of variation in Head Start funding to identify the program's effects on health and schooling. 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. The result was a large and lasting discontinuity in Head Start funding rates at the OEO cutoff for grant-writing assistance, but no discontinuity in other forms of federal social spending. We find evidence of a large negative discontinuity at the OEO cutoff in mortality rates for children ages 5–9 from causes that could be affected by Head Start, but not for other mortality causes or birth cohorts that should not be affected by the program. We also find suggestive evidence for a positive effect of Head Start on educational attainment in both the 1990 Census, concentrated among those cohorts born late enough to have been exposed to the program, and among respondents in the National Education Longitudinal Study of 1988.

Does Head Start Improve Children’s Life Chances? Evidence from a Regression Discontinuity Design

I. INTRODUCTION

This paper provides new evidence on the effects of the Head Start program on schooling and health by drawing on a new source of variation in program funding. We exploit a discontinuity in program funding across counties that resulted from how the Office of Economic Opportunity (OEO) launched the program during the spring of 1965 as part of the War on Poverty. Head Start was established to provide preschool, health and other social services to poor children ages 3 or 4. It represents one of the federal government’s primary means for addressing the large disparities in cognitive and non-cognitive skills observed along race and class lines well before children start school.¹ Head Start currently serves over 800,000 children each year at a cost of around \$7 billion (Haskins, 2004).

Early childhood may represent a particularly promising target for human capital interventions given evidence that neurodevelopmental plasticity may decline with age, together with the possibility of dynamic complementarities in learning (Duncan *et al.*, 1998, Shonkoff and Phillips, 2000, Carniero and Heckman, 2003). “Two generation” programs like Head Start, which target parents as well as children, could also change the stream of human capital investments that parents make in their children over the life course. Head Start’s medical services may also improve children’s health, an important outcome in its own right that may also be associated with lifetime economic outcomes (Case, Fertig and Paxson, 2003, Almond, 2005).

Whether Head Start yields lasting benefits to program participants in practice is an open question, the answer to which will necessarily rest on non-experimental evidence for the foreseeable future. The

¹For example the measured gap in achievement test scores between black and white children ages 3–4 years old is usually around 1 standard deviation, not much different from the gap observed among children in 9th or 12th grade (Jencks and Phillips, 1998, Phillips *et al.*, 1998). Disparities in non-cognitive outcomes such as physical aggression have been documented along family income lines for children as young as 17 months (Tremblay *et al.* 2004).

federal government has recently sponsored a randomized evaluation of Head Start's short-term impacts, which yields evidence of program impacts on achievement tests 1 year after participation on the order of .1 to .2 standard deviations (HHS, 2005).² However controversy about whether any Head Start benefits "fade out" date back to the program's beginning (for example Wolff and Stein, 1966), and has led to recent proposals for changing the program's focus from "comprehensive services to intellectual development" (Haskins, 2004). Even if the government's new randomized evaluation of Head Start included funding to study long-term impacts (which unfortunately it does not), experimental evidence on long-term effects would not be available for 15 or 20 years.

Policy makers might be tempted to draw inferences about Head Start's long-term effects from the encouraging experimental impact estimates found for other model early childhood programs such as Perry Preschool and Carolina Abecedarian, which seem to improve achievement test scores and educational attainment and reduce involvement with anti-social behavior well into adulthood.³ However these programs are much smaller and more intensive than Head Start, and as a result may not be very informative about what we might expect from the less-costly, larger-scale Head Start.⁴

To date the best available evidence for 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). CT find positive program impacts on test scores for children around 8–10 years old, but these improvements fade out for black children. Theirs

²Positive impacts were estimated for letter-word identification ($\sim .2sd$), pre-writing and vocabulary scores ($\sim .1sd$), and the frequency with which parents read to children ($\sim .15sd$), with no detectable effects on oral comprehension or math skills. Effects on problem behaviors were observed for children 3 but not 4 years old (HHS, 2005).

³Because there are so many excellent reviews of this literature already we do not discuss these findings in any detail in our paper. Interested readers are referred to Barnett (1995), Donohue and Siegelman (1998), Karoly *et al.* (1998), Campbell *et al.* (2002), Currie (2001), Ramey and Campbell (1994), Carniero and Heckman (2003), Krueger (2003a), Schweinhart *et al.* (2005), and Magnuson and Waldfogel (2005).

⁴Chicago Child-Parent Centers are less intensive but have not been experimentally studied (Reynolds *et al.*, 2001).

is also one of the few studies to consider Head Start effects on health, and they find that Head Start improves immunization rates for all children but has no statistically significant effects on height-for-age. The findings in GTC suggest long-term effects on educational attainment for whites, and reductions in crime among African-Americans.

While both CT and GTC substantially improve upon previous studies by controlling for unmeasured family fixed-effects associated with Head Start participation, as one recent review notes with respect to the program's long-term impacts: "The jury is still out on Head Start" (Currie, 2001, p. 213). There necessarily remains some uncertainty about what drives variation across siblings in participation, and whether enrollment is related to unmeasured child-specific or time-varying family attributes that also affect outcomes. For policy purposes it is important to develop a better understanding of Head Start's impacts on schooling, given the lack of evidence to date of lasting effects on African-Americans and the fact that more than a third of all black children are poor and nearly one-third of all poor children are black (Census Bureau, 1999, p. 483). Also important is to learn more about Head Start's effects on non-academic outcomes such as health, given debates about whether to increase the program's focus on educational services.

Our paper exploits a discontinuity in program funding across counties that resulted from OEO's implementation of the program during the spring of 1965. Unlike 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 staff to the 300 poorest counties in the country to identify potential service providers and help them develop proposals (Jones, 1979, pp. 6–7).

The first key result in this paper is to demonstrate that Head Start participation and funding rates are 50 percent to 100 percent higher in counties with poverty rates just above OEO's cutoff (the "treatment" group) compared to those just below (controls). These findings come from (somewhat noisy) county-level federal expenditure data obtained from the National Archives, as well as from the National Education Longitudinal Study (NELS), which provides information about a nationally representative

sample of 1988 eighth graders. The funding difference at the heart of our design appears to have persisted through the late 1970s.

We use this discontinuity in funding across counties to identify Head Start impacts by comparing outcomes for people in treatment and control counties “near” the OEO cutoff. Identification comes from the assumption that potential outcomes are smooth around the cutoff. Because the cutoff was based on a predetermined variable, common concerns about strategic behavior with the regression discontinuity (RD) research design are not likely to be an issue here.

Our main finding is that there is a large discontinuity in mortality rates to children 5–9 years of age during the period 1973–83 due to causes-of-death that were directly or indirectly screened for and addressed as part of Head Start’s health services. This mortality effect is quite large, equal to 33–75 percent of the control mean, enough to drive mortality rates from these causes in the treatment counties down to about the national average. To rule out other explanations we demonstrate there is no discontinuity for other causes-of-death or age groups that should not be affected by Head Start.

We also examine educational attainment drawing on two separate data sources, neither of which is ideal (as is often the case in working with historical data⁵) but which taken together seem to provide suggestive evidence for a positive effect of Head Start on schooling. The challenge is to measure adult schooling outcomes for those who were young children in counties “near” the OEO cutoff in the 1960s and 1970s. We first show that there is a discontinuity at the OEO cutoff in educational attainment in the 1990 Census, which is limited to cohorts born late enough to have been affected by Head Start. While the Census provides us with large samples near the cutoff, the disadvantage is that we only identify county of residence for people as adults.

⁵For example, Goldin and Katz (2000) measure schooling returns with data from a state census from 1915 available only for Iowa, and Collins and Margo (2004) examine the effects of the race riots of the 1960s on urban housing markets but cannot directly measure economic damage. Both studies are nonetheless interesting and informative.

To address the possibility of bias from selective migration we show there is also a discontinuity at the OEO cutoff in schooling in the NELS, which improves upon the Census by recording each respondent's county of residence at age 13 rather than adulthood, but at the cost of small samples local to the cutoff. Importantly, we do not find discontinuities in key parent characteristics that are predictive of children's outcomes, such as maternal education. This last finding does not seem consistent with a story in which our results for children's schooling outcomes in the NELS are driven by selective migration of families across county lines.

Do these estimates actually identify the effects of Head Start? One concern is that other forms of government spending might also vary discontinuously around the cutoff. Yet we show the estimated discontinuity at the OEO cutoff in other federal social spending is small and not significant. Another potential concern is the possibility that Head Start and other OEO activities might have affected community mobilization, which could have affected political behavior and in turn local or state spending patterns. Yet we show there is no evidence of a discontinuity in voter registration rates across counties in any year from 1968 through 1980.

Our estimates—like every RD study—rely on functional form assumptions for identification. We show that our results are generally robust to using different specifications to control for unobserved variables that vary with 1960 county poverty rates, including various flexible polynomials as well as non-parametric functions using the approach of Porter (2003). We also present a series of specification tests that demonstrate that there is no discontinuity in most “background” variables such as race or urbanicity. We show there are no discontinuities in child mortality or schooling at other “pseudo-cutoffs” where there are no discontinuities in Head Start funding, nor are there discontinuities in achievement test scores for children in the ECLS-K who were of Head Start age in the late 1990s after the discontinuity in funding had dissipated.

Ideally we would like to extend our analysis to consider Head Start impacts on other longer-term outcomes as well. However our data are not very informative about effects on labor market outcomes because those “treated” by Head Start in our data are still relatively young when we observe their long-

term outcomes (mid-20s in both the Census and NELS) and many are still enrolled in school. For reasons discussed in Section 3 below and Appendix A, the available data are also not very informative about other potential outcomes of interest such as achievement test scores or criminal activity. Understanding program impacts on these other outcomes remains an important topic for future research.

The remainder of the paper is organized as follows. The next section provides more background on Head Start. The third and fourth sections discuss our data and empirical strategy. Our main results for child mortality and educational attainment are presented in the fifth and sixth sections, a variety of specification checks and robustness tests are presented in section seven, and the final section discusses limitations and implications for public policy.

II. BACKGROUND ON HEAD START

Our study provides new evidence on the long-term effects of Head Start by exploiting a natural experiment generated by how the program was launched by OEO in 1964. Head Start began as a summer program for children around ages 3 or 4; by 1970 a majority of program participants were enrolled year-round (nine months).⁶ Widely perceived as a schooling program, Head Start also provides meals, social services, parenting services (sometimes jobs),⁷ and a variety of health services such as immunizations (DTP, polio, smallpox and measles) and unusually comprehensive health screenings that included blood samples, urinalysis and other tests for tuberculin, anemias, diabetes, and hearing and vision problems (North, 1979, O'Brien, Connell, and Griffin, 2004). Daily observation of children by teachers could also

⁶The 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).

⁷For example up to 47,000 parents were employed in Head Start centers the first year of the program and up to another 500,000 parents were part-time volunteers (Zigler and Valentine, 1979, p. 69). Head Start also helps parents improve parenting skills (GAO, 1981, p. 13) and provides subsidized child care and other social services.

help detect problems where parents might not realize the severity of the problem or how to access treatment.

Because Head Start involves federal funding of local service providers, the challenge for OEO in the spring of 1965 was to publicize Head Start, encourage local organizations such as schools and public health or welfare agencies 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 Johnson – all within the span of a few months. Despite OEO’s efforts to publicize Head Start, federal officials were concerned that applications from poor counties would be under-represented in a nationwide grant competition (see for example 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. Volunteers from the federal Presidential Management Intern (PMI) program were provided with funding to travel to the selected counties for two to six weeks 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. Importantly, this particular feature of Head Start’s launch is widely documented in historical accounts of the program’s beginning, suggesting that the discontinuity in grant-writing assistance at the heart of our research design is not the figment of a single historian’s imagination.⁸ Below we demonstrate that the result of this feature of Head Start’s launch is a discontinuity in county-level Head Start program funding and participation rates, which persists through the late 1970s.

We use the targeting of Head Start grant-writing assistance to just the 300 poorest counties to identify the effects of the program on people living in counties “near” the OEO cutoff using a standard

⁸See for example Jones (1979), pp. 6–7, Gillette (1996), p. 222, President Johnson’s speech on Head Start of May 18, 1965 (Zigler and Valentine, 1979, pp. 69–70) and GAO (1981) p. 17.

RD design. Often with RD designs we are concerned with the possibility of strategic behavior that may push observational units above or below whatever cutoff is used for the assignment of treatment. However there does not seem to be much room for strategic behavior with our design given that Head Start grant-writing assistance was assigned on the basis of a predetermined variable (poverty rates measured 5 years earlier), and the problem facing OEO administrators was one of excess funding supply rather than demand so there would be no incentive for “gaming” or favoritism by OEO. However generalizability may be more of an issue with our estimates: The 300th poorest county in the U.S. at the time Head Start was launched had a 1960 poverty rate of around 59 percent, so counties near this cutoff are very poor counties indeed, and most are located in the South.⁹ We return to this point in the conclusion.

III. DATA

In what follows we provide a brief overview of the county- and individual-level data sources used in our analysis. More details on each dataset are in Appendix A.

Perhaps the most important data question for our study is to understand what data source OEO used to identify the 300 poorest counties in the U.S. in 1960. The answer is not immediately obvious because the official federal poverty rate was only invented in 1964, and so the 1960 Census does not include a measure of persons living in poverty. OEO apparently identified the poorest counties using a special 1964 re-analysis of the 1960 Census conducted by the Census Bureau for OEO using the then-

⁹One-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.

newly-defined federal poverty rate, a copy of which we have obtained from the National Archives and Records Administration (NARA).¹⁰

Other county-level data for our study include an OEO file for all federal expenditures on every program for the years 1968 and 1972, also obtained from NARA. The expenditure data are noisy because of some obvious errors (which we have tried to correct) as well as a few cases where an organization ran Head Start programs in multiple counties but only received federal funding in the county with the group's headquarters. Data on child mortality comes from the Vital Statistics Compressed Mortality Files (CMF) for 1973–83, which include information on both cause-of-death and the decedent's age. To measure long-term schooling outcomes we use data from a special tabulation conducted by the Census Bureau for us from the 1990 Census, which include schooling and other outcomes disaggregated by county, age group, gender and race. Finally, we draw on biannual county-level voter registration records from 1968 to 1988.

Our main source of individual-level data is a restricted-use geo-coded (with county identifiers) version of the NELS, which provides information for a nationally representative sample of 8th graders in 1988 with follow-up interviews in 1990, 1992, 1994 and 2000. One key variable for our analysis is the NELS respondent's county of residence in the base year (1988) wave. A drawback with the NELS is the study is intended to provide a nationally representative sample, and so the baseline (1988) survey includes only 649 respondents who lived in counties with 1960 poverty rates among the 300 poorest, and 674 who lived in one of the next 300 poorest counties (see Appendix A). Other key variables include whether parent reports about whether the respondent has participated in Head Start (on average consistent with aggregate enrollment data), and educational attainment in the 2000 wave of the study (around 25 years of age).

¹⁰The alternative possibility is that OEO targeted grant-writing assistance using data on the proportion of families in the county with incomes below \$3,000, although our analysis reveals that a larger discontinuity in funding at the 300th poorest county using the official poverty rate.

Note that we do not use several other obvious data sources that at first glance might seem promising. We focus on the 1990 rather than 2000 Census because the Head Start funding discontinuity behind our design attenuates over time, and because across-county migration will be even more of a problem in 2000 than 1990. The High School and Beyond survey (unlike the NELS) is not geo-coded at the county level. The restricted version of the Panel Study of Income Dynamics provides annual geo-coded addresses but since the PSID has a national sampling frame the data winds up providing a very small and unrepresentative draw of people in counties near the OEO cutoff. The result is that we cannot detect a discontinuity in Head Start participation among the small number of PSID respondents who live in counties near the OEO cutoff, although we do use broader samples from the PSID below for descriptive mobility calculations.¹¹

IV. EMPIRICAL STRATEGY

The heart of our research design is to examine whether discontinuities in Head Start funding at the OEO cutoff are mirrored by discontinuities in other outcomes. Our analyses are conducted using county-level data, the level at which our Census and mortality data are reported. We aggregate the NELS to the county level as well, the simplest way to account within our semi-parametric framework for the fact that students within counties are not independent units.¹²

Let Y_c represent some average outcome for residents of county (c), such as child mortality rate, 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 occurs

¹¹If we restrict our attention to PSID respondents for whom we have addresses at age 3 plus adult educational outcomes, we are left with just around 97 people, who reside in only 19 of the 300 poorest counties and 123 people who reside in only 27 of the next 300 counties.

¹²That 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.

at $c=300$). The provision of grant-writing assistance is a deterministic function of the county's 1960 poverty rate, as in equation (1), where $P_{300}=59.1984$.

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

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 help in securing Head Start funding.

If the offer of grant-writing assistance had no effect beyond increasing the amount of funding, then in principle we could calculate the effect on long-term outcomes per dollar of extra Head Start funding by dividing the ITT effect for schooling or mortality by the ITT effect on program funds. In practice scaling the ITT estimates in this way is complicated by the fact that our estimates for the discontinuity in Head Start funding across counties are somewhat imprecise, but perhaps more importantly are only available for a few of the birth cohorts exposed to the Head Start funding discontinuity. The NELS captures discontinuities in Head Start participation but will miss across-county differences in funding that go towards increased spending per participant rather than expanded enrollment rates. In short we can generate fairly precise ITT estimates on outcomes but have less power to reliably scale these estimates to recover the effects per dollar of Head Start spending. We focus our analysis primarily on the “reduced form” ITT-style estimates for the overall discontinuity in outcomes at the OEO cutoff.¹³

¹³Estimating the effects of Head Start enrollment itself also requires a “stable unit treatment value assumption” (SUTVA) that children's outcomes are unaffected by participation rates among other county residents. This assumption could be violated if for example Head Start affects the probability that a given classroom contains a “rotten apple” that disrupts learning (Lazear, 2001) or if families share new parenting skills within social networks.

Our main estimating equation is given by (2), where $m(P_c)$ is an unknown smooth function of 1960 poverty, and α is the impact of grant writing assistance. The effect that we seek to identify is the one relevant for the poorest counties near the OEO cutoff.¹⁴

$$(2) \quad Y_c = m(P_c) + G_c\alpha + v_c$$

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). This is a plausible assumption, as the cutoff was defined on the basis of a predetermined variable (poverty rates 5 years before Head Start’s launch), and because (as we demonstrate below) this cutoff does not seem to have been used to distribute funding for other federal programs.¹⁵

One practical question is how to model $m(P_c)$. We present “parametric” estimates that estimate (2) using different polynomial functions of P_c calculated using counties “near” the OEO cutoff. We also refine the estimates by further controlling for observable county covariates from the 1960 Census, such as total population and age or race distribution. This parametric approach assumes we can adequately control for the other determinants of long-term outcomes that vary across counties using a sufficiently flexible polynomial function of P . We focus on linear and quadratic models, allowing the slope of these functions to vary on each side of the cutoff.

We also further relax these functional form assumptions and calculate our estimates using the nonparametric regression discontinuity approach of Porter (2003). This method uses locally linear

¹⁴Given that the unit of treatment is the county, our empirical analysis focuses 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.

¹⁵An alternative way to think about identification comes from Lee (2003). Under the smoothness assumption, 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} . So long as the mapping between P and Y is also locally smooth then potential outcomes will be independent of G given P , the necessary condition for identification (Hahn, Todd and van der Klaauw, 2001).

regressions (Fan 1992) to estimate the left and right limits of the discontinuity. The gap between these limits is the estimated treatment impact. This can be estimated in one step:

$$(3) \quad Y_c = b_0 + b_1(P_c - P_{300}) + \alpha G_c + b_2 G_c(P_c - P_{300}) + v_c$$

with kernel weights $w_c = K((P_c - P_{300})/h)$

for a chosen bandwidth h . We essentially estimate a kernel-weighted linear regression using data points to the left of the OEO cutoff, another kernel-weighted linear regression with data to the right of the OEO cutoff, and then calculate our estimate for the treatment impact (the parameter α) as the difference between the left and right limits of these regressions at the OEO cutoff. For $K(z)$ we use the Epanechnikov kernel, $(.75)(1-.2z^2)/\sqrt{5}$ for $|z| < \sqrt{5}$.

We present standard errors below that come from analytically estimating the variance of our parameter estimate using the formula as in equation (4), from Porter (2003). We have also calculated bootstrapped standard errors, which are typically very similar to those from equation (4).

$$(4) \quad \text{Var}[\alpha] = [(\sigma^{2+}(P_{300}) + \sigma^{2-}(P_{300})) \times c_K] / [(Nh) \times f(P_{300})]$$

where $\sigma^{2+}(P_{300})$ = right hand limit of variance of v_c at OEO cutoff

$\sigma^{2-}(P_{300})$ = left hand limit of variance of v_c at OEO cutoff

c_K = constant function based on shape of kernel¹⁶

$f(P_{300})$ = density of P at cutoff

The remaining estimation issue has to do with bandwidth selection. Because the RD design is identified only at the discontinuity, we try to balance the competing goals of staying as local to the OEO cutoff as possible while ensuring we have enough data to yield informative estimates. Unfortunately there is currently no widely agreed-upon method for selection of optimal bandwidths in the semi-parametric RD context, and so our strategy is to present results for a broad range of candidate bandwidths. We use a bandwidth range from 4 to 16 for most of our datasets, with a focal preferred bandwidth of 8. For

¹⁶We have calculated the constant to be 2.0116, for the kernel that we use.

example, a bandwidth of 8 assigns positive weight to the 958 counties in the NARA spending data that have 1960 poverty rates within $8 * \sqrt{5} = 17.9$ percentage points of the (59.1984) cutoff. Our tables show how many counties receive non-zero weight for each bandwidth and dataset. The 1990 Census provides more information near the cutoff so we use a range of 1.5 to 6, with a preferred bandwidth of 3.

Appendix B discusses the results of a more formal bandwidth selection process using leave-one-out cross-validation, which typically selects bandwidths more towards the upper end of the range that we present, although this cross-validation procedure itself has several important limitations as well. In any case, the overall pattern of findings across outcomes and datasets is generally consistent for different bandwidths and semi-parametric versus parametric approaches.

V. RESULTS FOR HEAD START FUNDING AND PARTICIPATION

What effect did OEO's Head Start grant-writing assistance to the 300 poorest counties have on Head Start funding? Historical documents note that 240 (80 percent) of these 300 counties received Head Start funding (GAO, 1981), compared to 43 percent of all counties nationwide (estimated from our 1968 federal expenditure data). Data from the NELS and for county-level federal expenditures are consistent with these historical accounts.

Table 1 shows that for the 349 "control" counties with 1960 poverty rates that are within 10 percentage points of the OEO cutoff from below (49.198 to 59.198), average Head Start spending per 4 year old in 1968 was equal to \$134 (in 1968 dollars). For the 228 "treatment" counties with 1960 poverty rates 10 percentage points above the OEO cutoff this figure is \$288 per 4 year old, more than twice as much. In 1972 Head Start spending per 4 year-old is still nearly 60 percent higher in the treatment than control counties.

In Figure 1 we show that the difference shown in Table 1 in Head Start funding around the OEO cutoff is driven in large part by a sharp drop-off in spending at the cutoff itself. The solid line in the top panel presents a histogram of county-level Head Start spending per capita in 1968, calculated using a bin width of 4; the bottom panel shows Head Start spending in 1972.

Table 1
County Characteristics

Variable	Counties with 1960 Poverty 49.198% to 59.198%		Counties with 1960 Poverty 59.1984% to 69.1984%	
	Mean	Std. Dev.	Mean	Std. Dev.
# observations (counties)	349		228	
Head Start Spending per 4-year-old 1968	134	(277.00)	288	(915.00)
Head Start Spending per 4-year-old 1972	183	(569.00)	289	(927.00)
Other Social Spending per capita 1972	446	(128.00)	483	(167.00)
Fraction of those 18 to 24 with High school or more	0.67	(0.09)	0.644	(0.07)
Fraction of those 25 to 34 with High school or more	0.74	(0.06)	0.709	(0.06)
Fraction of those 35 to 54 with High school or more	0.693	(0.07)	0.647	(0.07)
Fraction of those 55 and older with High school or more	0.396	(0.09)	0.36	(0.08)
Fraction of those 18 to 24 with some college or more	0.288	(0.12)	0.258	(0.09)
Fraction of those 25 to 34 with some college or more	0.321	(0.09)	0.289	(0.08)
Fraction of those 35 to 54 with some college or more	0.324	(0.08)	0.291	(0.07)
Fraction of those 55 and older with some college or more	0.174	(0.06)	0.16	(0.04)
Fraction of those 18 to 24 with a Bachelor's degree or more	0.0295	(0.02)	0.0257	(0.02)
Fraction of those 25 to 34 with a Bachelor's degree or more	0.0968	(0.05)	0.0847	(0.04)
Fraction of those 35 to 54 with a Bachelor's degree or more	0.124	(0.05)	0.109	(0.04)
Fraction of those 55 and older with a Bachelor's degree or more	0.0714	(0.03)	0.0685	(0.02)
1990 County population	24202	(24054.00)	21371	(29799.00)
Fraction ages 18–24	0.0958	(0.03)	0.0954	(0.02)
Fraction ages 25–34	0.148	(0.02)	0.149	(0.02)
Fraction ages 35–54	0.243	(0.02)	0.238	(0.02)
Fraction ages 55 plus	0.243	(0.05)	0.232	(0.05)
1990 Percent Urban	0.0254	(0.12)	0.0172	(0.10)
1990 Percent Black	0.163	(0.16)	0.266	(0.22)
1990 Per capita income	9520	(1537.00)	8488	(1434.00)
NELS results				
# observations (counties)	28		17	
Head start participation ("by" weights)	0.233	(0.16)	0.388	(0.24)
Head start participation ("fl" weights)	0.244	(0.18)	0.423	(0.24)

Standard deviations in parentheses. All means are unweighted. Data from the 1990 census STF4 file, and from the NELS

Figure 1A: Head Start Participation and 1968 Funding

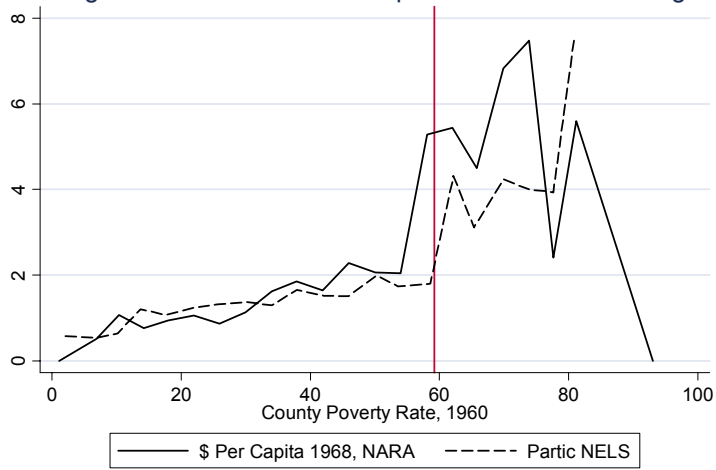


Figure 1B: Head Start Participation and 1972 Funding

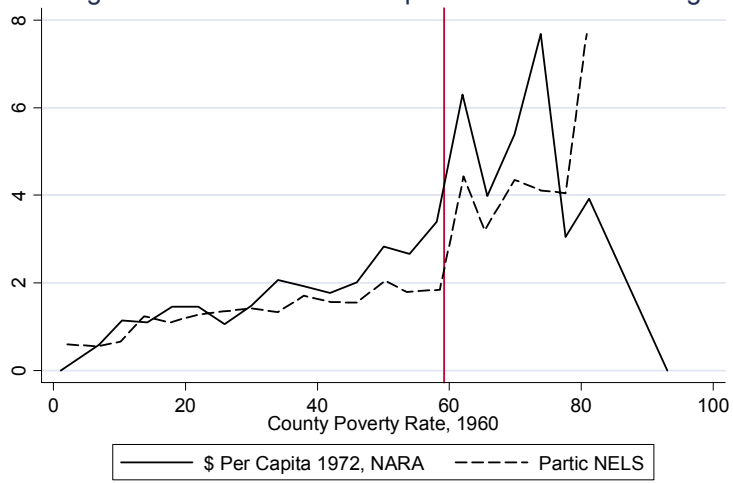


Figure 1 also demonstrates that the discontinuity in Head Start spending in 1968 and 1972 is mirrored by a discontinuity in parent-reported Head Start participation rates among children in the NELS (dashed line in both panels), calculated using the same bin width as with the spending data and re-scaled to fit on the same axis. In fact the overall similarity in patterns across counties is striking given that each data source is somewhat noisy (Appendix A) and capture different points in time.¹⁷ The fact that we observe a discontinuity in Head Start participation rates for NELS respondents, who would have been of Head Start age during the late 1970s, speaks to the stability of the cross-sectional differences in county Head Start resources, and implies that many cohorts were exposed to the “natural experiment” at the heart of our study.¹⁸

The top panel of Table 2 presents results from estimating the “first stage” effects of OEO’s grant-writing assistance on Head Start participation rates in the NELS using the different estimation approaches discussed in the preceding section. Since all of our results tables are formatted similarly, we describe this format for Table 2 in some detail. Each row presents results for a different dependent variable. Each column presents a different specification. The first five columns (“Porter”) show results from non-parametric regression discontinuity specifications, with varying bandwidths. For each bandwidth, we show the number of counties that receive non-zero weight in estimation. We also present the point estimate and standard error.

For example in the first row, second column, we use a bandwidth of 4 to non-parametrically estimate the discontinuity in NELS Head Start participation using the base year (“by”) weighted sample.

¹⁷Of the 1,346 counties that received any Head Start funding in 1968, 72 percent received Head Start funding in 1972. Of the 1,084 counties that received Head Start funding in 1972, 90 percent had received Head Start funding in 1968. The counties funded in 1968 and 1972 may not overlap perfectly because of errors in the data on federal spending, termination of OEO funding for some programs, and the addition of new Head Start programs in some areas.

¹⁸The persistence of the discontinuity in county-level program funding presumably results from the flat trajectory of Head Start appropriations over time after the first few years of the program’s launch (Haskins, 2004) together with a “hold harmless” rule that prevented states from losing Head Start funding from one year to the next (Jones, 1979).

Table 2
Regression Discontinuity Estimates of the Effect of Head Start Assistance on Head Start Spending and Participation
NELS County Averages

Variable	Porter				Parametric			
	LHS Mean	4	6	8	10	16	Flexible Linear	Flexible Quadratic
Bandwidth		4	6	8	10	16	8	16
Poverty Range		43	64	95	125	281	43	82
Number of observations (counties) with nonzero weight								
Head Start participation ("by")	0.206	0.168 (0.135)	0.120 (0.105)	0.137 (0.090)	0.144* (0.077)	0.128** (0.054)	0.168 (0.148)	0.217 (0.142)
Head Start participation ("fl")	0.22	0.238* (0.143)	0.219** (0.109)	0.176* (0.096)	0.169** (0.085)	0.151** (0.058)	0.238 (0.156)	0.316** (0.151)
National Archives Federal Spending Data								
Variable	Porter				Parametric			
	LHS Mean	4	6	8	10	16	Flexible Linear	Flexible Quadratic
Bandwidth		4	6	8	10	16	8	16
Poverty Range		524	734	958	1193	2155	484	863
Number of observations (counties) with nonzero weight								
1968 Head Start Spending per child	134.379	139.041 (125.468)	91.053 (104.618)	115.224 (89.378)	133.84* (79.295)	134.58** (60.697)	127.389 (120.098)	112.706 (112.603)
1972 Head Start Spending, per child	205.64	184.876 (144.428)	73.749 (115.090)	89.572 (99.617)	106.072 (87.362)	130.452** (65.584)	175.773 (142.363)	155.899 (132.292)
1972 other social spending, per capita	456.325	5.826 (26.145)	17.952 (22.397)	20.001 (19.215)	21.020 (17.269)	13.748 (14.487)	-2.361 (25.624)	5.659 (27.471)

Note: Standard errors in parentheses. RD estimates of treatment effects estimate the jump in educational outcomes associated with receiving 1965 PMI assistance. RD methodology based on Porter (2003), locally weighted kernel regression, as discussed in text, with an Epanechnikov kernel. Parametric models give equal weight to observations within the range of the cutoff. The model the last column includes controls for log(1960 population), 1960 %urban and 1960 %black. For the census data, the last column also includes state fixed effects. * = p<.1; ** = p<.05

There are 43 counties that contain NELS respondents and receive non-zero weight used to estimate this discontinuity. The point estimate of the impact is an increase in probability of participation by .168, with a standard error of .135. Is this estimated impact in Head Start participation large? To help answer this we show in the first column of the table (“LHS Mean”) the left-hand-side predictions of the non-parametric regression at the threshold.¹⁹ This represents the no-program counterfactual against which the program impact can be compared. Henceforth we refer to this as the “control mean.” As the non-parametric model is re-estimated using larger bandwidths, we see (reading across the first row) that the point estimate stays in the .120 to .168 range, and the standard error falls as more data are incorporated.

The final two columns, first row of Table 2 suggest that the estimate is not overly sensitive to using a more parametric approach instead, specifically using data within 8 percentage points of the threshold and modeling the direct impact of poverty as linear (allowing slopes to change at the threshold), and in the final column modeling the impact of poverty as a quadratic, again allowing the polynomial parameters to change at the threshold, using a larger poverty range to incorporate more data.²⁰

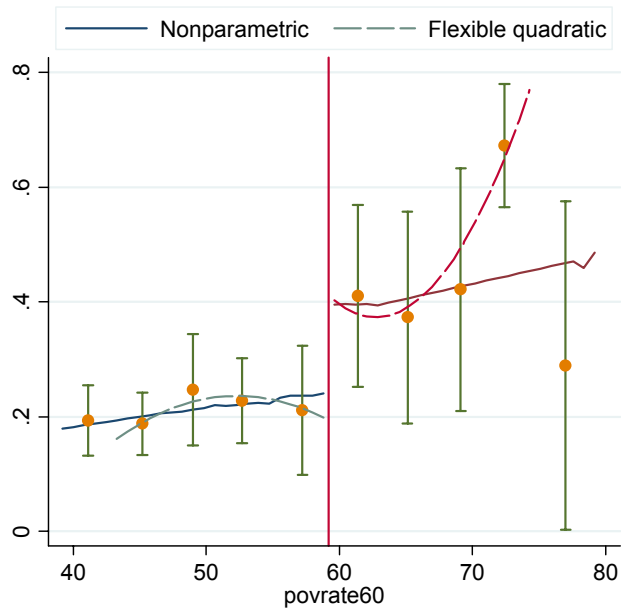
Figure 2 provides more intuition about these estimates by showing the semi-parametric estimates for $m(P)$ and α using a bandwidth of 8 (solid lines in the figure), the parametric estimates from the flexible quadratic specification (dashed lines), and raw cell means and their 95 percent confidence intervals from grouping the data into five categories on each side of the cutoff over the range for which we graph our estimates (40 to 80 percent poor). Notice the cell mean just to the right of the cutoff is almost twice the cell mean just left of the cutoff, consistent with the idea that our estimates are not driven by decisions about the functional form used to model $m(P)$. Note also that with a county poverty rate of about 60 percent at the OEO cutoff and a Head Start participation rate of around 40 percent at the right-

¹⁹Each specification and bandwidth choice has its own left-hand-side mean. For parsimony we only report the mean from the specification with a bandwidth of 8.

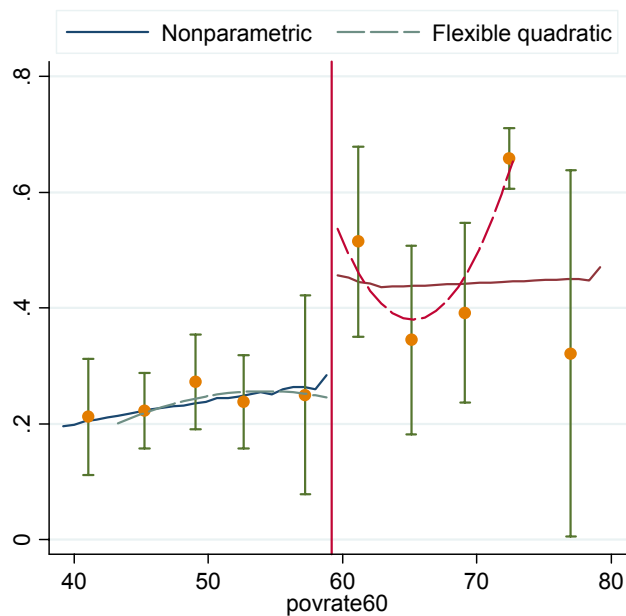
²⁰We have also estimated a specification similar to that of the final column, where we have also included covariates such as 1960 log county population, 1960 percent urban, and 1960 percent black. Our results are typically invariant to the use of these covariates, and so we do not report that specification in the tables.

Figure 2: Estimated Discontinuity in Head Start Participation in the NELS

Panel A: Discontinuity in Head Start participation, NELS base year sample (“by”)



Panel B: Discontinuity in Head Start participation, NELS first follow up sample (“fl”)



Note: Panel A Estimated nonparametric discontinuity = .136 T-stat = 1.51, bandwidth = 8.
 Panel B Estimated nonparametric discontinuity = .176 T-stat = 1.83, bandwidth = 8.

hand limit, the NELS data imply around two-thirds of all poor children participate in Head Start in the treatment counties at the OEO cutoff.

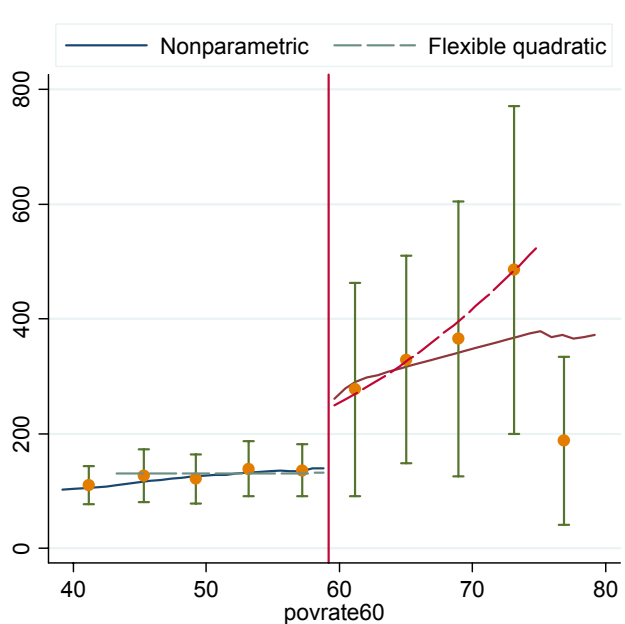
The second row of Table 2 shows that the estimated discontinuity in Head Start participation using the NELS first follow-up sample (“f1”) is even larger than with the base year sample, equal to between 70 percent and 140 percent of the control mean. These estimates may be more relevant for the estimated outcomes shown below that use data from later NELS follow-ups.

The bottom panel of Table 2 and Figure 3 shows that the estimated discontinuities in Head Start spending per 4-year-old are similarly large as a proportion of the control mean (left limit). Because the spending data are noisy and the series is “jumpy” given a number of counties have zero Head Start funding, these spending results are only statistically significant at larger bandwidths that do more smoothing, driven by reductions in the standard errors rather than changes in the point estimates. For example the discontinuity with a bandwidth of 16 is within 4 percent of the estimate for $bw=4$, yet the t -statistic is twice as larger with the former (2.2 versus 1.1).

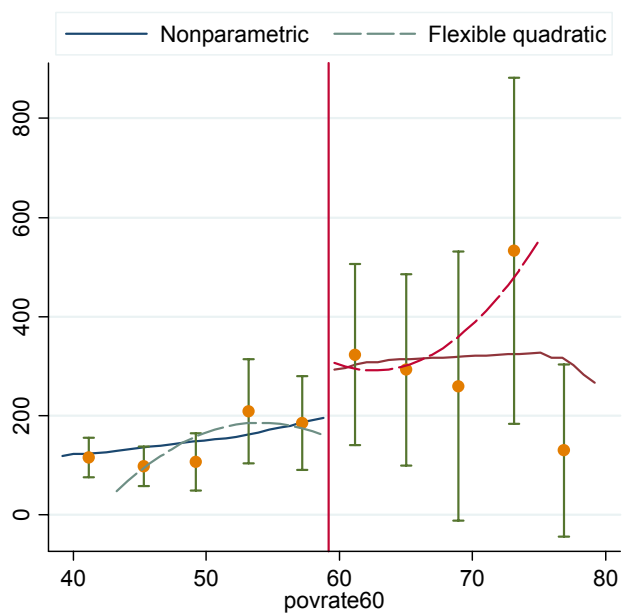
How long did this discontinuity in Head Start funding at the OEO cutoff last? Unfortunately we do not have high quality data that spans an extended time frame, and so must infer what we can from indirect sources. The fact that the NELS participation rates jump at the cutoff indicate that the Head Start funding discontinuity across counties persisted through 1979. While the Community Partnership Act of 1974 required Head Start funding to become more equalized across areas (at least across states), limited growth in program funding during the 1970s together with a “hold harmless” clause that prevented states from receiving funding below 1975 levels meant that funding equalization across areas began in earnest only in the late 1970s (Jones, 1979, GAO, 1981). By the late 1990s the Head Start funding disparity across counties at the OEO cutoff seems to have completely dissipated, as evidenced by the fact that we do not observe a discontinuity in Head Start participation rates among children in the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K).

Figure 3: Estimated Discontinuity in Head Start Funding per 4 year old, National Archives

Panel A: 1968 Head Start funding per 4 year old



Panel B: 1972 Head Start funding per 4 year old



Note: Panel A Estimated nonparametric discontinuity = 115.22 T-stat = 1.28, bandwidth = 8.
 Panel B Estimated nonparametric discontinuity = 89.57 T-stat = 0.89, bandwidth = 8.

VI. HEAD START AND CHILD MORTALITY

To estimate the impact of Head Start on health, we draw on county-level data from the Vital Statistics Compressed Mortality Files (CMF). We identify causes-of-death that may be affected by Head Start through discussions with pediatricians about what health problems the program's health services should influence, which include tuberculosis, whooping cough, infections, polio, measles, diabetes, malnutrition, meningitis, anemias, and respiratory problems such as asthma, bronchitis or pneumonia (a detailed list and ICD-9 codes is in Appendix A).²¹ We label these "Head Start susceptible causes" or "relevant causes." While most of these causes-of-death are now relatively rare, they were much more common during our observation period and are also generally more common in the South.²² We focus on mortality to children in the age group 5–9, which should capture ongoing health effects of Head Start services provided to children ages 3 or 4. (The alternative with the CMF would be to examine deaths to children ages 1–4, most of whom would be too young to have enrolled in Head Start). Our focus on mortality to children 5–9 also seems appropriate given that most of the causes of death susceptible to Head Start intervention are from chronic rather than acute conditions.

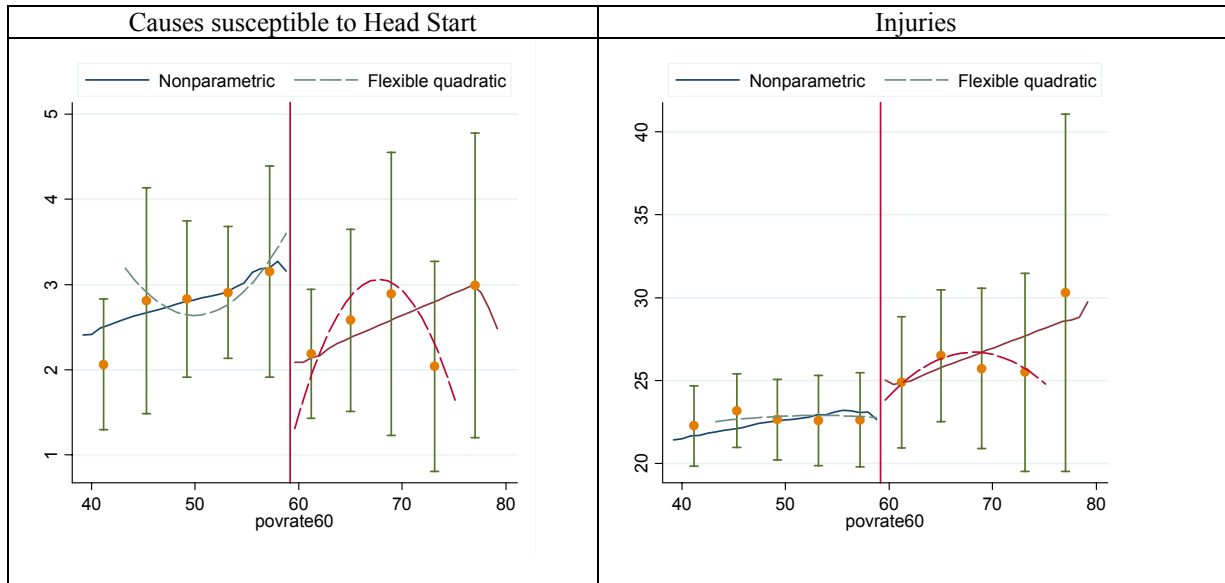
Table 3 presents evidence that the discontinuity in Head Start funding described above is mirrored by a discontinuity in child mortality rates over the period 1973–83 for Head Start susceptible causes. Graphs for the semi-parametric and quadratic polynomial estimates are presented in Figure 4. We choose this time period because by 1973 all children in the 5–9 age range would have been of Head Start age after the program was in existence, while the NELS results shown above suggest that this age group would have been exposed to the discontinuity in funding across the OEO cutoff through at least 1983.

²¹Thanks to Romi Webster, MD of Children's Hospital, Boston for helpful discussions.

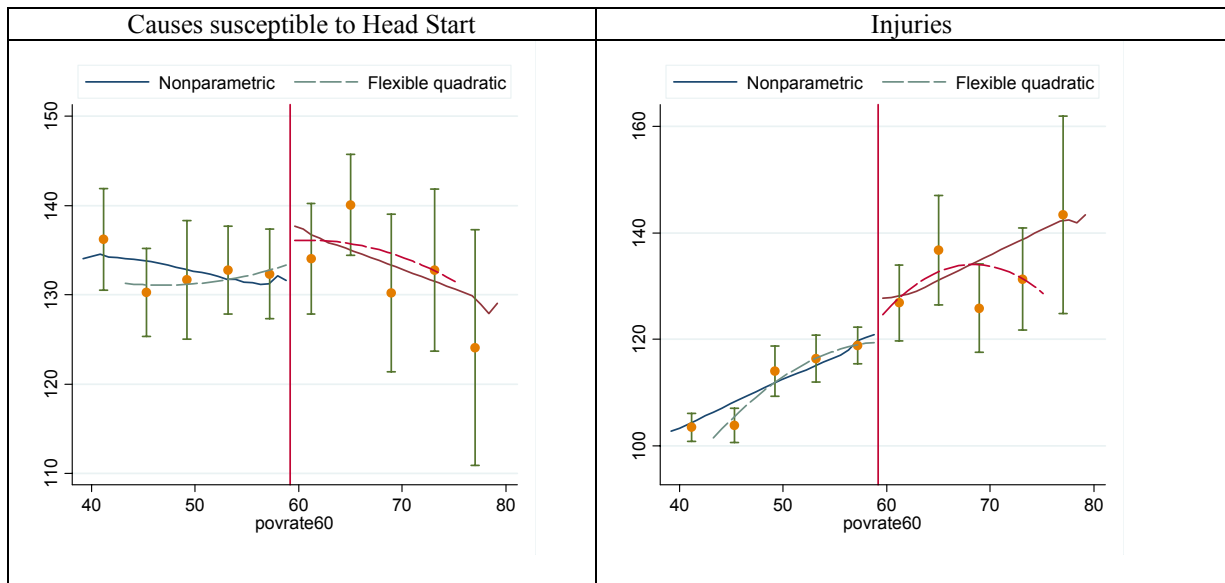
²²From 1960 to 2000 the all-age mortality rate in the U.S. declined by 80 percent for tuberculosis and 55 percent for influenza and pneumonia (edcp.org/tb/pdf/tb_Incidence_Rates_USMD_60-00.pdf and bcbs.com/mcrg/mcrg_stats.pdf, accessed 5/6/05). We assume these problems were more severe in the South during the 1960s and 1970s since mortality rates are still higher in the South today. For example Mississippi's mortality rate is 17 percent above the U.S. average, with infant mortality rates usually around 33–50 percent higher (US Statistical Abstracts, 1998, pp. 96, 98).

Figure 4: Mortality Rates per 100,000 from 1973-83, for Children and Adults, from Causes Affected by Head Start and from Injuries

Panel A: Children 5-9



Panel B: Adults 25+



Note: Panel A, Left: Estimated nonparametric discontinuity = 1.22 T-stat = 1.56, bandwidth = 8.
 Panel A, Right: Estimated nonparametric discontinuity = 2.31 T-stat = 0.95, bandwidth = 8.
 Panel B, Left: Estimated nonparametric discontinuity = 5.92 T-stat = 1.38, bandwidth = 8.
 Panel B, Right: Estimated nonparametric discontinuity = 7.34 T-stat = 1.65, bandwidth = 8.

Table 3
Regression Discontinuity Estimates of the Effect of Head Start Assistance on Mortality
Compressed Mortality Files, 1973–1983

Variable	LHS Mean	Porter				Parametric		
		4	6	8	10	16	Flexible Linear	Flexible Quadratic
Bandwidth		4	6	8	10	16	8	16
Poverty Range		524	734	958	1193	2155	484	863
Number of observations (counties) with nonzero weight								
<i>MAIN RESULTS</i>								
Ages 5–9, Head start related causes	3.415	-1.842* (0.954)	-1.711** (0.783)	-1.222 (0.779)	-1.167* (0.678)	-1.141** (0.528)	-2.201** (1.004)	-2.558** (1.261)
<i>SPECIFICATION CHECKS</i>								
Ages 5–9, Injuries	22.091	0.168 (3.373)	3.004 (2.745)	2.314 (2.413)	0.593 (2.187)	0.645 (1.732)	-0.164 (3.380)	0.775 (3.401)
Ages 5–9, All causes	39.828	-3.402 (4.195)	0.096 (3.417)	-0.166 (3.029)	-1.372 (2.772)	-1.646 (2.182)	-3.896 (4.268)	-2.927 (4.295)
Ages 25+, Head start related causes	132.892	2.678 (5.560)	3.949 (4.852)	5.916 (4.257)	6.834* (4.012)	5.703* (3.236)	2.091 (5.581)	2.574 (6.415)
Ages 25+, injuries	122.046	6.038 (6.348)	6.731 (5.192)	7.343* (4.438)	5.796 (4.098)	4.315 (3.152)	2.65 (6.206)	4.276 (6.059)
Whites age 5–9, Head start related causes	2.885	-1.052 (1.029)	-1.351 (0.868)	-0.919 (0.842)	-0.717 (0.732)	-0.774 (0.600)	-1.334 (1.061)	-1.746 (1.332)
Blacks age 5–9, Head start related causes	3.684	-2.169 (3.674)	-2.017 (2.741)	-2.703 (2.125)	-2.746 (1.741)	-1.535 (1.672)	-1.699 (4.094)	-1.93 (3.718)
Infants, Head start related causes	162.257	28.145 (21.499)	9.462 (18.463)	22.623 (16.692)	26.812* (14.955)	32.512** (11.162)	22.161 (21.610)	14.987 (22.825)

See note to Table 2 for methodology. Outcome of interest is 1-year mortality rates per 100,000 Head start related causes include deaths due to: tuberculosis, other infections, diabetes, nutritional causes, anemias, meningitis, and respiratory causes.

The first row of Table 3 shows that there is a proportionately very large discontinuity in mortality from relevant causes for children ages 5–9. The non-parametric estimates imply discontinuities equal to 33 percent to 54 percent of the control mean of 3.4 per 100,000, with t-statistics ranging from around 1.6 to 2.2. (The control mean implies that these causes of death account for about 9 percent of all deaths). The parametric estimates are larger (around 65–75 percent of the control mean) and significant at the usual 5 percent cutoff. This estimated impact is driven mostly by causes-of-death that children are tested and treated for (anemias, respiratory ailments, meningitis, and to a lesser extent diabetes) rather than vaccinated against (although some respiratory ailments can be prevented by DTP vaccination). The discontinuity for mortality from all causes for children 5–9 is usually negative and just slightly larger in absolute value than deaths from Head Start susceptible causes, but is much less precisely estimated.

One question is whether impacts of this proportional magnitude are plausible. For all counties in the U.S. from 1973–83 the mean and median mortality rate for children 5–9 from Head Start susceptible causes equaled around 2.0 and 1.8 per 100,000, respectively. So our estimates suggest that in the treatment counties near the cutoff, Head Start eliminates most of the “excess risk” of death from these causes (that is, the difference in mortality rates from the national average). Put differently, a 50–100 percent increase in Head Start funding leads to a 33–75 percent decline in mortality rates from relevant causes. This impact seems plausible given that poverty is a powerful risk factor for child health (Sing and Yu, 1996, Case, Lubotsky and Paxson, 2002).

Some support for the idea that these mortality differences are due to Head Start rather than other factors comes from the fact that we do not observe a similar discontinuity in mortality rates for children 5–9 from causes that should not be affected by Head Start—namely, injuries (second panel, Table 3).²³ Although we might expect Head Start enrollment to affect contemporaneous injury rate (see for example

²³In principle Head Start could have changed parent behaviors to affect injury rates. However in practice, as North (1979, p. 245) notes, “Head Start’s obvious potential for health education was, unfortunately, never attained.” Health education for parents often emphasizes nutrition and hygiene (O’Brien, Connell, and Griffin, 2004, p. 173).

Currie and Hotz, 2004), we do not expect the program to affect accident rates for children ages 5–9. The estimated discontinuities for injuries range from +8.8 percent to -8.7 percent of the control mean and are never close to statistically significant. Similarly there is no discontinuity in mortality from either relevant causes or injuries for people 25+ who would have been too old in 1973–83 to have received health services from Head Start.

The fifth and sixth rows of Table 3 show that the discontinuities in mortality for children 5–9 from causes plausibly affected by Head Start are found among both blacks and whites. The magnitudes of the point estimates (as a share of the control mean) are somewhat larger for blacks than whites with the semi-parametric estimates, although the reverse is true with our parametric results. In any case strong conclusions about any race differences in Head Start’s impact on child health are precluded by the fairly large standard errors for our estimates for blacks.

A final potential concern comes from the desegregation of hospitals in the South (particularly Mississippi) following the 1964 Civil Rights Act, which led to large declines in mortality to black infants (Almond, Chay and Greenstone, 2003). However hospital integration would only be a problem for our identification strategy if integration varied discontinuously at exactly the OEO cutoff used to allocate Head Start grant-writing assistance. Some evidence against this counter-explanation for our findings comes from the fact that the reduction in mortality among children 5–9 is not limited to blacks. Moreover while Almond *et al.* show large declines in deaths to infants or post-neonates from hospital desegregation, Table 3 shows the discontinuities at the OEO cutoff in infant mortality from Head Start-affected causes are relatively small (5–20 percent of the control mean), *positive* and usually not significant.²⁴

²⁴Restricting attention to just post-neonates yields similar results although the population data available with the CMF for 1973–83 does not allow us to restrict the denominator to children 1–12 months of age.

VII. HEAD START AND EDUCATIONAL ATTAINMENT

In this section we show that there is a discontinuity in educational attainment at the OEO cutoff in the 1990 Census among those cohorts born late enough to have been affected by Head Start, but not among earlier cohorts. The main drawback with the Census is the possibility of selective migration as people move across counties from childhood to adulthood (when we observe their place of residence). We show that the discontinuity in schooling outcomes for the Census is mirrored among respondents to the NELS, for whom we observe addresses at age 13.

A. Census Results

Estimates for educational attainment by cohort from the 1990 Census are presented in Table 4 and Figures 5 and 6. 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 while the program was in existence. We call this the “directly treated” group. Also of potential 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 parents of Head Start participants, who were also involved in the program in a variety of ways, and older siblings who might have benefited from spillover effects. We call this group the “indirectly treated” group. People 55 and over are unlikely to have been parents of Head Start participants and so are “untreated.”²⁵

The top left panel of Figure 5 summarizes our results for completion of high school diploma or equivalent for people ages 18–24 in 1990 (our directly treated group), while the top panel of Table 4

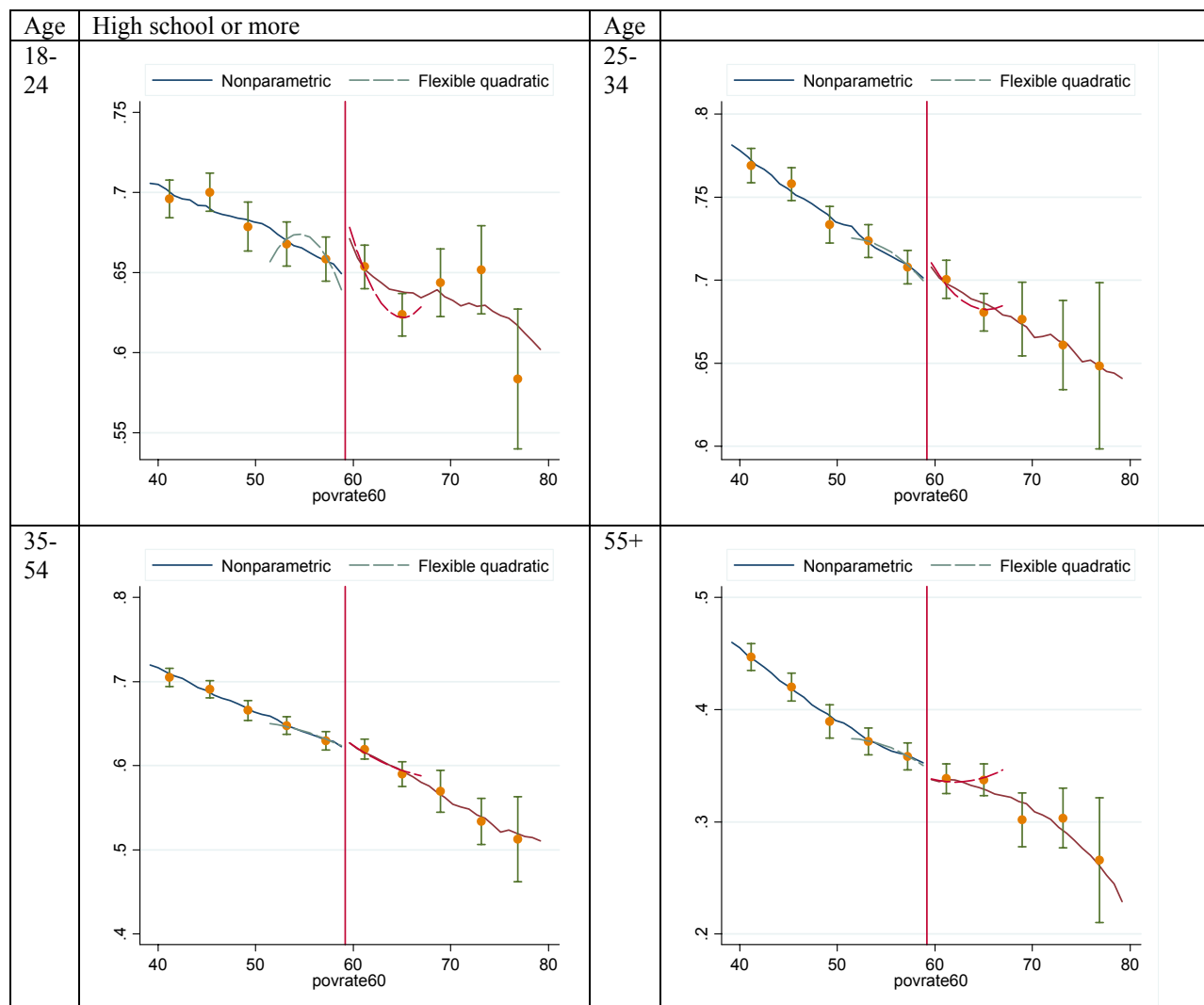
²⁵By 1990 the oldest person who could have participated in Head Start would have been about 30 years of age. Data from the Vital Statistics for 1960 suggests that about one-half of all births that year were to parents 25 or older (55+ in 1990), and only about 1/4 were to parents 30+ (HEW, 1960). This 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.

Table 4
Regression Discontinuity Estimates of the Effect of Head Start Assistance on Educational Outcomes, Directly Treated Group
 1990 Census

Variable	Porter						Parametric	
	LHS Mean	1.5	2.5	3	3.5	6	Flexible Linear	Flexible Quadratic
Bandwidth		202	343	411	476	734	244	484
Poverty Range							3	6
Number of observations (counties) with nonzero weight								
<i>ALL RACES</i>								
% High School or Equivalent, ages 18–24	0.640	0.038* (0.021)	0.034** (0.017)	0.029* (0.016)	0.019 (0.015)	0.003 (0.012)	0.043** (0.019)	0.055** (0.020)
% Some College, ages 18–24	0.227	0.044 (0.027)	0.045** (0.021)	0.037* (0.020)	0.029 (0.018)	0.011 (0.014)	0.051** (0.024)	0.057** (0.025)
% College completion, ages 18–24	0.022	0.002 (0.004)	0.002 (0.004)	0.004 (0.004)	0.002 (0.003)	0 (0.003)	0.004 (0.003)	0.004 (0.005)
% in same state of birth, ages 18–24	0.796	-0.033 (0.031)	-0.046* (0.026)	-0.039 (0.024)	-0.033 (0.023)	-0.023 (0.019)	-0.043 (0.028)	-0.042 (0.032)
<i>BLACKS ONLY</i>								
% High School or Equivalent, ages 18–24	0.559	0.057 (0.062)	0.072 (0.046)	0.038 (0.042)	0.016 (0.039)	-0.011 (0.031)	0.081 (0.051)	0.118** (0.053)
% Some College, ages 18–24	0.16	0.105* (0.059)	0.121** (0.048)	0.089** (0.042)	0.057 (0.039)	0.022 (0.031)	0.127** (0.050)	0.155** (0.052)
% College completion, ages 18–24	0.012	0.002 (0.006)	-0.003 (0.006)	-0.001 (0.005)	-0.001 (0.005)	-0.003 (0.004)	0.002 (0.006)	-0.001 (0.007)
% in same state of birth, ages 18–24	0.9	-0.09* (0.054)	-0.117** (0.045)	-0.111** (0.040)	-0.108** (0.036)	-0.069** (0.029)	-0.101** (0.045)	-0.105** (0.048)

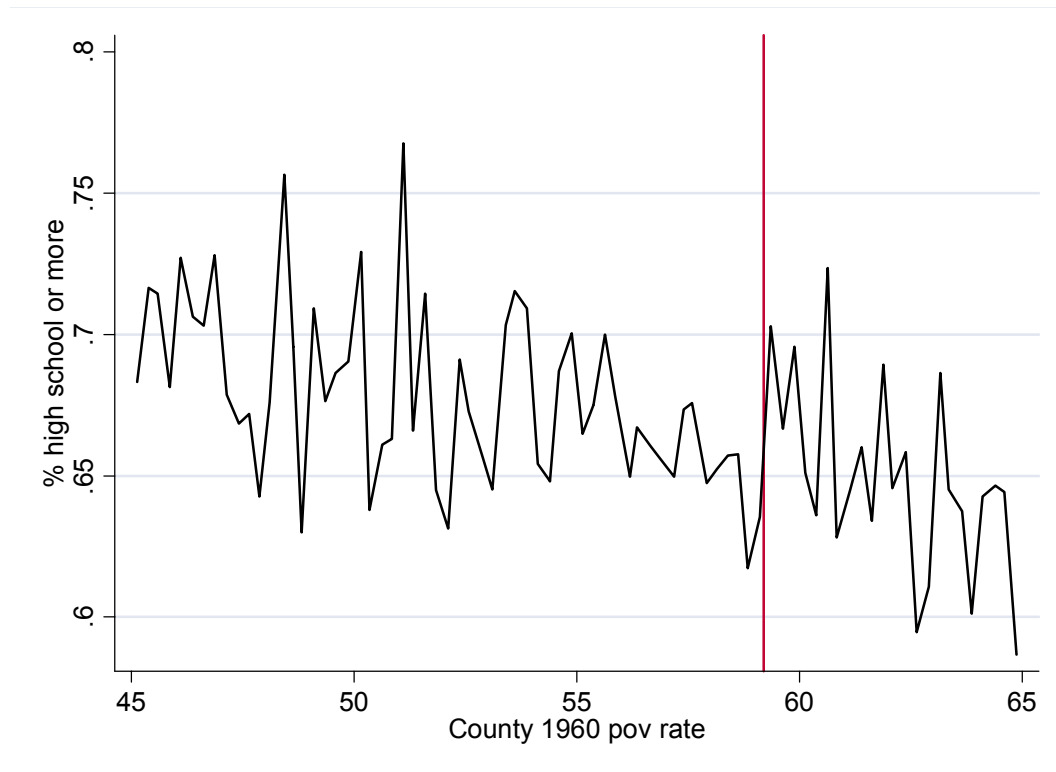
See note to Table 2.

Figure 5: Discontinuity in High School Completion by Age, 1990 Census



Note: Ages 18-24: Estimated nonparametric discontinuity = .029 T-stat = 1.85, bandwidth = 3.
 Ages 25-34: Estimated nonparametric discontinuity = .011 T-stat = 0.90, bandwidth = 3.
 Ages 35-54: Estimated nonparametric discontinuity = .007 T-stat = 0.56, bandwidth = 3.
 Ages 55+: Estimated nonparametric discontinuity = .016 T-stat = 1.11, bandwidth = 3.

Figure 6
Histogram of High School Graduation Rates, 18-24 Year olds in 1990 Census



Notes: Figure presents mean high school graduation rates for 18-24 year olds in the 1990 Census, with means calculated using a bin width of .25.

summarizes our full set of results for alternative estimation approaches. There is a relatively large discontinuity in high school completion for this directly treated group that ranges from around 2 to 5 percentage points across most of our different estimation approaches (the one exception being our semi-parametric estimator at the largest bandwidth). If Head Start spending is 50 percent to 100 percent higher in treatment than control counties our results imply an elasticity of high school completion to Head Start spending of around $+0.05$ to $.15$. While the smoothed trends and cell means in Figure 5 convey the visual impression that our discontinuity estimate is driven by the functional form of $m(P)$ —specifically, the “flattening out” of the general downward trend in high school completion as 1960 county poverty increases—this impression is incorrect: Figure 6 gives a less-aggregated picture of the data (a histogram calculated with a bandwidth of 0.25) and shows a positive jump in high school completion rates at the OEO cutoff.

Note that in principle these schooling results may be biased by the Head Start effects on mortality documented in the previous section. However any compositional bias is unlikely to be important in practice given that the proportionately large mortality effect is still small in an absolute sense, equal to around 1 or 2 deaths per 100,000 children. In comparison the estimated discontinuity for high school completion equals between 2,000 and 5,000 per 100,000.

For college attendance the estimates are typically between 3 and 6 percentage points (again smaller with a large bandwidth in the semi-parametric estimator), compared to a control mean of 23 points. We do not find a discontinuity in college completion rates for this group, not surprising given that most in this group will be too young to have graduated from college.

The second panel of Table 4 shows that the point estimates for high school completion and college attendance are typically about twice as large for blacks as for the full sample, although these estimates have large standard errors. Nevertheless the results are intriguing in part because they differ from the sibling-difference estimates from CT and GTC, which show long-term schooling impacts from Head Start for whites but not blacks.

Table 5 shows that the estimated discontinuities in educational attainment are much smaller for older age groups than for the directly treated group of 18–24 year olds shown in Table 4. We find some evidence of statistically significant effects on high school completion and college attendance for the partially directly treated group (25–34) and the indirectly treated group (35–54), which are typically less than half as large as the estimates for our directly treated group. In contrast, there are no discontinuities in educational attainment for any of our schooling outcomes or estimation approaches for the “untreated” group (55 and older in 1990). This pattern of estimates across cohorts in Tables 4 and 5 is consistent with what we would predict if Head Start has a positive schooling effect for participants and a more modest impact on their siblings and parents.

One concern with these results comes from the possibility of migration across counties between when people were of Head Start age (3–4) and when they are observed as adults in the 1990 Census. If exposure to Head Start made individuals more likely to move out the state, then we might expect to find a discontinuity in the fraction of individuals born in the same state. The estimated discontinuities in the share of residents born in the state are usually not statistically significant (Tables 4 and 5), but they tend to be somewhat larger in absolute value for younger than older age groups, and larger for blacks than with the full-race sample.

B. NELS Results

One way to address concerns about selective migration is to replicate our schooling estimates using the NELS, which allows us to identify county of residence for respondents in 8th grade when most students were around 13 years old, about 8–10 years after Head Start age.

Figure 7 and Table 6 show that the results from the NELS are qualitatively similar to those from the 1990 Census presented above, using the schooling variables measured in the 2000 wave of the NELS when respondents were around 25. Using a bandwidth of 6 yields a semi-parametric estimate for completion of a high school diploma of around 19 points (around 1/3 of the control mean), with a t-statistic of 2.0. Larger bandwidths yield smaller semi-parametric estimates, although the parametric

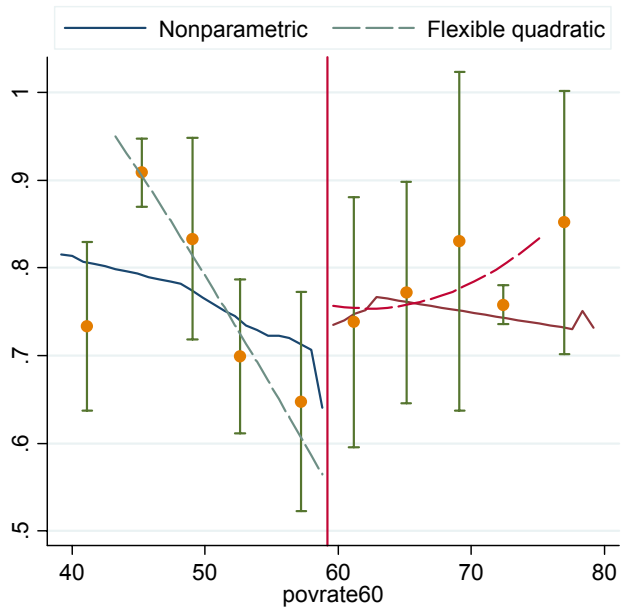
Table 5
Regression Discontinuity Estimates of the Effect of Head Start Assistance on Educational Outcomes, Older Age Groups
 1990 Census

Variable	Porter						Parametric	
	LHS Mean	1.5	2.5	3	3.5	6	Flexible Linear	Flexible Quadratic
Bandwidth		202	343	411	476	734	244	484
Poverty Range							3	6
Number of observations (counties) with nonzero weight								
% High School or Equivalent, ages 25–34	0.702	0.006 (0.018)	0.009 (0.013)	0.011 (0.013)	0.006 (0.012)	0.005 (0.010)	0.005 (0.015)	0.018 (0.016)
% Some College, ages 25–34	0.206	0.005 (0.016)	0.014 (0.013)	0.016 (0.012)	0.011 (0.011)	0.011 (0.009)	0.012 (0.014)	0.014 (0.016)
% College completion, ages 25–34	0.062	0.011 (0.008)	0.01 (0.007)	0.01 (0.007)	0.007 (0.006)	-0.001 (0.005)	0.011 (0.008)	0.014* (0.008)
% in same state of birth, ages 25–34	0.754	-0.014 (0.032)	-0.022 (0.027)	-0.009 (0.024)	0.003 (0.023)	0 (0.019)	-0.02 (0.028)	-0.025 (0.032)
% High School or Equivalent, ages 35–54	0.624	-0.011 (0.019)	0.011 (0.015)	0.008 (0.014)	0.001 (0.013)	0.013 (0.011)	-0.004 (0.016)	0.009 (0.018)
% Some College, ages 35–54	0.181	0.008 (0.013)	0.019* (0.011)	0.014 (0.010)	0.007 (0.009)	0.01 (0.008)	0.017 (0.011)	0.021 (0.013)
% College completion, ages 35–54	0.062	0.005 (0.007)	0.007 (0.006)	0.008 (0.005)	0.004 (0.005)	0.003 (0.004)	0.008 (0.006)	0.01 (0.007)
% in same state of birth, ages 35–54	0.786	-0.009 (0.029)	-0.02 (0.024)	-0.005 (0.022)	0.003 (0.021)	-0.004 (0.018)	-0.014 (0.026)	-0.016 (0.029)
% High School or Equivalent, ages 55+	0.35	-0.023 (0.021)	-0.007 (0.016)	-0.016 (0.015)	-0.016 (0.014)	0.003 (0.011)	-0.011 (0.018)	-0.008 (0.019)
% Some College, ages 55+	0.092	0.001 (0.009)	0.001 (0.007)	-0.002 (0.006)	-0.003 (0.006)	0.002 (0.005)	0.002 (0.007)	0.002 (0.008)
% College completion, ages 55+	0.043	-0.005 (0.005)	-0.002 (0.004)	-0.002 (0.004)	-0.002 (0.003)	0.001 (0.003)	-0.002 (0.004)	-0.003 (0.005)
% in same state of birth, ages 55+	0.794	0.014 (0.036)	0.001 (0.029)	0.012 (0.026)	0.002 (0.024)	-0.008 (0.020)	0.009 (0.032)	0.027 (0.034)

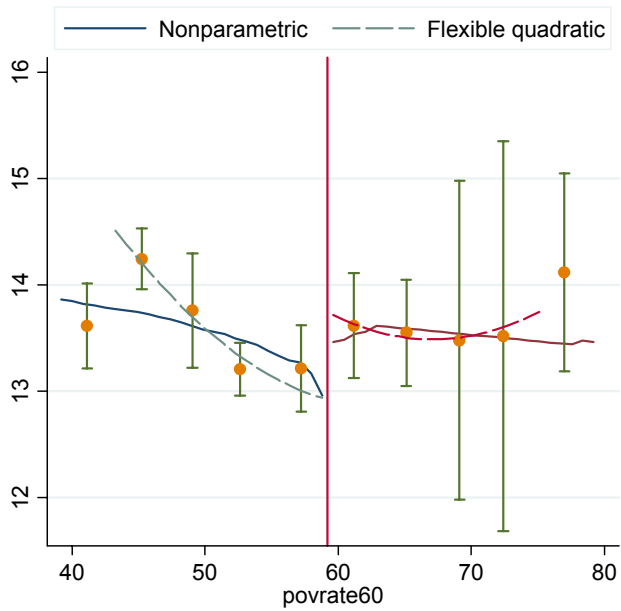
See note to Table 2

Figure 7
Discontinuity in Schooling, NELS

Panel A: High school or more (2000 survey)



Panel B: Years of school completed (2000 survey)



Note: Panel A Estimated nonparametric discontinuity = .065 T-stat = 1.10, bandwidth = 8.

Panel B Estimated nonparametric discontinuity = .602 T-stat = 1.74, bandwidth = 8.

Table 6
NELS Outcomes
 NELS

Variable	Porter					Parametric		
	LHS Mean	4	6	8	10	16	Flexible Linear	Flexible Quadratic
Bandwidth		4	6	8	10	16	8	16
Poverty Range		43	64	95	125	281	43	82
Number of observations (counties) with nonzero weight		43	64	95	125	281	43	82
Completed high school, 2000	0.551	0.171 (0.122)	0.189** (0.093)	0.12 (0.077)	0.09 (0.076)	0.03 (0.052)	0.171 (0.137)	0.205 (0.127)
High school or equivalent, 2000	0.731	0.134 (0.103)	0.123 (0.075)	0.066 (0.059)	0.072 (0.055)	0.032 (0.035)	0.134 (0.120)	0.183* (0.104)
Some post-secondary school, 2000	0.515	0.084 (0.128)	0.257** (0.106)	0.157* (0.094)	0.159* (0.086)	0.098 (0.061)	0.084 (0.138)	0.165 (0.151)
Years of schooling, 2000	12.691	0.581 (0.414)	1.082** (0.372)	0.602** (0.345)	0.462 (0.324)	0.207 (0.227)	0.581 (0.449)	0.821 (0.563)
Enrolled in school, 2000	0.086	0.104 (0.090)	0.148* (0.087)	0.068 (0.073)	0.072 (0.061)	0.108** (0.043)	0.104 (0.100)	0.183 (0.131)

See note to Table 2.

estimates are as large or larger (though imprecise). The results in the next row for completion of either a high school degree or GED are somewhat smaller. This suggests that part of Head Start's impact is to move people along the diploma versus GED margin, which is consistent with what has been found for Perry Preschool (Schweinhart et al., 2005) and important given evidence that GED-holders do not fare as well as high school graduates in the labor market (see Carneiro and Heckman, 2003).

The patterns for post-secondary school attendance are similar to those for high school (second row). Among NELS respondents in the 600 poorest counties, about half of post-secondary school attendance does not result in completion of a degree or certificate, while 15–20 percent results in a certificate or license, around 10 percent results in an associate's degree and only around 20 percent results in the completion of a bachelor's degree or more. Overall the NELS data suggest a discontinuity in educational attainment of between .5 and 1.0 years of school (third row). Given that the NELS has small samples near the OEO cutoff, disaggregating further by race with NELS yields very noisy estimates.

The NELS results for high school or equivalent or college attendance are usually on the order of 2–3 times as large as the point estimates from the Census, although our estimates for the effects of doubling Head Start funding per county on these measures are of about the same magnitude as those estimated by GTC for Head Start participation for whites. The difference in our estimates from the NELS versus Census data could be due in part to greater attenuation with the Census results from migration across counties. Also possible are cohort differences in Head Start effects: NELS respondents would have been younger in 1990 than 18–24, and we can see from comparing Tables 4 and 6 that the control means for completion of high school or equivalent are about 10 percentage points higher in the NELS than for the Census cohorts.

Although in principle we might wish to examine other long-term outcomes, most other long-term outcome measures of interest are problematic with our data. For example labor market outcomes for the directly treated cohorts in the Census and NELS (in their 20s) are not discontinuous at the OEO cutoff because as Table 6 shows, a larger share of people in treatment counties appears to be still enrolled in

school at the time.²⁶ In sum, where our data are strongest—for schooling and especially mortality—estimated Head Start effects are strongest as well.

VIII. EXTENSIONS AND SPECIFICATION CHECKS

In this section we address three concerns with the findings presented above: (1) the possibility of selective migration across counties between early childhood and when we first observe people's county of residence; (2) whether our estimates are artifacts of functional form assumptions; and (3) whether there are discontinuities in other social spending at the OEO cutoff.

A. Selective Migration

If migration were random (independent of Head Start participation or outcomes) then our estimates for program impacts would be attenuated, since treatment or control county residence is a noisy measure of actual exposure. More worrisome is the possibility of non-random migration, which is not a real concern with our mortality results (given we measure outcomes at ages 5–9, just a few years after Head Start participation) although may be more of a problem with our schooling estimates. But there are a few reasons to believe that our schooling estimates are not simply artifacts of selective migration across counties.

One data point against the selective-migration hypothesis comes from the qualitative similarity in results between the 1990 Census, which assigns people to treatment and control counties based on county of residence during adulthood, and the NELS, which assigns people to counties based on where they live

²⁶In other results not shown we have used the NELS to examine achievement test scores in 8th and 10th grade as well as self-reported arrests. The available arrest data in the NELS are self-reported and so may substantially understate youth involvement with crime (Kling, Ludwig and Katz, 2005). The test score data are limited by the fact that the NELS sampling frame excludes many children in special education, and so any effect of Head Start to reduce the chances children are in special education (as is found with more intensive early childhood programs such as Abecedarian or Perry Preschool) may depress average test scores by including more marginal students.

at age 13. Migration across poor counties between ages 3–4 and age 13 appears to be modest at least according to the PSID, which provides small and unrepresentative samples “close” to the OEO cutoff but at least lets us compare people in all 300 poorest counties with the next 300 poorest counties. Of the respondents in the 300 poorest counties at age 3, fully 86 percent were in one of these counties at age 13; of those in one of these counties at age 13, 64 percent were in one at age 3. Importantly, the results are similar for PSID respondents in the next 300 poorest counties, and if anything slightly more low-education people are leaving the “control” than “treatment” counties. With the caveats about the PSID sample in these counties kept in mind, these data do not seem consistent with a selective migration explanation for the schooling findings reported above.^{27,28}

B. Specification Tests

Another way to test for selective migration is to test for balance at the discontinuity in exogenous characteristics that should not be affected by the intervention, since selective migration with respect to such characteristics would lead to imbalance. As McCrary and Royer (2003) note, the specification tests available with RD designs are also useful for examining the more general continuity assumptions that provide identification. Under our identifying assumptions about smoothness we would expect no discontinuity in exogenous characteristics at the OEO cutoff. Just as in an experimental setting, we can

²⁷Of those at age 3 in counties with 1960 poverty rates that rank them 301–600, 78 percent were in one of these counties at age 13. Of those in one of counties ranked 301–600 at age 13, 71 percent were in such a county at age 3. For completion of high school diploma or equivalent, we find identical rates for in-movers into the treatment and control counties (92 percent), defined as those in a treatment or control county at age 13 but not age 3. The high school completion rate is lower for out-movers from the control than treatment counties (58 versus 91 percent), and consistent with the idea of more out-migration of low-education people from controls, the high school completion rate among stayers is slightly higher in the control than treatment counties (88 versus 79 percent). However restricting attention to PSID respondents for whom we observe addresses at ages 3 and 13 and educational outcomes leaves us with just 100 people in the poorest 300 counties at age 13 and 128 in the next 300 poorest counties at age 13.

²⁸The NELS data suggest that around 70 percent of people living in counties with poverty rates just above (or below) the OEO cutoff at age 13 are in the same set of counties at age 25 in the 2000 wave of NELS.

test the assumption that the assignment to treatment was as good as “random” by checking for balance on the covariates.

On the whole, exogenous covariates seem fairly well balanced at the OEO cutoff (Table 7). There is no systematic evidence for discontinuities in any of the Censuses from 1950 to 1990 in total county population (in thousands), percent black, or percent urban, measures that are often associated with economic outcomes and may change within a county in response to selective migration. Table 8 shows that with the NELS there are a few model specifications that yield significant point estimates for one exogenous characteristic (either black or urban), but there is no bandwidth that suggests imbalance across multiple exogenous characteristics. NELS is also balanced for parent outcomes like maternal schooling that predict children’s outcomes, which is consistent with the very modest effects on schooling for the “indirectly treated” cohorts in the 1990 Census and which is not consistent with a story in which parents systematically move across counties in ways related to children’s outcomes.

Another specification check is suggested by the fact that the across-county funding discontinuity at the OEO cutoff seems to have dissipated over time (see for example the recommendations in GAO, 1981), and so among kindergartners in 1998 who were surveyed as part of the ECLS-K we do not observe any discontinuity in Head Start participation rates. Reassuringly, we also do not observe discontinuities in children’s test scores at the OEO cutoff, although we note the number of counties sampled by the ECLS-K is smaller than with the NELS.

Yet another specification test comes from examining whether there are statistically significant discontinuities in outcomes that may be affected by Head Start at other “pseudo-cutoffs,” which might be the case if our estimates are an artifact of the functional form assumptions behind our estimation approach. Table 9 shows the results using a pseudo-cutoff of a 1960 county poverty rate of 40 percent, chosen (before looking at any of our outcome data) because we do not see any evidence of discontinuities

Table 7
Specification Tests with Census Data
 1990 Census

Variable	Porter						Parametric	
	LHS Mean	1.5	2.5	3.0	3.5	6.0	Flexible Linear	Flexible Quadratic
Bandwidth		202	343	411	476	734	244	484
Poverty Range							3	6
Number of observations (counties) with nonzero weight								
Population (1000s), 1950	22.654	-3.202 (5.062)	1.101 (3.587)	1.553 (3.523)	2.543 (3.261)	1.183 (2.437)	-2.389 (4.027)	-2.126 (4.437)
Population, 1960	23.476	-6.447 (7.480)	-0.848 (4.753)	0.063 (4.515)	0.876 (4.067)	0.425 (2.885)	-4.827 (5.861)	-4.783 (5.679)
Population, 1970	18.977	-2.175 (5.081)	0.137 (3.544)	0.933 (3.908)	1.835 (3.750)	0.178 (2.738)	-1.494 (4.002)	-2.259 (5.099)
Population, 1980	20.855	0.988 (6.132)	5.348 (4.244)	4.323 (4.838)	4.362 (4.497)	2.077 (3.242)	3.108 (4.834)	3.148 (6.068)
Population, 1990	21.856	1.262 (7.105)	5.99 (4.909)	4.02 (5.964)	3.993 (5.505)	1.788 (3.935)	4.043 (5.641)	3.768 (7.397)
Percent non-white, 1950	0.25	-0.007 (0.062)	0.031 (0.050)	0.034 (0.046)	0.017 (0.042)	0.036 (0.034)	-0.002 (0.052)	0.038 (0.058)
Percent black, 1960	0.253	-0.026 (0.060)	0.006 (0.048)	0.011 (0.044)	0.003 (0.041)	0.014 (0.033)	-0.019 (0.051)	0.014 (0.055)
Percent black, 1970	0.219	-0.021 (0.059)	-0.022 (0.046)	-0.011 (0.042)	-0.017 (0.039)	-0.008 (0.031)	-0.04 (0.049)	0 (0.052)
Percent black, 1980	0.217	-0.034 (0.054)	-0.004 (0.043)	0.005 (0.039)	0.005 (0.036)	0.011 (0.029)	-0.036 (0.045)	-0.006 (0.049)
Percent black, 1990	0.215	-0.039 (0.054)	-0.009 (0.043)	0.007 (0.039)	0.007 (0.036)	0.013 (0.029)	-0.044 (0.045)	-0.015 (0.049)
Percent urban, 1950	0.138	0.014 (0.047)	0.015 (0.036)	0.021 (0.034)	0.028 (0.031)	0.003 (0.025)	0.003 (0.040)	0.011 (0.044)
Percent urban, 1960	0.164	0.028 (0.053)	0.002 (0.041)	0.015 (0.039)	0.024 (0.036)	-0.004 (0.030)	0.011 (0.045)	-0.003 (0.050)
Percent urban, 1970	0.177	0.041 (0.059)	-0.007 (0.047)	0.011 (0.044)	0.007 (0.040)	-0.015 (0.033)	0.021 (0.050)	0.016 (0.055)
Percent urban, 1980	0.014	0.002 (0.026)	0.005 (0.020)	0.012 (0.021)	0.006 (0.019)	0.009 (0.016)	0.012 (0.022)	0.008 (0.026)
Percent urban, 1990	0.016	0.016 (0.028)	0.016 (0.023)	0.022 (0.024)	0.013 (0.021)	0.008 (0.018)	0.025 (0.024)	0.023 (0.029)

Notes: See Table 2. Population figures in thousands.

Table 8
Specification Tests with NELS Data
 NELS

Variable	Porter				Parametric			
	LHS Mean	4	6	8	10	16	Flexible Linear	Flexible Quadratic
Bandwidth							8	16
Poverty Range								
Number of observations (counties) with nonzero weight		43	64	95	125	281	43	82
Maternal schooling (years)	13.088	-0.068 (0.274)	0.266 (0.289)	0.267 (0.274)	0.113 (0.231)	0.045 (0.175)	-0.068 (0.312)	-0.138 (0.553)
Black	0.426	-0.095 (0.158)	-0.214* (0.127)	-0.191* (0.102)	-0.117 (0.088)	-0.103* (0.056)	-0.095 (0.182)	-0.146 (0.186)
Hispanic	0.122	-0.085 (0.140)	0.03 (0.105)	0.044 (0.088)	0.029 (0.079)	0.059 (0.060)	-0.085 (0.153)	-0.054 (0.130)
Urban	0.033	-0.103** (0.045)	-0.024 (0.065)	-0.004 (0.069)	0.034 (0.073)	0.042 (0.071)	-0.103** (0.049)	-0.115 (0.159)

See note to Table 2.

Table 9
Regression Discontinuity Estimates, False cutoff (1960 county poverty rate = 40%)
 1990 Census

Variable	LHS Mean	Porter				Parametric		
		4	6	8	10	16	Flexible Linear	Flexible Quadratic
Mortality outcomes								
Bandwidth		4	6	8	10	16		
Poverty Range							8	16
Number of observations (counties) with nonzero weight		922	1413	1915	2392	3108	825	1716
Ages 5–9, Head start related causes	2.7	-0.280 (1.052)	-0.431 (0.765)	-0.330 (0.612)	-0.226 (0.538)	0.092 (0.447)	-0.251 (1.124)	-0.371 (0.962)
Ages 5–9, Injuries	22.099	-0.054 (2.396)	-0.050 (1.882)	0.321 (1.634)	0.252 (1.499)	-0.449 (1.201)	-0.174 (2.488)	-1.235 (2.513)
Ages 25+, Head start related causes	138.395	-1.922 (5.641)	-4.229 (4.325)	-5.347 (3.632)	-8.931** (3.173)	-15.818** (2.461)	-1.599 (5.849)	-2.009 (5.654)

(table continues)

Table 9, continued

Variable	1990 Census Educational Outcomes					
	Porter			Parametric		
LHS Mean	1.5	2.5	3	3.5	6	6
	328	570	699	811	1422	828
Bandwidth						
Poverty Range						
Number of observations (counties) with nonzero weight						
% High School or equivalent, ages 18–24	0.69	0.011 (0.021)	0.008 (0.015)	0.008 (0.014)	0.011 (0.010)	0.026 (0.018)
% High School or equivalent, ages 25–34	0.787	0.032* (0.019)	0.009 (0.013)	0.014 (0.011)	0.004 (0.008)	0.035** (0.016)
% High School or equivalent, ages 35–54	0.767	0.019 (0.019)	-0.001 (0.013)	0.005 (0.012)	0.003 (0.009)	0.023 (0.016)
% High School or equivalent, ages 55+	0.472	0.032 (0.021)	0.017 (0.015)	0.021 (0.013)	0.009 (0.010)	0.031* (0.018)
% Some College, ages 18–24	0.313	0.003 (0.025)	-0.012 (0.019)	-0.011 (0.018)	-0.008 (0.014)	-0.001 (0.021)
% Some College, ages 25–34	0.374	0.039* (0.023)	0.007 (0.017)	0.009 (0.015)	-0.004 (0.011)	0.04** (0.019)
% Some College, ages 35–54	0.389	0.023 (0.021)	-0.005 (0.016)	0.000 (0.014)	-0.004 (0.011)	0.018 (0.018)
% Some College, ages 55+	0.2	0.014 (0.015)	0.004 (0.011)	0.007 (0.010)	-0.001 (0.008)	0.015 (0.013)
% College completion, ages 18–24	0.031	0.011** (0.005)	0.009** (0.004)	0.009** (0.004)	0.005* (0.003)	0.012** (0.004)
% College completion, ages 25–34	0.115	0.017 (0.012)	0.008 (0.010)	0.006 (0.009)	0.000 (0.007)	0.016 (0.010)
% College completion, ages 35–54	0.15	0.005 (0.013)	0.000 (0.010)	0.000 (0.009)	-0.003 (0.007)	0.006 (0.011)
% College completion, ages 55+	0.077	0.002 (0.008)	0.001 (0.006)	0.000 (0.006)	-0.003 (0.005)	0.002 (0.007)

(table continues)

Table 9, continued
NELS

Variable	Porter				Parametric		
	4	6	8	10	16	Flexible Linear	Flexible Quadratic
Bandwidth	126	201	277	349	571	16	16
Poverty Range							
Number of observations (counties) with nonzero weight							
	126	201	277	349	571	111	243
LHS Mean							
High school or equivalent, 2000	0.873	-0.073 (0.069)	-0.017 (0.051)	-0.016 (0.045)	-0.024 (0.032)	-0.214** (0.068)	-0.272** (0.076)
Some post-secondary school, 2000	0.789	-0.044 (0.073)	-0.003 (0.050)	-0.042 (0.044)	-0.017 (0.033)	-0.056 (0.072)	-0.082 (0.079)
Years of schooling, 2000	14.04	-0.108 (0.319)	0.174 (0.208)	0.007 (0.178)	0.096 (0.134)	-0.192 (0.319)	-0.392 (0.333)
Still in school, 2000	0.109	0.011 (0.055)	0.037 (0.039)	0.02 (0.034)	0.002 (0.026)	0.027 (0.057)	0.061 (0.060)

See note to Table 2

in Head Start spending here.²⁹ For the mortality variables, 2 out of 21 point estimates are statistically significant at the 5 percent level at our pseudo-cutoff. With our Census data only 5 out of 85 point estimates are statistically significant, and the significant point estimates are not concentrated among cohorts exposed to Head Start as with our estimates for the true OEO cutoff. The only significant point estimate in the NELS is for high school completion and is of the opposite sign to what we find at the actual OEO cutoff.

C. Endogenous Break Point Tests

We have strong *a priori* reasons to believe that the discontinuity in Head Start assistance occurred at the 300th poorest county. However, we are also interested in whether the data offers supportive evidence for identifying the location of the threshold. In contrast to standard parametric goodness-of-fit tests for break points (Piehl et. al. 2003, Bai and Perron 1998), in our application computing the degrees-of-freedom correction is complicated by the “local” nature of our estimation and by the non-uniform density of county poverty. To the best of our knowledge there is no literature on threshold or break-point tests for non-parametric RD models. As such, we performed a goodness-of-fit exercise that is in the same spirit of the parametric methods.

We examined each potential threshold from the 50th percentile of 1960 county poverty rates (31 percent poor) through the 95th percentile (74 percent poor), in increments of .2 percent poor. For each potential threshold, we estimated our discontinuity model at that point. We then estimated a nonparametric locally-linear (Fan 1992) model at each data point, and came up with a predicted value. Points to the left of the potential discontinuity were predicted using only data to the left of the potential discontinuity, and similarly for points to the right. Based on these predicted values, we created a measure

²⁹We must be careful in conducting this sort of specification check because there are significant differences in Head Start funding across counties even away from the OEO cutoff (Figure 1). If we arbitrarily choose a cutoff where there is an actual program funding difference we might detect in part the effects of Head Start funding.

of loss, equal to the sum of squared errors over the whole data set. We then examined how this loss function varied across potential discontinuity points.

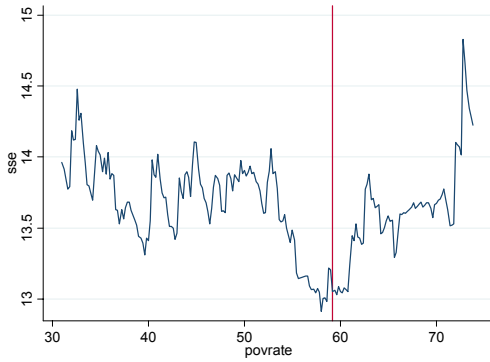
We performed this series of calculations for each of our three first stage variables: NELS Head Start attendance, and 1968 and 1972 NARA records of Head Start funding per child. Plots of the loss functions are presented in Figure 8, where lower values indicate better fit. For all three variables, the loss function makes a noticeable improvement in the neighborhood of the discontinuity. This improvement typically covers an area of plus or minus 3 percentage points around the discontinuity. Although the local minimum is not exactly at 59.1984, it is close: 58.0 in the case of the NELS, 60.4 in the case of 1968 head start funding, and 59.4 in the case of 1972 head start funding. For the NELS, this local minimum is also a global minimum over the range we examined. For 1968 and 1972 head start funding, there are also other minima in the range of 34 percent poor. However we are not concerned that these represent a discontinuity related to the Head Start program, as this roughly corresponds to the 56th percentile of county poverty rates. In addition, these minima are located in a region with much greater density of data, where the degrees-of-freedom differences (due to greater density) can confound the goodness of fit measure. Finally, we note the only region that experiences an improved goodness of fit across all three first stage measures is the region around the 300th poorest county. On the whole, then, we take this evidence as confirmatory that there is a discontinuity happening at our threshold.

D. Discontinuities in Other Spending

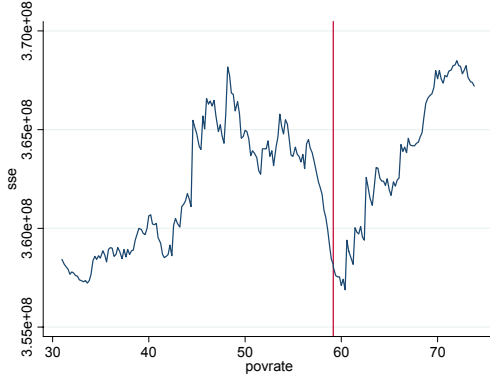
A final potential challenge to our identification of Head Start impacts with our research design comes from the possibility of discontinuities in other forms of social spending at the OEO cutoff. This seems unlikely since the decision to focus Head Start grant-writing assistance in the 300 poorest counties seems to have been made arbitrarily within the Head Start office rather than as some part of larger OEO-

Figure 8
Endogenous cut-point results

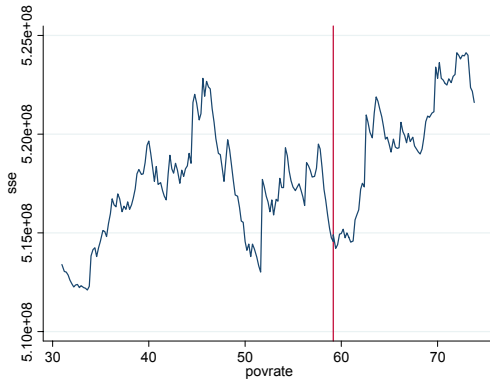
a. NELS parent reported Head Start participation



b. 1968 NARA Head Start funding per 4-year-old



c. 1972 NARA Head Start funding per 4-year-old



Note: Graphs show loss function, derived from estimating regression discontinuity models at varying potential discontinuity points. See Section VII.C of the text for details.

wide policy. And in fact Figure 9 shows that the discontinuity in other forms of federal social spending³⁰ at the OEO cutoff is very small expressed as a share of the control mean (<2 percent, vs. 50–100 percent for Head Start), in absolute dollars (~\$10 per capita, vs. \$100 or more per 4-year-old with Head Start), or compared to the standard error ($t=0.16$).³¹

A related concern is the possibility that because Head Start was part of OEO's Community Action Program, differences across counties in Head Start funding could generate differences in political mobilization that could in turn affect local and state spending. Community mobilization could also directly affect children's schooling if "social capital" matters for long-term outcomes. To address this concern we analyze data on the ratio of registered voters to age-eligible (18+) county residents biannually from 1968 to 1980 (see Appendix A), and do not find any statistically significant discontinuities in this measure in any year (Table 10).

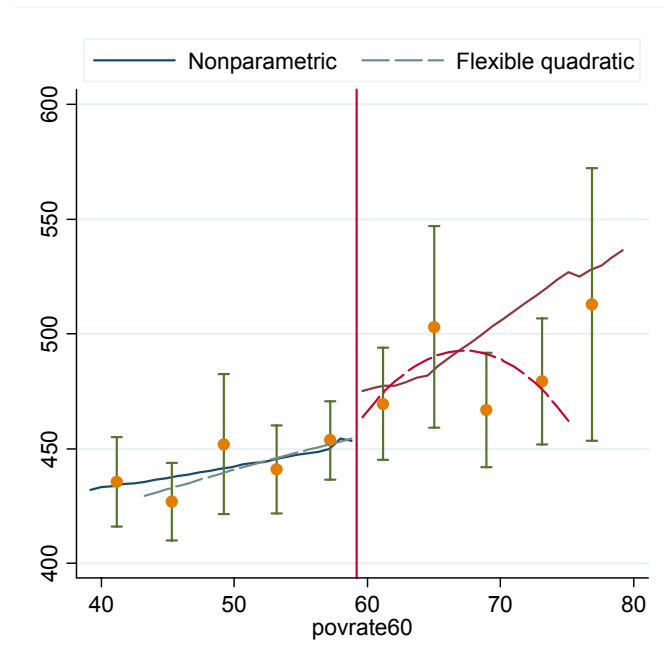
E. Additional Estimation Issues

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 NELS Head Start participation, for the mortality results, and for the educational outcomes for the directly treated cohort in the Census 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). The results are weaker than those shown above for the NARA funding results and for the NELS educational outcomes. However, the weighted estimates show somewhat more pronounced discontinuities for the population measures in

³⁰This category includes all appropriations by HEW, HUD, the US Department of Labor, and OEO, plus some selected programs run by the Department of Agriculture such as low-income housing programs and school lunches.

³¹Focusing just on Federal education expenditures per child, including from the Title 1 program, yields a statistically insignificant discontinuity estimate that is equal to around 5 percent of the left limit, while the estimated discontinuity in spending specific to child and maternal health or nutrition programs are also proportionally small and statistically insignificant.

Figure 9
Discontinuity in Other Federal Social Spending, 1972 National Archives Data



Note: Estimated nonparametric discontinuity = 20.00 T-stat = 1.04, bandwidth = 8.

Table 10
Regression Discontinuity Estimates, Voter Registration Rates,
 Voter Registration Data from ICPSR, 1968–1980

Variable	LHS Mean	Porter				Parametric		
		4	6	8	10	16	Flexible Linear	Flexible Quadratic
Bandwidth							8	16
Poverty Range								
Number of observations (counties) with nonzero weight		526	738	962	1198	2164	486	867
Voter registration rate, 1968	0.797	-0.051 (0.032)	-0.037 (0.028)	-0.027 (0.024)	-0.04* (0.023)	-0.028 (0.018)	-0.051 (0.032)	-0.045 (0.035)
Voter registration rate, 1970	0.763	-0.031 (0.029)	-0.016 (0.025)	-0.015 (0.022)	-0.027 (0.021)	-0.008 (0.016)	-0.033 (0.030)	-0.02 (0.032)
Voter registration rate, 1972	0.785	-0.004 (0.025)	0.001 (0.022)	0.004 (0.019)	-0.01 (0.018)	0.011 (0.015)	-0.004 (0.025)	0.001 (0.027)
Voter registration rate, 1974	0.759	-0.045* (0.025)	-0.028 (0.022)	-0.011 (0.019)	-0.016 (0.018)	0.016 (0.014)	-0.059** (0.025)	-0.047 (0.029)
Voter registration rate, 1976	0.762	-0.017 (0.025)	0.002 (0.022)	0.004 (0.019)	0.004 (0.017)	0.021 (0.014)	-0.018 (0.026)	-0.006 (0.030)
Voter registration rate, 1978	0.726	-0.005 (0.025)	0.001 (0.022)	-0.002 (0.019)	-0.005 (0.017)	0.013 (0.013)	-0.005 (0.025)	0.003 (0.029)
Voter registration rate, 1980	0.766	-0.014 (0.023)	0.002 (0.020)	-0.018 (0.018)	-0.021 (0.016)	-0.002 (0.013)	-0.006 (0.023)	0.005 (0.026)

Notes: Voter registration rate is defined as the ratio of registered voters to population 18 and older in a county. States for which county-level voter registration data are available increases over time (see Appendix A).

Table 7, and for educational outcomes for the directly treated group at the pseudo-cutoff used in Table 9. To the extent to which this serves as a diagnostic test, this finding provides further empirical justification for preferring the un-weighted to the weighted estimates.

Finally, conditioning our parametric estimates on exogenous covariates like percent of the county that is black or urban does not do much to our results, consistent with findings above that these covariates are balanced at the OEO cutoff.

IX. CONCLUSIONS

One contribution of our paper is to highlight a new source of identifying information for the long-term effects of Head Start, generated by the discontinuity in program funding across counties by virtue of the grant-writing assistance given in 1965 to just the poorest 300 counties. We also demonstrate that the discontinuity in Head Start funding at the OEO cutoff is mirrored by a large discontinuity in child mortality rates for causes of death plausibly affected by Head Start. The impacts are proportionately large among both blacks and whites, although imprecisely estimated for blacks. There are no significant discontinuities in either other causes of death or age groups that should not be affected by Head Start.

Our evidence for positive Head Start impacts on educational attainment is suggestive, although neither of the data sources available to us is quite ideal. We show that there are statistically significant positive discontinuities in high school completion and college attendance in the 1990 Census, concentrated among cohorts exposed to Head Start. We find similar results for schooling outcomes in the NELS, which first identifies county of residence for people at age 13, and other data suggest there is a great deal of overlap between those living in the poorest counties at age 3 (Head Start age) versus 13. The qualitative similarity in findings across data sources strengthens the findings from each. Importantly, unlike the sibling-difference estimates from GTC, we find suggestive evidence for schooling impacts on blacks as well as whites.

A final potential concern with our results is the possibility that the grant-writing assistance provided by OEO could have generated a difference in the Head Start “production function” across

counties. While we cannot definitively rule this out, this seems unlikely in part because the PMIs sent out in Spring 1965 were not in the field long enough to do much training. Moreover once Head Start grants were awarded that spring, OEO conducted a massive national training effort that would be expected to help standardize production technologies across areas.

What do our findings imply for public policy? One aspect of this question has to do with generalizability of our estimated treatment effects. The gain from Head Start will be positively related to the difference between the quality of the program's educational and health services versus the counterfactual family or child-care services children experience otherwise, and perhaps as well to the quality of the post-Head Start schooling experience. If our study population of poor children in the South in the 1960s and 1970s is more disadvantaged than the average Head Start child today, then our estimates may provide an upper bound for the effects of expanding Head Start funding for more recent national cohorts. Working in the opposite direction may be improvements over time in average school quality for poor children.

With this caveat in mind, some rough calculations suggest that Head Start passes a benefit-cost test at least for the cohorts of children for whom we estimate treatment effects. We assume the "treatment" we study here is an extra \$400 in spending on Head Start per 4-year-old child in 2003 dollars (discontinuity in funding we estimate using 1972 data). On the benefit side our estimates suggest mortality improvements yield benefits of around \$180 per child.³² If our NELS estimate is correct that the funding discontinuity in Head Start increases schooling by about 0.5 years per child then we expect the present value of the associated earnings gains to be around \$9,000,³³ although our study population is still

³²Our Vital Statistics estimates for children 5–9 imply an effect on overall mortality rates of about 3 per 100,000. Using a value per statistical life of \$6 million, which is about the value used by the EPA and FDA (Sunstein, 2004), these mortality results would imply a modest benefit of around \$180 per 4-year-old in a county.

³³If we use a conservative estimate for the returns to an additional year of schooling of 5 to 10 percent (Card, 1999), and if Head Start does in fact increase educational attainment by around .5 years, then the benefits from increased earnings (even assuming a relatively high discount rate of 6 percent and no annual productivity

too young (mid-20s) in our data to measure earnings gains directly. For the poorest children in the 1960s–70s, Head Start seems to yield net benefits even if there are distortions from raising tax revenues to support the program.

Over the past several years there has been growing concern in policy circles about the ability of Head Start in its original incarnation to achieve lasting impacts on children, as summarized by one recent policy brief: “When the emotional appeals [about Head Start] are cleared from the table, what is left is a costly but unsuccessful experiment. . . . Head Start is not working. Accept that, and let it go.”³⁴ Our findings together with the sibling-difference estimates from CT and GTC paint a less pessimistic picture about the ability of Head Start to achieve lasting benefits for poor children. Importantly, this new research suggests that the non-educational components of Head Start may contribute to the program’s benefits, including the mortality impacts presented in our paper, and that lasting impacts on educational attainment may arise for blacks as well as whites. Many of the recent assessments of Head Start’s value would appear to be overly pessimistic.

growth over time) would be on the order of nearly \$9,000 per 4-year-old (results from Krueger, 2003b, Table 5, multiplied by around 3).

³⁴“Don’t Cry for Me, Head Start,” Darcy Olsen and Eric Olsen, April 15, 1999, Daily Commentary, Cato Institute. Concerns about the limited evidence in support of lasting Head Start effects are common in Washington, DC policy discussions; see, for example, Haskins (2004) or Douglas Besharov’s testimony to the U.S. Senate Committee on Health, Education, Labor and Pensions, Subcommittee on Children and Families, April 11, 2000.

Appendix A: Data Sources

In what follows we discuss the various county-level data sources that we use to estimate the long-term impacts of Head Start, and then discuss our student-level micro-data sources.

A. County-Level Data

Perhaps the most important data question for our study is to understand what data source OEO used to identify the 300 poorest counties in the U.S. in 1960. The answer is not immediately obvious because the official federal poverty rate was only invented in 1964, and so the 1960 Census does not include a measure of persons living in poverty. OEO apparently identified the poorest counties using a special 1964 re-analysis of the 1960 Census conducted by the Census Bureau for OEO using the then-newly-defined federal poverty rate, a copy of which we have obtained from the National Archives and Records Administration (NARA).³⁵ The alternative possibility is that OEO targeted grant-writing assistance using data on the proportion of families in the county with incomes below \$3,000, although our analysis reveals that a larger discontinuity in funding at the 300th poorest county using the official poverty rate, suggesting OEO used that measure instead.³⁶

In order to document the discontinuity in Head Start funding around the OEO cutoff, and to examine whether there is a similar discontinuity in other forms of federal spending, we have also obtained from NARA a series of OEO data files on federal expenditures per county for the years 1967 through

³⁵NARA, Records of the Community Services Administration, Record Group 381: Putnam Print File, 1960.

³⁶Since 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.

1980.³⁷ The accuracy of these data is less than perfect. The electronic data files have some obvious errors and are poorly documented.³⁸ In the end only spending data from 1968 and 1972 were usable, in the sense that the data from the electronic files matched published figures for total federal spending and Head Start spending at the national level, and the data matched for Head Start at the state level as well. Another problem is that some federal spending is passed through state governments. In these cases OEO pro-rated state spending across counties, which might be reasonable on average but lead to error in measuring spending in the poorest areas. A related challenge specific to the 1968 spending data is that the Child Development Group of Mississippi (CDGM) operated Head Start programs in a number of counties in Mississippi, yet CDGM's total funding is concentrated in the Mississippi county (Clay) that housed CDGM's headquarters.³⁹ In any case, in addition to constructing a measure of Head Start spending we create variables for spending on other social programs.⁴⁰

³⁷Federal 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).

³⁸The 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.

³⁹We use information from the appendix map in Greenberg (1969) to identify other Mississippi counties in which CDGM operated and interpolate spending across these counties, assuming that CDGM allocated their Head Start spending across counties in proportion to county population. Results setting CDGM spending to missing are qualitatively similar. The CDGM problem is not an issue for the 1972 data because the agency lost OEO funding in the late 1960s.

⁴⁰For 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 by HEW, OEO (excluding those made through Head Start), the Department of Labor, HUD, and selected programs from the Department of Agriculture.

Our primary data source on long-term schooling outcomes comes from a special tabulation conducted for us by the U.S. Census Bureau for data from the 1990 Census. These data include information on educational attainment, employment, earnings, and residence 5 years ago disaggregated by age group, gender, and race. We use data from the 1990 rather than 2000 Census because the funding disparity across counties seems to have smoothed out over time, which means that the discontinuity in funding will be less pronounced for young people in 2000 compared to 1990. At the same time the additional 10 years worth of mobility for older cohorts from 1990 to 2000 will further attenuate our estimates for these age groups.

We also draw on county-level data from the Vital Statistics, which provides information from a census of all death certificates in the U.S., including detailed cause-of-death codes recorded during our observation period (1973–83) using the International Classification of Diseases, 8th or 9th edition (ICD-8 and ICD-9) systems. Using data from the Compressed Mortality Files we focus on mortality to children in the age group 5–9, which should capture the health effects of Head Start services to children ages 3–4. We identify the causes-of-death that may plausibly be affected by Head Start by consulting descriptions of the health services offered by the program (for example North, 1979) together with discussions with pediatricians about what health problems those services should influence. The detailed list of ICD-9 codes and cause-of-death descriptions that we count as plausibly affected by Head Start is given in Appendix Table A1.

Finally, we draw on county-level voter registration records collected from state officials (ICPSR Study 9405) for the period 1968–88. We focus on the ratio of registered voters to population 18 and over for the years 1968 through 1980 to capture possible effects of Head Start or related OEO activities on voter registration rates. (The population data for the denominators of these ratios come from the CMF discussed above). Note that voter data are not available for all states although this becomes less of a

problem over time,⁴¹ and voter registration totals may be subject to some upward bias from voters who have left the county or state but are not immediately removed from the voting rolls (see for example Piven and Coward, 1989). Note also that trends in voter registration rates over time will diverge from trends in voting rates because of variation over time in the propensity of registered voters to vote.

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 around 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.

The disadvantage of the NELS is that the study 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. 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

⁴¹The ICPSR dataset is missing voter information for Alabama for 1968–72, Iowa for 1968–74, Kansas for 1968–70, Minnesota for 1968–72, Texas for 1968–74, and Virginia for 1968–72. Note that by 1976 our dataset will include complete information for all of the Southern states that account for most of the 600 with the poorest 1960 county poverty rates, as mentioned above. The number of counties for which we have valid voter registration data in our sample increases over time as follows: 1968, N=1,991; 1970, N=1,994; 1972, N=2,111; 1974, N=2,135; 1976, N=2,732; 1978, N=2,879; and 1980, N=2,845.

Asian Islander students (U.S. Department of Education, 1994). 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).

Overall the base year sample includes students drawn from 568 different counties. The base year sample includes 649 students who lived in counties with 1960 poverty rates among the 300 poorest, and 674 respondents who lived in one of the next 300 poorest counties with respect to 1960 poverty. The NELS base year sample sizes for the 100 counties with poverty rates just above the cutoff is N=183; for 200 counties above, N=493; for 100 counties with poverty rates just below the cutoff, N=185; and for the 200 counties just below the cutoff, N=361.

Another potential concern with the NELS is that Head Start effects on grade retention could affect the age distribution across counties of who is enrolled in 8th grade in 1988, although in practice we do not find a discontinuity in age at the OEO cutoff.

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.⁴² The other key explanatory variable for

⁴²This 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 (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

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 1988.⁴³

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 1988.⁴⁴

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 the state (much less county) level, 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

government: If we assume each Head Start participant 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.

⁴³For 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 percent of base-year NELS respondents.

⁴⁴We 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).

larger PSID sample in all poor counties to explore the problem of selective across-county migration with our NELS and county-level data.⁴⁵

⁴⁵An 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.

Appendix Table A1
Causes of Death Susceptible to Head Start Interventions

ICD-9 Codes	Cause of Death
10-18	Tuberculosis
33	Whooping cough
34-35	Streptococcal sore throat, scarlatina, erysipelas
36	Meningococcal infection
38	Septicemia
45	Acute poliomyelitis
55	Measles
70	Viral hepatitis
250	Diabetes mellitus
260-269	Nutritional deficiencies
280-285	Anemias
320-322	Meningitis
466	Acute bronchitis, bronchiolitis
480-487	Pneumonia and influenza
490-491	Bronchitis, chronic and unspecified
492	Emphysema
493	Asthma
494-496	Other chronic obstructive pulmonary diseases and allied conditions

Appendix B:
“Optimal” Bandwidth Selection for Semi-Parametric Regression Discontinuity Estimates

In our non-parametric regression discontinuity estimates, we have primarily chosen the range of bandwidths for which to present results based on visual inspection of the data. Our goal has been to look as “locally” as possible, while still using enough data to obtain reasonable statistical power. We have also explored a leave-one-out cross validation selection procedure that we describe in this appendix, although in the end we reject the mechanical approach to bandwidth selection that comes from this procedure for the reasons discussed below.

As is well known, cross-validation (henceforth referred to as “C-V”) is one method that in principle can be used to select “tuning parameters” such as bandwidths. However, as is also known, the asymptotic properties of C-V methods imply extremely slow rates of convergence. This problem can represent itself in “noisy” measures of the optimal tuning parameter (Hastie *et. al.*, 2001). As we will describe below, the C-V estimates of the loss function are typically rather flat, which we take to be an indication of this problem.

In addition, there is further complication due to the fact that we are trying to select the optimal bandwidth for a non-parametric regression discontinuity estimator. To the best of our knowledge existing literature does not provide specific guidance for this particular problem. Applying the optimal bandwidth from a “typical” C-V exercise would seem to provide misleading results, as the typical case involves estimating a function at the interior of the support of the data, while our regression discontinuity involves two estimations at a boundary. Because the bias/variance tradeoffs are likely to be different for boundary estimation than for interior estimation, we should use a C-V technique that focuses on boundary predictions. Finally, we are interested in the local nature of the data, and in identifying the optimal bandwidth for the local density, conditional variance, etc. With these concerns in mind, we performed a C-V exercise as described below.

Given a potential bandwidth, we form a loss function by looking at prediction errors for each data point within 5 percentage points of 1960 county poverty from the 59.1984 cutoff. The predictions for each

point are made as follows: for points to the left of (less poor than) the cutoff, we predict their values using a locally weighted regression using only data to their left (exclusive of the data point being predicted). That is, we predict their values as if they were at the boundary of the estimation. For points to the right of the cutoff, we predict their values using only data to their right. As a result, our loss function is an average of boundary prediction errors, and is local to the cutoff of interest. We create this loss for a range of potential bandwidths from 1 to 16, and then examine the results. We do this for several different left hand side variables of interest.

Appendix Table B presents the results from this exercise for several key variables presented in the paper's main tables. Several patterns emerge. First, the bandwidths selected by our C-V procedure are typically toward the upper end of the range of bandwidths for which results are presented in the body of the paper. Second, the estimated loss function from the C-V exercise is very "flat." To see this, we show the range of bandwidths (we only consider bandwidths in the range of 1 to 16) whose loss function is no more than 5 percent greater than the "optimal." This set is typically very wide. The third point that emerges is that for most variables, the estimated treatment impact from using the C-V bandwidth is similar to that from using our preferred bandwidth.

References

- Almond, Douglas V. 2005. "Is the 1918 Influenza Pandemic Over? Long-Term Effects of In Utero Influenza Exposure in the Post-1940 U.S. Population." Working Paper, Department of Economics, Columbia University.
- Almond, Douglas V., Kenneth Y. Chay, and Michael Greenstone. 2003. "Civil Rights, the War on Poverty, and Black-White Convergence in Infant Mortality in Mississippi." Working Paper, Department of Economics, University of California at Berkeley.
- Bai, Jushan, and Pierre Perron. 1998. "Estimating and Testing Linear Models with Multiple Structural Changes." *Econometrica* 66(1): 47–78.
- Barnett, W. Steven. 1995. "Long-Term Effects of Early Childhood Programs on Cognitive and School Outcomes." *The Future of Children* 5(3): 25–50.
- Boozer, Michael A., Alan B. Krueger, and Shari Wolkin. 1992. "Race and School Quality Since Brown vs. Board of Education." In *Brookings Papers on Economic Activity: Microeconomics*, edited by Martin Bailey and Clifford Winston. Washington, DC: Brookings.
- 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 *Handbook of Labor Economics, Volume 3A*, edited by Orley Ashenfelter and David Card. Amsterdam: Elsevier. pp. 1801–1864.
- Carneiro, Pedro, and James J. Heckman. 2003. "Human Capital Policy." In *Inequality in America: What Role for Human Capital Policies?* edited by James J. Heckman and Alan B. Krueger. Cambridge, MA: MIT Press. pp. 77–240.
- Case, Anne, Darren Lubotsky, and Christina Paxson. 2002. "Economic Status and Health in Childhood: The Origins of the Gradient." *American Economic Review* 92(5): 1308–1334.
- Case, Anne, Angela Fertig, and Christina Paxson. 2003. "From Cradle to Grave? The Lasting Impact of Childhood Health and Circumstance." NBER Working Paper 9788.
- Citro, Constance F., and Robert T. Michael. 1995. *Measuring Poverty: A New Approach*. Washington, DC: National Academy Press.
- Coleman, James S. et al. 1996. *Equality of Educational Opportunity*. Washington, DC: U.S. Department of Health, Education, and Welfare.
- Collins, William, and Robert A. Margo. 2004. "The Economic Aftermath of the 1960s Riots: Evidence from Property Values." NBER Working Paper 10493.
- Currie, Janet. 2001. "Early Childhood Education Programs." *Journal of Economic Perspectives* 15(2): 213–238.
- Currie, Janet, and Joseph V. Hotz. 2004. "Accidents Will Happen? Unintentional Injury, Maternal Employment and Child Care Policy." *Journal of Health Economics* 23(1): 25–59.

- 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.
- Fan, Jianqing. 1992. "Design-Adaptive Nonparametric Regression." *Journal of the American Statistical Association* 87(420): 998–1004.
- Fryer, Roland G., and Steven D. Levitt. 2004. "Understanding the Black-White Test Score Gap in the First Two Years of School." *Review of Economics and Statistics* 136(2): 447–464.
- 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.
- Goldin, Claudia, and Lawrence F. Katz. 2000. "Education and Income in the Early Twentieth Century: Evidence from the Prairies." *Journal of Economic History* 60: 782–818.
- Greenberg, Polly. 1969. *The Devil Has Slippery Shoes*. New York: Macmillan.
- Grogger, Jeffrey, and Derek Neal. 2000. "Further Evidence on the Effects of Catholic Secondary Schooling." In *Brookings-Wharton Papers on Urban Affairs*, edited by William Gale and Janet Rothenberg Pack. Washington, DC: Brookings Institution Press. pp. 151–202.
- Guryan, Jonathan. 2001. "Does Money Matter? Regression-Discontinuity Estimates from Education Finance Reform in Massachusetts." National Bureau of Economic Research Working Paper 8269.
- Guryan, Jonathan. 2004. "Desegregation and Black Dropout Rates." *American Economic Review* 94(4): 919–943.
- 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.
- Haskins, Ron. 2004. "Competing Visions." *Education Next* 4(1): 26–33.
- Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2001. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer.

- Jencks, Christopher, and Meredith Phillips. 1998. "The Black-White Test Score Gap: An Introduction." In *The Black-White Test Score Gap*, edited by Christopher Jencks and Meredith Phillips. Washington, DC: Brookings Institution Press. pp. 1–54.
- Jones, Jean Yavis. 1979. *The Head Start Program—History, Legislation, Issues and Funding, 1964–1978*. Washington, DC: Congressional Research Service Report 79-14 EPW.
- Karoly, Lynn A. et al. 1998. *Investing in Our Children: What We Do and Don't Know About the Costs and Benefits of Early Childhood Interventions*. Santa Monica, CA: RAND.
- Kling, Jeffrey R., Jens Ludwig, and Lawrence F. Katz. 2005. "Neighborhood Effects on Crime for Male and Female Youth: Evidence from a Randomized Housing Voucher Experiment." *Quarterly Journal of Economics* 120(1): 87–130.
- Krueger, Alan B. 2003a. "Inequality, Too Much of a Good Thing." In *Inequality in America: What Role for Human Capital Policies?* edited by James J. Heckman and Alan B. Krueger. Cambridge, MA: MIT Press.
- Krueger, Alan B. 2003b. "Economic Considerations and Class Size." *Economic Journal* 113: 34–63.
- Ladd, Helen F., and Janet S. Hansen, Eds. 1999. *Making Money Matter: Financing America's Schools*. Washington, DC: National Academies Press.
- 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.
- Magnuson, Katherine A., and Jane Waldfogel. 2005. "Early Childhood Care and Education: Effects on Ethnic and Racial Gaps in School Readiness." *The Future of Children* 15(1): 169–196.
- McCrary, Justin, and Heather Royer. 2003. "Does Maternal Education Affect Infant Health? A Regression Discontinuity Approach Based on School Age Entry Laws." Working Paper, University of Michigan.
- 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.
- Neal, Derek. 2005. "Why Has Black-White Skill Convergence Stopped?" Cambridge, MA: NBER Working Paper 11090.
- Nelson, Charles A. 2000. "Neural Plasticity and Human Development: The Role of Early Experience in Sculpting Memory Systems." *Developmental Science* 3(2): 115–136.
- North, A. Frederick. 1979. "Health Services in Head Start." In *Project Head Start: A Legacy of the War on Poverty*, edited by Edward Zigler and Jeanette Valentine. NY: Free Press. pp. 231–257.
- O'Brien, Robert, David B. Connell, and James Griffin. 2004. "Head Start's Efforts to Improve Child Health." In *The Head Start Debates*, edited by Edward Zigler and Sally J. Styfco. Baltimore: Paul H. Brooks. pp. 161–178.

- 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.
- Piehl, Anne Morrison, Suzanne J. Cooper, Anthony A. Braga, and David M. Kennedy. 2003. "Testing for Structural Breaks in the Evaluation of Programs." *The Review of Economics and Statistics* 85(3): 550–558.
- Piven, Frances Fox, and Richard Cloward. 1989. "Government Statistics and Conflicting Explanations of Nonvoting." *PS: Political Science and Politics* 22(3): 580–588.
- Porter, Jack. 2003. "Estimation in the Regression Discontinuity Model." Working Paper, Harvard University Department of Economics, draft date September 25, 2003.
- Ramey, Craig T., and Frances A. Campbell. 1984. "Preventive Education for High-Risk Children: Cognitive Consequences of the Carolina Abecedarian Project." *American Journal of Mental Deficiency* 88(5): 515–523.
- 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.
- Rock, Donald A., and A. Jackson Stenner. 2005. "Assessment Issues in the Testing of Children at School Entry." *The Future of Children* 15(1): 15–34.
- Schweinhart, Lawrence J., Jeanne Montie, Zongping Xiang, W. Steven Barnett, Clive R. Belfield, and Milagros Nores. 2005. *Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40*. Ypsilanti, Michigan: High/Scope Press.
- Shonkoff, J. P., and D. A. Phillips. 2000. *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Washington, DC: National Academy Press.
- Singh, Gopal K., and Stella M. Yu. 1996. "U.S. Childhood Mortality, 1950 through 1993: Trends and Socioeconomic Differentials." *American Journal of Public Health* 86(4): 505–512.
- Sundquist, James L. 1969. *On Fighting Poverty: Perspectives from Experience*. New York: Basic Books.
- Sunstein, Cass R. 2004. "Are Poor People Worth Less Than Rich People? Disaggregating the Value of Statistical Lives." University of Chicago Law School, John M. Olin Law and Economics Working Paper 207.
- Tremblay, Richard E., Daniel S. Nagin, Jean R. Seguin, Mark Zoccolillo, Philip D. Zelazo, Michael Boivin, Daniel Perusse, and Christa Japel. 2004. "Physical Aggression During Early Childhood: Trajectories and Predictors." *Pediatrics* 114(1): e43–e50.
- Trochim, W. 1984. *Research Design for Program Evaluation: The Regression Discontinuity Approach*. Beverly Hills, CA: Sage Publications.

- U.S. Census Bureau. 1999. *Statistical Abstract of the United States, 119th Ed.* Washington, DC.
- U.S. Department of Education. 1994. *National Education Longitudinal Study of 1988 Second Follow-Up: Student Component Data File User's Manual (NCES 94-374)*. Washington, DC: Government Printing Office.
- U.S. Department of Health, Education, and Welfare. 1960. *Vital Statistics of the United States, 1960. Volume 1: Natality*. Washington, DC: United States Department of Health, Education, and Welfare.
- U.S. Department of Health and Human Services, Administration for Children and Families. 2005. *Head Start Impact Study: First Year Findings*. Washington, DC.
- U.S. Statistical Abstract. 1999. *Statistical Abstract of the United States, 119th Ed.* Washington, DC: U.S. Census Bureau.
- 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.
- Zigler, Edward, and Jeanette Valentine. 1979. *Project Head Start: A Legacy of the War on Poverty*. New York: Free Press.
- Zigler, Edward, and Sally J. Styfco. 2004. "Preface." *The Head Start Debates*. Baltimore: Paul H. Brookes.