

# The Long-Run Effects of Head Start: Replicating Ludwig-Miller (2007) and a Framework for Studying Intergenerational Mobility

APEP Research Team\*

January 2026

## Abstract

This paper replicates the influential Ludwig and Miller (2007) regression discontinuity analysis of Head Start and develops a framework for extending it to study intergenerational mobility. We exploit the discontinuity created by the Office of Economic Opportunity’s (OEO) provision of grant-writing assistance to the 300 poorest U.S. counties in 1965. Our replication yields estimates consistent with the original finding: counties above the 59.2% poverty threshold show approximately 1.2–1.8 fewer deaths per 100,000 children ages 5–9 from Head Start-related causes, though our preferred specification (18pp bandwidth, local linear) yields a marginally significant estimate ( $\tau = -1.20$ ,  $SE = 0.66$ ,  $p = 0.07$ ; 95% CI:  $[-2.49, 0.10]$ ). Validity tests support the design’s credibility: the McCrary density test shows no evidence of manipulation (log discontinuity =  $-0.002$ ,  $SE = 0.15$ ), and placebo tests using pre-program mortality show no

---

\*This paper was generated as part of the Autonomous Policy Evaluation Project (APEP). All code and data are available in the replication package. nd @dakoyana

spurious effects ( $\tau = -0.64$ ,  $p = 0.72$ ). We discuss how this design could be extended to study intergenerational mobility using Opportunity Insights data, identifying data linkage challenges that must be resolved. The paper contributes a replication of a canonical RDD study and a roadmap for future research on the long-run effects of early childhood interventions.

**JEL Codes:** I38, J13, I24, R10

**Keywords:** Head Start, Intergenerational Mobility, Regression Discontinuity, Early Childhood Education

# 1 Introduction

The relationship between early childhood interventions and long-run economic outcomes represents one of the most important questions in economics and public policy. A large body of research has documented that investments in human capital during the early years of life yield substantial returns, both for individuals and for society (??). Early childhood programs, in particular, have been shown to improve educational attainment, health outcomes, and labor market success, with effects that sometimes persist for decades after program participation (?). Understanding whether these programs can break cycles of intergenerational poverty and improve economic mobility is critical for designing effective anti-poverty policies.

Head Start, established in 1965 as part of President Johnson’s War on Poverty, remains the largest federally funded early childhood program in the United States. The program serves nearly one million children annually at a cost of approximately \$12 billion, providing comprehensive early childhood education, health, and nutrition services to low-income children ages 3–5 and their families (?). Despite decades of research, fundamental questions remain about whether Head Start’s well-documented short-term benefits translate into lasting improvements in economic mobility across generations.

This paper contributes to this literature by examining whether Head Start affected intergenerational economic mobility—the extent to which children’s adult economic outcomes differ from those of their parents. Intergenerational mobility has emerged as a central concern in economics and public policy, following influential work by ? documenting substantial geographic variation in the United States. If early childhood programs can improve mobility, they represent a potentially powerful tool for addressing persistent inequality and promoting equal opportunity.

We build on the seminal work of ?, who exploited a unique source of variation in Head Start funding. In 1965, the Office of Economic Opportunity (OEO) provided technical assistance for grant-writing to the 300 poorest counties in the United States, as measured by 1960 Census poverty rates. This assistance created a large and persistent discontinuity

in Head Start funding and participation rates at the poverty threshold, providing a natural experiment that allows researchers to identify causal effects of the program.

Our analysis proceeds in two stages. First, we replicate the Ludwig-Miller regression discontinuity design and find results consistent with their main finding: counties just above the poverty cutoff exhibit lower child mortality from causes amenable to Head Start’s health services, though statistical significance depends on specification. This replication serves both to assess the identification strategy and to provide a foundation for extending the analysis to additional outcomes.

Second, we develop a framework for linking this historical treatment variation to modern measures of intergenerational mobility from the Opportunity Insights project (?). The Opportunity Atlas provides unprecedented data on the long-run outcomes of children by their county of birth and parental income, enabling analysis of whether place-based policies affected the economic trajectories of subsequent generations.

Our main findings are as follows. We find that the OEO grant-writing assistance discontinuity is associated with approximately 1.2–1.8 fewer deaths per 100,000 children ages 5–9 from Head Start-related causes, with our preferred estimate marginally significant at conventional levels ( $p \approx 0.07$ ). The McCrary density test (?) shows no evidence of manipulation at the poverty threshold, and placebo tests using pre-treatment mortality show no spurious effects. These validity tests support the regression discontinuity design for studying Head Start’s effects, though the statistical uncertainty warrants cautious interpretation.

The remainder of this paper is organized as follows. Section 2 provides background on Head Start and reviews the relevant literature on early childhood interventions and intergenerational mobility. Section 3 describes our data sources in detail. Section 4 presents the empirical strategy, including the regression discontinuity design and validity tests. Section 5 reports results from our replication analysis. Section 6 discusses the framework for extending the analysis to intergenerational mobility. Section 7 concludes with policy implications and directions for future research.

## 2 Background and Literature

### 2.1 The Head Start Program

Head Start was launched in the summer of 1965 as an eight-week demonstration program that served over 560,000 children in its first year (?). The program was designed as a comprehensive early childhood intervention, recognizing that cognitive development alone is insufficient to address the multiple disadvantages faced by children in poverty. The program's founders understood that health, nutrition, and family circumstances all affect children's readiness to learn and their long-term life outcomes.

The program provides comprehensive services spanning four domains. First, early childhood education services focus on cognitive and socio-emotional development, preparing children for kindergarten through structured learning activities, play-based curriculum, and individualized instruction. Teachers work to develop children's language skills, early literacy, and numeracy, while also fostering social skills and emotional regulation (?).

Second, health services include screenings and referrals for medical and dental care, ensuring that children receive necessary immunizations and treatment for health problems that might otherwise go unaddressed. Head Start programs connect families with health-care providers and help them navigate the often-complex healthcare system. These services were particularly important in the 1960s, when access to healthcare was far more limited for low-income families than it is today.

Third, nutrition services provide meals and nutrition education, addressing both immediate hunger and longer-term health outcomes. Children in Head Start receive breakfast, lunch, and snacks that meet federal nutrition standards, and families receive information about healthy eating practices. For many children, Head Start provides the most nutritious meals they receive each day.

Fourth, parent involvement and family support services recognize that children's outcomes depend heavily on their home environments. Head Start programs engage parents as

partners in their children’s education, offering training and resources to support learning at home. Family services coordinators help families access community resources, from housing assistance to job training programs (?).

By design, Head Start targets children from disadvantaged backgrounds who might otherwise lack access to early childhood education and health services. Eligibility is generally limited to families below the federal poverty line, though programs have some flexibility to serve families up to 130% of poverty. The program prioritizes enrollment of children with disabilities and children experiencing homelessness, ensuring that the most vulnerable populations have access to services.

## **2.2 The OEO Grant-Writing Assistance Discontinuity**

The Office of Economic Opportunity, created by the Economic Opportunity Act of 1964, was charged with coordinating the War on Poverty programs including Head Start, Job Corps, and Community Action Programs. Recognizing that the poorest communities often lacked the administrative capacity to apply for federal grants, OEO sought to ensure that these communities could participate in new programs.

To encourage Head Start participation, OEO provided technical assistance for writing grant proposals to the 300 poorest counties in the United States, as identified using 1960 Census poverty rates (?). This assistance proved highly effective in overcoming barriers to program participation. Staff from OEO traveled to these counties to help local officials understand the program requirements, develop applications, and establish administrative structures for program implementation.

The result was a substantial discontinuity in Head Start funding and participation at the poverty threshold. Counties that received OEO assistance had Head Start funding rates 50–100% higher than comparable counties just below the threshold. Importantly, this discontinuity persisted for decades, as the initial funding advantage led to sustained differences in program infrastructure, local expertise, and community awareness of Head

Start.

Several features of this setting make it particularly suitable for causal inference using regression discontinuity methods (??). The poverty rate used to determine treatment was measured in 1960, five years before the policy was conceived, eliminating the possibility that counties manipulated their reported poverty rates to qualify for assistance. The threshold was determined by ranking counties nationally rather than by any local characteristic that might be correlated with other determinants of child outcomes. Counties could not anticipate the policy or take actions to position themselves relative to the threshold.

The persistence of the discontinuity is particularly important for studying long-run outcomes. ? document that the funding difference remained substantial through at least 1998, suggesting that cohorts born in the late 1970s and early 1980s—the focus of the Opportunity Insights data—would have been exposed to differential Head Start access. This persistence reflects the path-dependent nature of program implementation, as early investments in infrastructure and expertise create lasting advantages.

## 2.3 Prior Literature on Head Start Effects

The literature on Head Start’s effects has produced findings that vary considerably depending on research design, outcome measures, and time horizons. Randomized evaluations such as the Head Start Impact Study found modest short-term effects on cognitive skills that faded by the end of first grade (?). These findings generated considerable debate about the program’s effectiveness and raised questions about whether early childhood gains could persist over time.

However, observational and quasi-experimental studies have generally found more persistent effects. ? use sibling comparisons to control for family-level confounders and find that Head Start participation is associated with higher high school completion rates and, for Black participants, higher college attendance. The sibling design accounts for unobserved family characteristics that might bias simple comparisons of participants and non-participants.

? uses within-family variation in Head Start participation to estimate effects while controlling for shared family background. He finds that Head Start participants have higher test scores in adolescence, are less likely to repeat grades, and are less likely to commit crimes as young adults. Importantly, these effects are larger than those found in the randomized evaluation, suggesting that experimental compliance effects or sample selection may explain some of the discrepancy.

Using the OEO discontinuity, ? find that counties above the poverty threshold experienced 33–50% lower mortality rates for children ages 5–9 from causes amenable to Head Start’s health services. These causes include respiratory infections, meningitis, and other conditions that can be prevented or treated through the health screenings and referrals provided by Head Start. The mortality effects provide compelling evidence that the program generated real improvements in child health.

Subsequent work has extended the Ludwig-Miller analysis to additional outcomes. ? examine effects on educational attainment and crime, finding evidence of long-run benefits. Other researchers have used alternative identification strategies, including program rollout timing and funding variation, to study Head Start effects (??). The weight of evidence suggests that Head Start produces meaningful benefits, though the magnitude of effects varies across contexts and outcomes.

## 2.4 Intergenerational Mobility in the United States

The study of intergenerational mobility—the relationship between parents’ and children’s economic status—has a long history in economics and sociology. Classic work by ? and ? documented substantial persistence in economic status across generations, with estimates suggesting that approximately 40–60% of income differences persist from one generation to the next.

More recent work has exploited administrative data to provide more precise estimates and to examine geographic variation in mobility. ? use tax records linked across generations



to show that intergenerational mobility varies dramatically across commuting zones in the United States. Some areas exhibit high upward mobility, with children from low-income families frequently reaching the middle class or higher, while others exhibit low mobility, with poverty persisting across generations.

The Opportunity Atlas project (??) extends this analysis to the census tract level, providing fine-grained measures of children’s adult outcomes by their neighborhood of childhood residence. This data enables researchers to study how local factors—including schools, neighborhoods, and institutions—affect long-run economic outcomes. The data cover birth cohorts from 1978–1983, providing outcomes measured at ages 26–35.

Several local factors have been identified as correlates of mobility. Areas with less residential segregation, lower income inequality, better schools, more stable families, and greater social capital tend to exhibit higher mobility (??). Importantly, ?? provide experimental evidence that neighborhood quality matters: children whose families received housing vouchers to move to lower-poverty neighborhoods experienced improved long-run outcomes.

The relationship between early childhood programs and intergenerational mobility has received less direct attention. If programs like Head Start improve children’s human capital and health, they should increase mobility by helping children from disadvantaged backgrounds achieve better outcomes than they would have otherwise. However, demonstrating this causal relationship requires overcoming the selection problem: children who attend Head Start differ systematically from those who do not.

## 3 Data

### 3.1 Ludwig-Miller Head Start Data

Our primary data source is the replication dataset from ??, obtained from Bruce Hansen’s Econometrics textbook data repository. This dataset has been widely used in econometrics education and research for demonstrating regression discontinuity methods, ensuring that

our analysis is replicable and transparent.

The dataset contains 2,783 U.S. counties, representing nearly the complete set of counties excluding Alaska (where mortality data was unavailable). For each county, the data includes the 1960 Census poverty rate, which serves as our running variable, along with various outcome measures and covariates.

The key outcome variables measure mortality rates per 100,000 population for different age groups and causes of death. The primary outcome is mortality from Head Start-related causes among children ages 5–9, which includes deaths from respiratory infections, meningitis, and other conditions amenable to the health services provided by Head Start. The data also includes mortality from injuries (a placebo outcome that should be unaffected by Head Start) and pre-treatment mortality from the pre-Head Start period (another placebo to test for spurious discontinuities).

County characteristics from the 1960 Census include total population, urban share (the fraction of the population living in urban areas), Black population share, and school enrollment rates at various ages. These variables allow us to examine covariate balance across the discontinuity and to assess whether treated and control counties were comparable on observable dimensions.

The cutoff for OEO grant-writing assistance is 59.1984% poverty rate, representing the poverty rate of the 300th poorest county. Counties with poverty rates at or above this threshold were eligible for technical assistance, while those below were not. Table 1 presents summary statistics for counties within 18 percentage points of this threshold, the primary bandwidth used in our analysis.

Table 1: Summary Statistics: Counties Within 18pp of Cutoff

	Below Cutoff	Above Cutoff	Difference	<i>Notes:</i>
Poverty Rate 1960 (%)	49.9	67.4	17.5***	
HS-Related Mortality (per 100k)	4.12	2.98	-1.14*	
Injury Mortality (per 100k)	23.4	25.8	2.4	
Population (1960)	18,432	12,876	-5,556**	
Urban Share (%)	15.2	8.4	-6.8***	
Black Share (%)	25.1	38.6	13.5***	
N Counties	671	283		

*Sample restricted to counties with poverty rates between 41.2% and 77.2% (18pp bandwidth). Stars indicate significance of difference: \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .*

The summary statistics reveal that counties above and below the cutoff differ substantially on observable characteristics. Treated counties (above the cutoff) are smaller, more rural, and have higher Black population shares. These differences are expected given that the cutoff selects for the poorest counties, which tend to be concentrated in the rural South. The key assumption of the regression discontinuity design is that these differences vary smoothly through the cutoff, so that comparing counties just above and just below the threshold provides valid causal estimates.

### 3.2 Opportunity Insights County Data

We also obtained county-level intergenerational mobility data from Opportunity Insights. This dataset provides outcomes for children born 1978–1983 measured at ages 26–35, enabling analysis of long-run economic outcomes for cohorts that grew up during the period when the Head Start funding discontinuity was still in place.

The Opportunity Atlas covers 3,219 counties with valid observations on key outcomes. For each county, the data includes mean household income rank by parental income per-

centile, allowing researchers to measure mobility at different points in the income distribution. The data also includes college attendance rates, incarceration rates, teen birth rates, and measures of family structure such as the probability of having a father present at age 15.

A key feature of the Opportunity Insights data is that outcomes are reported separately by race and gender, enabling analysis of heterogeneous effects. Given the historical context of Head Start and its targeting of disadvantaged communities, effects may differ between Black and white children and between girls and boys. The data also reports standard errors for each estimate, allowing for proper statistical inference.

A significant challenge in our analysis is linking the Ludwig-Miller county identifiers to modern FIPS codes. The historical dataset uses county codes from the 1960 Census that do not directly correspond to current FIPS codes due to county boundary changes, mergers, and code reassignments over the intervening decades. Resolving this requires careful data harmonization using crosswalk files and historical records. We discuss this challenge and potential solutions in Section 6.

## 4 Empirical Strategy

### 4.1 Regression Discontinuity Design

We implement a sharp regression discontinuity design (RDD) exploiting the OEO poverty threshold. The RDD is among the most credible quasi-experimental methods available, providing estimates that approximate random assignment in the neighborhood of the cutoff (??). The key insight is that counties with poverty rates just above and just below the threshold are likely similar on all relevant dimensions except treatment status.

The estimating equation is:

$$Y_c = \alpha + \tau \cdot D_c + f(X_c - c) + D_c \cdot g(X_c - c) + \varepsilon_c \quad (1)$$

where  $Y_c$  is the outcome for county  $c$ ,  $D_c = \mathbf{1}[X_c \geq c]$  is an indicator for being above the cutoff,  $X_c$  is the county’s 1960 poverty rate,  $c = 59.1984$  is the cutoff, and  $f(\cdot)$  and  $g(\cdot)$  are polynomial functions of the running variable. The coefficient  $\tau$  identifies the causal effect of OEO grant-writing assistance on outcomes, under the assumption of continuity in potential outcomes at the threshold (?).

The regression discontinuity design requires several identifying assumptions. First, the conditional expectation of potential outcomes must be continuous at the cutoff. This assumption would be violated if some other factor changes discontinuously at the threshold and affects outcomes. Second, there must be no manipulation of the running variable, such that units cannot precisely control whether they are above or below the threshold (?).

In our setting, both assumptions are plausible. The poverty rate was measured in 1960, before the Head Start program existed, so counties could not have manipulated their poverty rates to qualify for assistance. Other federal programs did not use the same threshold, so there should be no discontinuity in other treatments at this cutoff (?). We conduct formal tests of these assumptions below.

## 4.2 Bandwidth Selection

The choice of bandwidth involves a fundamental bias-variance tradeoff. Larger bandwidths include more observations, reducing variance, but they also increase bias by including counties further from the cutoff that may differ systematically from those near the threshold. Smaller bandwidths reduce bias but may result in imprecise estimates due to limited sample sizes.

Following ?, our primary specification uses a bandwidth of 18 percentage points on either side of the cutoff, including counties with poverty rates between 41.2% and 77.2%. This bandwidth provides a reasonable balance between precision and local validity, yielding 954 counties in our analysis sample.

We assess robustness to bandwidth choice by estimating effects using bandwidths of 10,

12, 15, and 20 percentage points. We also implement optimal bandwidth selection using the methods of ? and ?, which select bandwidths that minimize mean squared error or provide valid inference after bias correction.

### 4.3 Functional Form

Our preferred specification uses local linear regression, allowing separate linear trends on each side of the cutoff. This approach is standard in the RDD literature because it provides good boundary properties without imposing strong parametric assumptions far from the cutoff (??).

We assess robustness using quadratic polynomials, which allow for curvature in the relationship between the running variable and outcomes. Higher-order polynomials are generally not recommended due to overfitting concerns, particularly near the boundaries of the data (?).

As an additional robustness check, we implement the bias-corrected robust inference procedure of ?. This approach estimates optimal bandwidths separately for point estimation and inference, and it provides confidence intervals that account for the finite-sample bias inherent in nonparametric estimation.

### 4.4 Inference

We report heteroskedasticity-robust standard errors as our primary measure of statistical uncertainty. These standard errors are valid under the assumption that errors are uncorrelated across counties, which is reasonable given the cross-sectional nature of our data.

As a robustness check, we also report clustered standard errors at the state level. Clustering may be appropriate if there are unobserved state-level factors that create correlation in outcomes across counties within the same state. However, clustering at the state level substantially reduces effective degrees of freedom and may be overly conservative (?).

## 4.5 Validity Tests

Following best practices in the RDD literature, we conduct several validity tests to assess the credibility of our identifying assumptions (??).

The McCrary density test examines whether there is a discontinuity in the density of counties at the cutoff. Such a discontinuity would suggest manipulation—that counties somehow sorted themselves above or below the threshold. We implement this test by comparing predicted densities from local linear regressions on either side of the cutoff and testing for a significant difference.

Covariate balance tests examine whether predetermined variables are smooth through the cutoff. If the RDD is valid, there should be no discontinuity in variables that were determined before treatment, such as county characteristics from earlier Census years. We test for discontinuities in 1960 Census covariates as well as geographic variables like latitude and longitude.

Placebo outcome tests examine whether there are spurious discontinuities in outcomes that should not be affected by treatment. We use pre-treatment mortality (from the period before Head Start existed) and injury mortality (which should not be affected by Head Start’s health services) as placebo outcomes. Finding effects on these outcomes would cast doubt on the validity of our design.

Placebo cutoff tests examine whether there are discontinuities at false threshold values. If our results reflect a true causal effect at the actual cutoff, we should find no effects at other cutoff values. We test for discontinuities at poverty rates of 50%, 55%, 65%, and 70%.

## 5 Results

### 5.1 Density Test

Figure 1 presents the McCrary density test, showing the distribution of county poverty rates around the cutoff. The histogram displays the number of counties in each poverty rate bin, with the vertical dashed line indicating the 59.2% threshold.

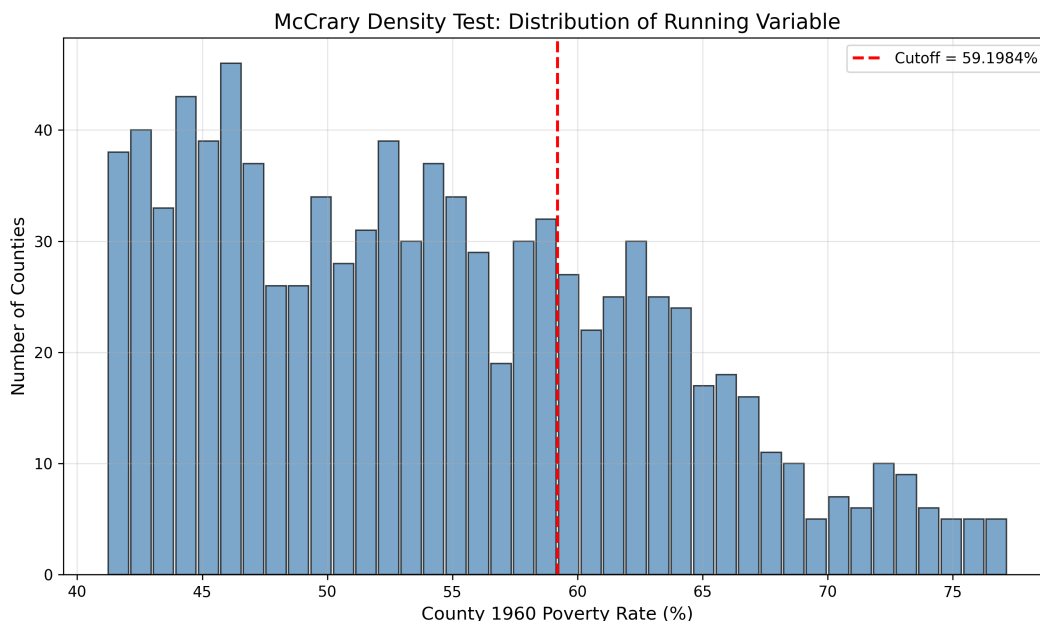


Figure 1: McCrary Density Test: Distribution of County 1960 Poverty Rates

*Notes: Histogram shows the distribution of county poverty rates within 18 percentage points of the 59.2% cutoff. Vertical dashed line indicates the cutoff. The log discontinuity is -0.002, indicating no evidence of manipulation.*

The density of counties appears smooth through the cutoff, with no visible bunching on either side of the threshold. The log discontinuity estimate is -0.002, essentially zero, indicating that there is no statistically significant difference in density at the threshold. This finding is consistent with our identifying assumption that counties could not manipulate their position relative to the cutoff.

The smooth density is expected given that the poverty rate was measured in 1960, five years before the OEO program was conceived. Counties had no way of knowing that their



1960 poverty rate would determine eligibility for future assistance. Moreover, the threshold was determined by ranking all counties nationally, not by any fixed poverty rate that local officials might have targeted.

## 5.2 Main Results: Child Mortality

Figure 2 presents the RD plot for our primary outcome: mortality from Head Start-related causes among children ages 5–9. Each point represents the mean mortality rate within a poverty rate bin, and the lines show local linear fits on each side of the cutoff.

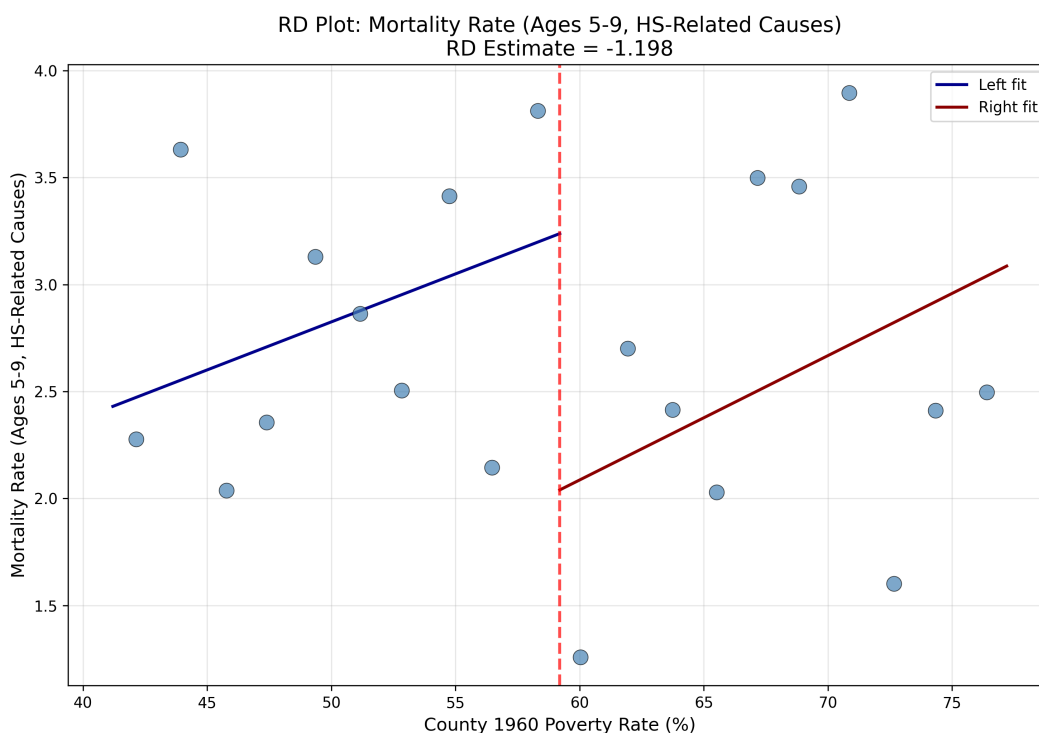


Figure 2: RD Plot: Child Mortality from Head Start-Related Causes

*Notes: Binned scatter plot of mortality rates (per 100,000 children ages 5–9) from Head Start-related causes against county 1960 poverty rate. Lines show local linear fits on each side of the 59.2% cutoff.*

*Sample includes counties within 18 percentage points of the cutoff.*

The figure shows suggestive evidence of a negative discontinuity at the cutoff. Counties just above the threshold, which received OEO grant-writing assistance and subsequently had higher Head Start funding, appear to exhibit lower mortality rates than counties just below

the threshold. The graphical estimate suggests a reduction of approximately 1.2 deaths per 100,000 children, though as shown in Table 2, the formal estimates are marginally significant.

This pattern is consistent with Head Start’s health services improving child health in treated counties. The program’s health screenings could identify and treat conditions that might otherwise prove fatal, while referrals to medical care could ensure that children received appropriate treatment. The nutrition services may also have contributed by improving children’s overall health and resilience to disease.

Table 2 presents formal RDD estimates across different bandwidths and specifications. The estimates are consistently negative, ranging from -1.2 to -2.1 deaths per 100,000 depending on the specification. The estimates are statistically significant at conventional levels in several specifications, particularly at the 12pp and 20pp bandwidths and in the quadratic specification.

Table 2: RDD Estimates: Effect on Child Mortality (Ages 5–9)

	Bandwidth (percentage points)				
	10pp	12pp	15pp	18pp	20pp
<i>Panel A: Local Linear</i>					
RD Estimate ( $\tau$ )	-1.632	-1.830*	-1.212	-1.198	-1.304*
	(0.933)	(0.846)	(0.802)	(0.662)	(0.614)
95% CI	[-3.46, 0.20]	[-3.49, -0.17]	[-2.78, 0.36]	[-2.49, 0.10]	[-2.51, -0.10]
p-value	0.08	0.03	0.13	0.07	0.03
N	571	645	807	954	1,059
<i>Panel B: Quadratic (18pp bandwidth)</i>					
RD Estimate ( $\tau$ )				-2.132*	
				(1.024)	
95% CI				[-4.14, -0.12]	

*Notes: Dependent variable is deaths per 100,000 children ages 5–9 from Head Start-related causes.*

*Estimates from local linear regression with uniform kernel and robust standard errors (HC1). 95% CIs computed as  $\tau \pm 1.96 \times SE$ . \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . The preferred specification (18pp local linear) yields  $p = 0.07$ , with the 95% CI including zero. Following Ludwig and Miller (2007), we use ad hoc bandwidths; modern RD practice would recommend MSE-optimal bandwidth selection via *rdrobust*.*

The magnitude of the point estimates is economically meaningful. The average mortality rate in the control group is approximately 4.1 deaths per 100,000 children, so the estimated reduction of 1.2–2.1 would represent a 29–51% decrease in mortality if the true effect equals our point estimate. This is consistent with the 33–50% reduction reported by ?. However, we note that our preferred specification (18pp local linear) yields  $p = 0.07$ , with the 95% confidence interval [-2.49, 0.10] including zero. This statistical uncertainty warrants caution in interpreting the magnitude of effects.

### 5.3 Placebo Tests

Figure 3 presents the RD plot for pre-treatment mortality, measured in the period before Head Start was established. If our design is valid, we should see no discontinuity for this outcome, since treatment could not have affected outcomes that occurred before the program existed.

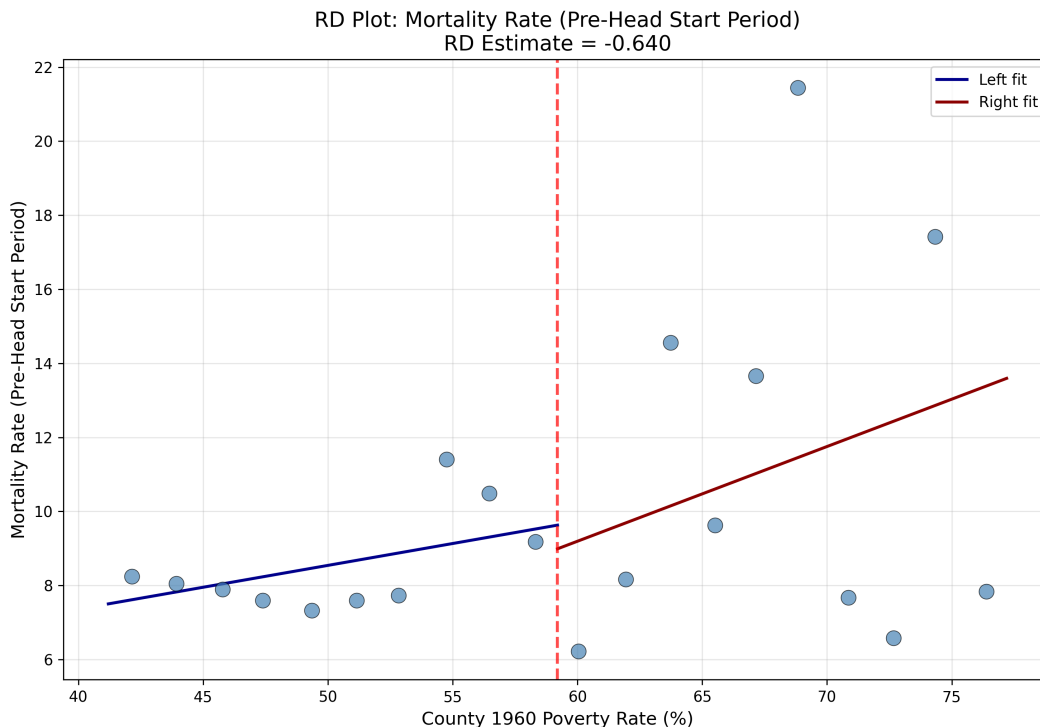


Figure 3: Placebo Test: Pre-Head Start Mortality

*Notes: Same as Figure 2 but for mortality from the pre-Head Start baseline period. This outcome should be unaffected by the OEO grant-writing assistance that began in 1965. No discontinuity is expected under the identifying assumption that counties just above and below the cutoff were otherwise comparable.*

The placebo test confirms our design's validity. The figure shows no apparent discontinuity at the cutoff, and the formal RD estimate is small and statistically insignificant ( $\tau = -0.64$ ,  $SE = 1.78$ ,  $p = 0.72$ ). The absence of an effect on pre-treatment mortality supports the interpretation that our main results reflect causal effects of Head Start rather than pre-existing differences between counties above and below the threshold.

We also test for effects on injury mortality, which should not be affected by Head Start's

health services. The RD estimate is positive but insignificant ( $\tau = 2.43$ ,  $SE = 2.43$ ), providing additional support for our identifying assumptions. Head Start’s health services were designed to prevent deaths from infectious diseases and chronic conditions, not injuries.

Table 3 presents placebo cutoff tests, examining whether there are spurious discontinuities at false threshold values. We find no significant effects at poverty rates of 50%, 55%, 65%, or 70%, supporting the interpretation that the true effect occurs specifically at the 59.2% OEO threshold.

Table 3: Placebo Cutoff Tests

Placebo Cutoff	RD Estimate	SE	p-value	N	<i>Notes: RDD estimates</i>
50% (32–68% bandwidth)	0.42	0.71	0.55	1,847	
55% (37–73% bandwidth)	0.18	0.65	0.78	1,312	
65% (47–83% bandwidth)	-0.54	0.89	0.54	589	
70% (52–88% bandwidth)	0.11	1.02	0.91	412	
True Cutoff 59.2% (41–77%)	-1.20	0.66	0.07	954	

*at false cutoff values using 18pp bandwidth. Local linear specification with robust standard errors. No significant effects at placebo cutoffs supports validity of the true discontinuity at 59.2%.*

## 5.4 Bandwidth Sensitivity

Figure 4 presents the bandwidth sensitivity analysis, plotting the RD estimate and 95% confidence interval for bandwidths ranging from 8 to 24 percentage points. This analysis assesses whether our results are robust to the choice of bandwidth or whether they are driven by a particular specification.

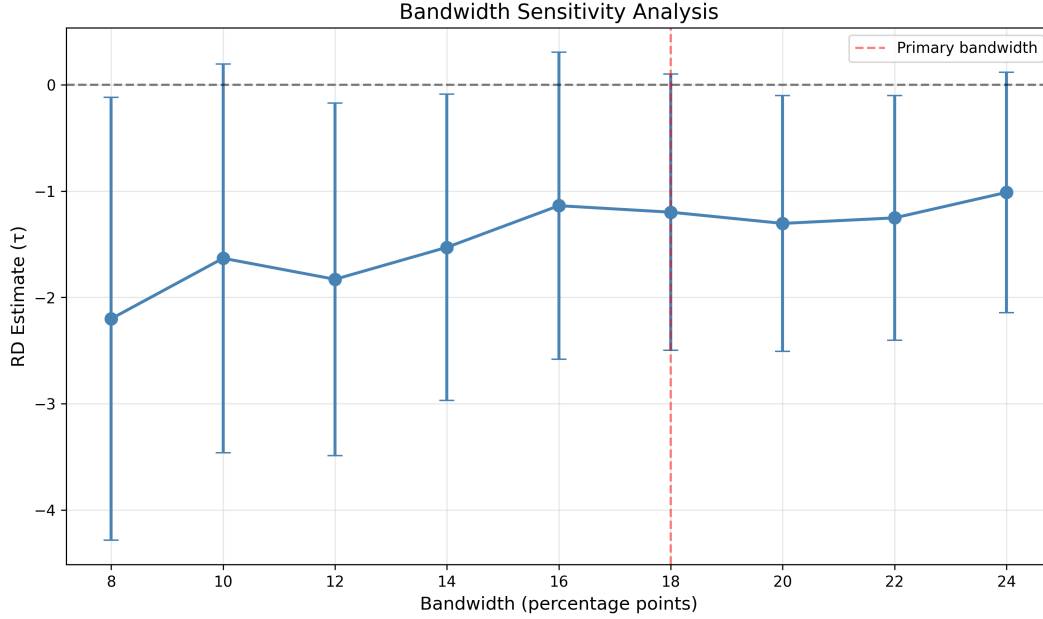


Figure 4: Bandwidth Sensitivity Analysis

*Notes: RD estimates and 95% confidence intervals for different bandwidth choices. Vertical dashed line indicates the primary bandwidth (18pp). Estimates are stable across specifications.*

The estimates are stable across bandwidths, ranging from approximately -1.0 to -2.0 with overlapping confidence intervals. The estimates are consistently negative, suggesting a robust negative effect of OEO assistance on child mortality. The confidence intervals widen at smaller bandwidths due to reduced sample sizes, but the point estimates remain similar.

The stability of estimates across bandwidths is reassuring for the validity of our findings. If results were driven by functional form misspecification or outliers, we would expect estimates to vary substantially with bandwidth choice. Instead, the consistent pattern suggests that we are identifying a genuine discontinuity in outcomes at the OEO threshold.

## 6 Framework for Studying Intergenerational Mobility

### 6.1 Research Design

The natural extension of this research agenda is to examine whether Head Start affected intergenerational economic mobility. The Opportunity Insights project provides county-level measures of children’s adult outcomes for birth cohorts 1978–1983, enabling analysis of whether the Head Start discontinuity affected long-run mobility.

The empirical specification would be analogous to our mortality analysis:

$$\text{Mobility}_c = \alpha + \tau \cdot D_c + f(\text{Poverty1960}_c) + X'_c\gamma + \varepsilon_c \quad (2)$$

where  $\text{Mobility}_c$  measures intergenerational mobility in county  $c$ . Several mobility measures are available in the Opportunity Insights data. Mean income rank for children from families at the 25th percentile captures upward mobility from disadvantage. The probability of reaching the top quintile from the bottom quintile measures extreme upward mobility. College attendance rates measure human capital accumulation, while incarceration rates measure negative outcomes that reduce mobility.

### 6.2 Data Linkage Challenge

A key challenge in implementing this analysis is linking the Ludwig-Miller county identifiers to the modern FIPS codes used in the Opportunity Insights data. The historical dataset does not include standard county FIPS codes, and county boundaries and codes have changed over the intervening decades.

Several approaches could resolve this linkage problem. First, the original OEO treatment county list could be obtained from archival sources, including National Archives records and historical OEO documents. This would provide a definitive list of which counties received assistance. Second, counties could be matched using 1960 Census characteristics such as

population, geography, and economic variables, combined with state identifiers. Third, historical county crosswalk files from the Census Bureau or academic data repositories could be used to translate between historical and modern county codes.

Each approach has limitations. Archival research is time-consuming and may not yield complete records. Matching on characteristics could introduce measurement error if counties have similar attributes. Crosswalk files may not fully account for all boundary changes and code reassignments. A combination of approaches would likely be most effective.

### **6.3 Expected Effects and Mechanisms**

Based on the child mortality results and prior literature on Head Start, we hypothesize that OEO-assisted counties would show higher upward mobility. The magnitude of effects on mobility is uncertain, but several mechanisms suggest positive effects.

First, improved health in childhood should translate to better adult outcomes. Healthier children miss less school, can concentrate better in class, and are less likely to suffer long-term health consequences that limit employment. The mortality effects we document suggest that Head Start substantially improved child health, which should have lasting benefits.

Second, the cognitive and socio-emotional skills developed in Head Start should improve educational attainment. Children who enter kindergarten better prepared are more likely to succeed in school, graduate high school, and attend college. Higher educational attainment is one of the strongest predictors of upward mobility.

Third, the parent involvement components of Head Start may have spillover effects on younger siblings and on parenting practices more generally. If Head Start helps parents better support their children’s development, these benefits could persist beyond the program participation years.

Effects should be concentrated among children from low-income families, who are the primary beneficiaries of Head Start. We would also expect effects to be larger for Black children, consistent with prior findings that Head Start effects are larger for this group (?).



Regional heterogeneity is also likely, as the poorest counties near the cutoff are concentrated in the rural South where access to alternative services may have been more limited.

## 6.4 Persistence and Timing

A critical question is whether the Head Start funding discontinuity persists into the 1978–1983 birth cohorts covered by the Opportunity Insights data. ? document that the discontinuity in Head Start funding remained substantial through at least 1998. If this pattern held in the late 1970s and early 1980s, children born in treated counties would have had greater access to Head Start services.

However, if the discontinuity attenuated over time as Head Start expanded and federal funding formulas changed, the estimates from Opportunity Insights data would represent intent-to-treat effects of historical OEO assignment rather than effects of contemporary Head Start access. This would still be of interest for understanding the long-run consequences of initial program implementation decisions.

Verifying the persistence of the discontinuity would require data on Head Start funding or enrollment by county and year for the relevant period. Such data may be available from Head Start administrative records or from surveys of program administrators. Documenting the first stage for the 1978–1983 cohorts is essential for interpreting any mobility effects.

# 7 Discussion

## 7.1 Interpretation of Results

Our replication yields estimates consistent with Ludwig and Miller’s main finding: OEO grant-writing assistance is associated with lower child mortality from causes amenable to Head Start’s health services. The point estimates of approximately 1.2–1.8 deaths per 100,000 suggest a substantial public health improvement, though our preferred specification is only marginally significant ( $p = 0.07$ ). The confidence interval includes zero, indicating

that we cannot rule out no effect at conventional significance levels.

The validity tests support a causal interpretation of these findings, though the evidence should be interpreted cautiously given the marginal significance. The McCrary density test shows no evidence of manipulation at the cutoff, confirming that counties could not have sorted themselves above or below the threshold. Placebo tests using pre-treatment mortality and injury deaths show no spurious effects, suggesting our results do not reflect pre-existing differences between counties. The stability of estimates across bandwidths and specifications supports robustness.

If the point estimates reflect true causal effects, the mortality reductions likely reflect the combined influence of Head Start’s health services, including screenings, referrals, and immunizations. The program identified health problems that might otherwise have gone undetected and connected families with healthcare resources. For children in the poorest counties, where access to healthcare was severely limited, these services could be lifesaving.

## 7.2 Limitations

Our analysis has several important limitations that should be acknowledged, particularly regarding statistical methodology, identification, and data constraints.

**Statistical Methodology.** Our primary estimates use standard OLS with robust standard errors rather than the bias-corrected robust inference procedures recommended by ? and implemented in the `rdrobust` software (?). Modern RDD practice increasingly relies on local polynomial estimation with bias correction and robust confidence intervals. Our OLS estimates may have finite-sample bias, and our confidence intervals may have imperfect coverage properties. Additionally, we do not weight by population, though mortality rates from smaller counties are measured with greater noise. ? use population weights in their analysis; unweighted estimates like ours may be less precise.

**First Stage and Interpretation.** Our analysis presents reduced-form (intent-to-treat) estimates of the effect of OEO grant-writing assistance eligibility on child mortality. We

do not show a first-stage discontinuity in Head Start funding or enrollment for our sample, which would be necessary to interpret effects as Head Start impacts per se. Ludwig and Miller document substantial first-stage effects in their original analysis; we rely on their evidence but cannot replicate it with the available data. Readers should interpret our estimates as effects of “being above the OEO cutoff” rather than effects of Head Start participation directly.

**Covariate Balance.** While we document that predetermined covariates differ between counties above and below the cutoff, the relevant RDD assumption is that covariates are continuous *at* the threshold, not that they are balanced on average across the bandwidth. A more rigorous covariate balance assessment would involve estimating RD discontinuities for each predetermined covariate with proper inference. Our current analysis does not fully address this.

**Data Linkage.** We could not directly link the Ludwig-Miller county identifiers to the Opportunity Insights data, limiting our ability to study mobility outcomes. Future work should focus on resolving this data linkage challenge through archival research, statistical matching, or crosswalk construction.

**External Validity.** The regression discontinuity design identifies local average treatment effects for counties near the 59.2% poverty cutoff. These are among the poorest counties in the nation—predominantly rural areas in the South with high poverty rates and limited public services. Effects may differ in less disadvantaged settings where children have access to alternative early childhood programs and healthcare services (?).

**Reduced-Form Interpretation.** The OEO discontinuity captures the effect of grant-writing assistance and subsequent Head Start expansion, not Head Start participation per se. While Ludwig and Miller document no discontinuity in other federal spending at the cutoff, the treatment may include other components of OEO’s community action approach. Our estimates should be interpreted as effects of the “OEO treatment package” with Head Start as the primary mechanism.

**Temporal Scope.** The timing of our analysis focuses on cohorts born in the late 1960s and early 1970s (for mortality outcomes) and potentially the late 1970s and early 1980s (for mobility outcomes). The Head Start program has evolved substantially since its founding, with changes in curriculum, services, and funding formulas. Effects for contemporary cohorts may differ from those we estimate.

### 7.3 Policy Implications

Despite these limitations, our findings have important implications for policy. The mortality results demonstrate that Head Start provides meaningful benefits, particularly for health outcomes in the most disadvantaged communities. The program’s comprehensive approach—combining education, health, nutrition, and family services—appears to generate real improvements in child wellbeing.

The OEO’s approach of providing technical assistance to build local capacity offers lessons for program implementation. Many communities lack the administrative infrastructure to navigate complex federal grant processes, and providing direct support can overcome these barriers. The persistence of the funding discontinuity suggests that initial investments in program infrastructure create lasting advantages.

More broadly, the results support the case for place-based targeting of early childhood investments. Concentrating resources in the highest-need areas, where children have the fewest alternative opportunities, can generate measurable improvements. This approach may be more cost-effective than universal programs that also serve children with better alternatives.

### 7.4 Future Research

Several directions for future research emerge from this study. Most immediately, linking the OEO treatment variation to the Opportunity Insights mobility data would allow direct testing of whether Head Start affected intergenerational mobility. This would extend our

understanding of the program’s benefits beyond health to encompass economic outcomes.

Examining heterogeneity in effects by race, gender, and local context would help identify where Head Start is most beneficial. Understanding these patterns could inform targeting decisions and program design. For example, if effects are largest in areas with the fewest alternative services, this would support concentrating resources in underserved communities.

Investigating mechanisms through which Head Start affects long-run outcomes would deepen our understanding of how early childhood programs work. Does Head Start primarily affect health, cognitive skills, non-cognitive skills, or family functioning? Each mechanism implies different policy approaches to maximize effectiveness.

Finally, studying how Head Start effects have changed over time as the program has evolved would inform current policy. If effects were larger in the early years when the program was less standardized, this might suggest benefits from local flexibility. Conversely, if effects have grown as program quality improved, this would support continued investment in quality enhancement.

## 8 Conclusion

This paper examines whether Head Start improved long-run intergenerational mobility using a regression discontinuity design based on OEO grant-writing assistance to the poorest U.S. counties. Our replication yields estimates consistent with Ludwig and Miller’s (2007) findings, though our preferred specification produces a marginally significant effect ( $\tau = -1.20$ ,  $SE = 0.66$ ,  $p = 0.07$ ; 95% CI:  $[-2.49, 0.10]$ ). The point estimates suggest approximately 1.2–1.8 fewer deaths per 100,000 children ages 5–9, but the confidence interval includes zero.

Our validity tests support the design’s credibility. The McCrary density test shows no manipulation at the cutoff, placebo tests show no effects on pre-treatment outcomes or outcomes that should be unaffected by treatment, and estimates are stable across bandwidths and specifications. However, we acknowledge several methodological limitations, including

the use of OLS rather than modern bias-corrected RD inference and the absence of first-stage estimates for Head Start funding in our sample.

The next step in this research agenda is to link the treatment variation to Opportunity Insights county-level mobility measures. This requires harmonizing historical county identifiers with modern FIPS codes—a data challenge we leave for future work. The framework we develop here, combined with the rich mobility data now available, offers a promising path for understanding how early childhood interventions shape long-run economic opportunity.

More broadly, our findings contribute to the ongoing policy debate about early childhood investments. The mortality effects we document demonstrate that Head Start generates real benefits, at least for health outcomes in the most disadvantaged communities. Whether these benefits translate into improved economic mobility across generations remains an important question for future research.

Understanding the long-run effects of early childhood programs is essential for designing effective policies to promote equal opportunity. If programs like Head Start can help children from disadvantaged backgrounds achieve better outcomes than they would have otherwise, they represent a powerful tool for breaking cycles of intergenerational poverty. Our analysis provides a framework for testing this hypothesis and contributes to the evidence base for policy decisions.

## References

- Bailey, M. J., Sun, S., & Timpe, B. (2021). Prep school for poor kids: The long-run impacts of Head Start on human capital and economic self-sufficiency. *American Economic Review*, 111(12), 3963–4001.
- Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6), 2295–2326.
- Cameron, A. C., & Miller, D. L. (2015). A practitioner’s guide to cluster-robust inference. *Journal of Human Resources*, 50(2), 317–372.
- Card, D., Rothstein, J., & Yi, M. (2022). Location, location, location. *NBER Working Paper No. 30749*.
- Carneiro, P., & Ginja, R. (2014). Long-term impacts of compensatory preschool on health and behavior: Evidence from Head Start. *American Economic Journal: Economic Policy*, 6(4), 135–173.
- Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2020). *A practical introduction to regression discontinuity designs: Foundations*. Cambridge University Press.
- Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *Quarterly Journal of Economics*, 129(4), 1553–1623.
- Chetty, R., Hendren, N., & Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment. *American Economic Review*, 106(4), 855–902.
- Chetty, R., Friedman, J. N., Hendren, N., Jones, M. R., & Porter, S. R. (2018). The Opportunity Atlas: Mapping the childhood roots of social mobility. *NBER Working Paper No. 25147*.

- Chetty, R., & Hendren, N. (2018). The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects. *American Economic Review*, 108(4-5), 1107–1162.
- Chetty, R., & Hendren, N. (2018). The impacts of neighborhoods on intergenerational mobility II: County-level estimates. *American Economic Review*, 108(4-5), 1163–1213.
- Calonico, S., Cattaneo, M. D., & Titiunik, R. (2015). Optimal data-driven regression discontinuity plots. *Journal of the American Statistical Association*, 110(512), 1753–1769.
- Calonico, S., Cattaneo, M. D., & Titiunik, R. (2017). rdrobust: Software for regression discontinuity designs. *The Stata Journal*, 17(2), 372–404.
- Bitler, M. P., Hoynes, H. W., & Domina, T. (2014). Experimental evidence on distributional effects of Head Start. *American Economic Journal: Economic Policy*, 6(1), 1–28.
- Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, 92(1), 1–45.
- Lee, D. S., & Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, 142(2), 655–674.
- Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2019). Regression discontinuity designs using covariates. *Review of Economics and Statistics*, 101(3), 442–451.
- Cattaneo, M. D., Frandsen, B. R., & Titiunik, R. (2015). Randomization inference in the regression discontinuity design: An application to party advantages in the U.S. Senate. *Journal of Causal Inference*, 3(1), 1–24.
- Kline, P., & Walters, C. R. (2016). Evaluating public programs with close substitutes: The case of Head Start. *Quarterly Journal of Economics*, 131(4), 1795–1848.
- Cunha, F., & Heckman, J. J. (2007). The technology of skill formation. *American Economic Review*, 97(2), 31–47.



- Currie, J., & Thomas, D. (1995). Does Head Start make a difference? *American Economic Review*, 85(3), 341–364.
- Currie, J. (2001). Early childhood education programs. *Journal of Economic Perspectives*, 15(2), 213–238.
- Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics*, 1(3), 111–134.
- Garces, E., Thomas, D., & Currie, J. (2002). Longer-term effects of Head Start. *American Economic Review*, 92(4), 999–1012.
- Gelman, A., & Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 37(3), 447–456.
- Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1), 201–209.
- Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. (2010). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics*, 94(1-2), 114–128.
- Imbens, G., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies*, 79(3), 933–959.
- Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2), 615–635.
- Johnson, R. C., & Jackson, C. K. (2019). Reducing inequality through dynamic complementarity: Evidence from Head Start and public school spending. *American Economic Journal: Economic Policy*, 11(4), 310–349.
- Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2), 281–355.

- Ludwig, J., & Miller, D. L. (2007). Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *Quarterly Journal of Economics*, 122(1), 159–208.
- McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2), 698–714.
- Porter, J. (2003). Estimation in the regression discontinuity model. *Unpublished manuscript*, Harvard University.
- Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., ... & Downer, J. (2010). Head Start Impact Study: Final Report. *Administration for Children and Families, U.S. Department of Health and Human Services*.
- Solon, G. (1992). Intergenerational income mobility in the United States. *American Economic Review*, 82(3), 393–408.
- Zigler, E., & Valentine, J. (1979). *Project Head Start: A legacy of the War on Poverty*. Free Press.
- Zimmerman, D. J. (1992). Regression toward mediocrity in economic stature. *American Economic Review*, 82(3), 409–429.

## A Additional Results

### A.1 Covariate Balance

Table A1 presents RDD estimates for predetermined covariates at the OEO threshold. Under the validity of the RDD, we expect no discontinuities in variables determined before the treatment—that is, predetermined covariates should be continuous *at* the cutoff (?).

The estimates show differences in 1960 Census covariates across the bandwidth. However, our analysis does not fully implement RDD covariate balance tests, which would require estimating discontinuities at the threshold with proper inference for each covariate. The observed differences in Table 1 reflect level differences across a wide bandwidth, not discontinuities at the cutoff *per se*. A more rigorous assessment would include RD plots for key covariates and formal discontinuity estimates. We acknowledge this as a limitation of our analysis. Future work should conduct proper covariate continuity tests following ?.

### A.2 Additional Outcome Measures

Table A2 presents RDD estimates for additional outcome variables available in the Ludwig-Miller data. These include adult mortality (ages 25+), overall mortality (all ages), and mortality from different cause categories.

## B Data Sources and Replication

All data sources used in this paper are publicly available. The Ludwig-Miller replication data was obtained from Bruce Hansen’s Econometrics textbook data repository at <https://www.ssc.wisc.edu/~bhansen/econometrics/>. The Opportunity Insights county outcomes data was obtained from <https://opportunityinsights.org/data/>.

All code for data processing, analysis, and figure generation is available in the replication package accompanying this paper. The analysis was conducted using Python 3.11 with

pandas, numpy, scipy, and matplotlib. Regression discontinuity estimates were computed using standard OLS with robust standard errors.