Early Childhood Intervention and Income Inequality: An Analysis on the Intergenerational Mobility of Head Start Participants

PAUL HINTON

Carleton College

ABSTRACT: The Head Start program is an early childhood intervention program funded by the federal government. Designed for low-income families, it promotes school readiness among its participants. In my research, I examine the effects of Head Start on the intergenerational mobility of its participants. Using data from the National Longitudinal Survey of Youth (NLSY79) and the NLSY Child and Young Adult Supplement (CNLSY), I measure the degree of earnings persistence between parents and their children for students in preschool and Head Start. I find that participation in Head Start is associated with an increase in mobility, which remains true after controlling for family effects.

I. Introduction

Countries with lower intergenerational mobility tend to have higher income inequality. I investigate whether increased access to education through the Head Start program leads to more intergenerational mobility. This higher intergenerational mobility could lead to a decrease in income inequality, a growing problem in the United States. The Head Start Program is a comprehensive early childhood intervention program that educates low-income children to prepare them for primary school. In addition to a traditional preschool education, the program provides health, nutrition, and parent involvement services. Past literature has investigated the long-term effects of Head Start on metrics such as participant earnings, test scores, and crime rates.

Researchers judge the immediate success of programs like Head Start through measures such as test scores. While research finds that test scores increase right after a child’s participation in the program, these increases dissipate over time in what has been called the “Head Start fade.” However, test scores are not the only determinant of success for the program. Children can still benefit in the longer run as a result of better school preparation. Increased school readiness for students could cause a change in trajectory throughout primary school and make them more likely to have positive future outcomes. If Head Start participants improve their economic status relative to their parents more than students who do not enroll in Head Start, then enrollment in Head Start could decrease income inequality in the United States for future generations.

I believe my research contributes to the literature surrounding Head Start. Previous research focuses on test scores and individual young adult outcomes such as high school graduation and teen parenthood. My research provides an alternate lens for looking at its potential benefits. By looking at the intergenerational mobility of Head Start participants compared to non-participants, I find that Head Start participants have more economic mobility. That is, the income of a Head Start participant’s parents is less predictive of the participant’s future income than it is for someone who does not participate in Head Start. This suggests that early childhood intervention programs such as Head Start accrue lasting benefits for its participants.
II. Context and Theory

Connecting Mobility, Inequality, and Education

Previous research finds associations between income inequality and other phenomena that help explain rising income inequality in countries such as the United States. A country’s intergenerational mobility is one potential link that could account for this rise. Corak (2013) discusses this relationship through the Great Gatsby Curve. Among OECD countries, those with higher income inequality also tend to experience lower intergenerational mobility. As a measure of intergenerational mobility, Corak uses the intergenerational earnings elasticity (IGE) of a country. The IGE is the coefficient $\beta$ from the regression equation of one generation’s income $Y_{i, t}$ versus a previous generation’s income $Y_{i, t-1}$:

$$\ln Y_{i, t} = \alpha + \beta \ln Y_{i, t-1} + \epsilon_i.$$ 

In this equation, $i$ represents each individual, $t$ represents the current generation and $t - 1$ represents the previous generation. The coefficient $\alpha$ is the trend in average income across generations and $\epsilon$ is an error term that captures any other influences on the current generation’s income. This is the primary empirical strategy I use in my analysis of Head Start participation and intergenerational mobility; I expand on the strategy in Section III.

The coefficient $\beta$ in the above equation measures the degree to which a parent’s earnings allow us to predict a child’s earnings. Higher values of $\beta$ imply lower intergenerational mobility. For some percent variation in a parent’s income, the intergenerational elasticity tells us the expected percent variation in a child’s income. For example, if a parent’s income is 50 percent above the mean income for his generation, then an IGE of 1 implies a child will also have an income 50 percent above the mean income for the child’s generation. An IGE of 0, on the other hand, implies a parent’s income has no relationship to the child’s income. A parent’s income bracket is thus more likely to be the same as the child’s future income bracket if we observe low intergenerational mobility. Corak (2013) compares the IGE to a country’s level of income inequality using the Gini coefficient. In his data, countries with a higher Gini coefficient also have a higher IGE. The United States is at the upper end of the relationship: it has high income inequality and low intergenerational mobility compared to other countries. Since Corak’s study, research has found multiple explanations for the relationship between inequality and mobility.

One such study by Jerrim and Macmillan (2015) suggests education is the primary link between income inequality and intergenerational mobility. They present empirical evidence of this relationship across countries that participate in the Programme for the International Assessment of Adult Competencies (PIACC). Instead of regressing log child income on log parent income, the authors instead regress log child income on the highest level of parental education, claiming this method will capture financial, social, and cultural effects on the child. They analyze two different effects in their paper: (1) the association between educational attainment of parents and children and (2) the overall returns on education for the child. Jerrim and Macmillan find both effects are positively associated with income inequality among the countries in the survey, with the persistence of educational attainment between parents and children being the stronger result. From this, they conclude transmission of education is important for intergenerational mobility, particularly through financial investments in education. If these financial investments are primarily
at the top of the income distribution, then low-income families are left behind in education and income inequality is expected to be higher.

One potential consequence of differential investment in education is differing school quality. Grawe (2010) investigates the relationship between intergenerational earnings mobility and school quality for male students, which he proxies using the pupil-to-teacher ratio.\textsuperscript{17} Theoretically, this proxy’s effect on the IGE could be either positive or negative. A higher pupil-to-teacher ratio could imply a lower school quality because schools are unable to hire more teachers. In these scenarios, each student may receive less individual attention and would therefore not learn as much as he would at a school with a lower ratio. On the other hand, a lower pupil-to-teacher ratio may imply that, while a school hires many teachers, the average teacher has lower qualifications than in an identical school with the same budget that hires fewer teachers. Grawe finds a positive relationship between intergenerational mobility and the pupil-to-teacher ratio; for an increase in the pupil-to-teacher ratio, the mobility of sons increases.\textsuperscript{18} While this study may say nothing about the effects of improvements in early childhood intervention, it suggests that school quality could affect mobility.

Assuming that communities with different average income levels have different levels of financial investment in education, socioeconomic segregation represents a potential hindrance to improvements in educational investment and attainment for low-income groups. In their paper “Understanding the Great Gatsby Curve,” Durlauf and Seshadri (2018) argue that social factors determine education and human capital, with education acting as the mediator between parental income and the future income of their offspring.\textsuperscript{19} The authors show that an equilibrium outcome of the Great Gatsby Curve can occur as a result of high socioeconomic segregation.\textsuperscript{20} Central to this analysis is the assumption that people prefer to live with more affluent neighbors in larger areas. These preferences lead to socioeconomic segregation of neighborhoods.\textsuperscript{21} For neighborhoods with a higher average income, members contribute more to public investments in education than neighborhoods with a lower average income. These investments will translate into greater overall investment in human capital for higher-income children in more affluent neighborhoods. The authors propose this difference in investment causes a disparity in accumulated human capital between children of low-income neighborhoods and high-income neighborhoods.\textsuperscript{22} By assuming early accumulation of human capital predicts labor market outcomes in adulthood, Durlauf and Seshadri conclude that socioeconomic segregation leads to low intergenerational mobility and persistent income inequality.\textsuperscript{23} Empirical research could further test this theory, but my research focuses only on the individual decisions of parents to invest in their children’s human capital, not societal decisions as a whole.

Many students in the United States go to school depending on their residential district, implying residential segregation and school segregation are closely connected.\textsuperscript{24} Arenas and Hindriks (2021) investigate the latter phenomenon in their research on intergenerational mobility and unequal school opportunity.\textsuperscript{25} They simulate different outcomes of intergenerational mobility and efficiency to investigate the relationship between income inequality and low mobility. In this instance, the authors define a change in efficiency as the change in average human capital.\textsuperscript{26} They use the US parental income distribution to assess the tradeoff between school desegregation and efficiency.\textsuperscript{27} In their model, unequal school opportunity includes both unequal school quality and different probabilities of access to the best schools, which is very similar to the story of socioeconomic segregation told by Durlauf and Seshadri.\textsuperscript{28} This unequal opportunity causes a gap between high- and low-income families where higher-income families will match with higher-quality schools. Arenas and Hindriks assume that investment in human capital and school quality
are complements.\textsuperscript{20,30} Therefore, a higher rate of positive assortative matching implies higher school quality and higher average investment in human capital. In their simulation, Arenas and Hindriks test which effect dominates by changing school segregation and school inequality variables. They find the increase in mobility from “creating” more equal schools is higher than the loss in efficiency from this equalization.\textsuperscript{31} Finding ways to make schools more equal could therefore improve economic mobility. Later, I discuss how early childhood interventions such as Head Start could improve overall school quality.

*Traditional Models of Intergenerational Mobility*

My research focuses on a parent’s individual decision to invest in his or her child’s human capital and how this decision relates to intergenerational mobility. Becker and Tomes (1979) build the primary economic theory for intergenerational mobility.\textsuperscript{32} Starting with the first principles of utility maximization, a parent faces the decision to invest either in his own consumption or in his child’s consumption. The parent obtains some discounted utility from his child’s utility—this depends on an intergenerational discount factor that represents the parent’s degree of altruism toward his child. As an equation, we represent this as

\[ U = u(c) + \beta * u(c') \]

where \( U \) is the parent’s overall utility, \( c \) is the parent’s consumption, \( c' \) is the child’s consumption, and \( \beta \) is the intergenerational discount factor. Because this equation separates the utility of a parent’s consumption from the utility of his child’s consumption, we can just focus on maximizing the child’s utility in our analysis. By assumption, the child’s consumption \( c' \) is equivalent to his future earnings:

\[ c' = w = ah^\gamma + x(1 + r). \]

Here, \( a \) is inherent ability, \( h^\gamma \) is the parent’s investment in human capital, \( x \) is the endowment that a parent gives to a child, and \( r \) is the interest rate on the endowment. If a parent maximizes his child’s future earnings, then the parent also maximizes his utility from his child’s consumption. A parent determines his child’s future earnings through investment in the child’s human capital and endowments to the child. Investment in human capital, \( h^\gamma \), is an increasing function that exhibits decreasing marginal returns because \( 0 < \gamma < 1 \). For two children with the same investment in human capital, the child with a higher inherent ability \( a \) will also have higher future wages. As a simplifying assumption, we assume the price of each unit of human capital is $1. The endowment given to the child, \( x \), earns interest \( r \) over time at a rate of \((1 + r)\). Solving for the optimal investment in a child’s human capital is therefore reduced to taking the partial derivatives of our first order condition:

\[
\frac{\partial w}{\partial h} = a\gamma h^{\gamma-1} \\
\frac{\partial w}{\partial x} = 1 + r
\]

From these conditions, setting \( \frac{\partial w}{\partial h} \) equal to \( \frac{\partial w}{\partial x} \) gives the optimal investment in a child’s human capital. Here, \( \frac{\partial w}{\partial h} \) is the marginal benefit of an additional unit of investment in human capital and \( \frac{\partial w}{\partial x} \) is the marginal cost of the additional investment. We think of this marginal
cost as an opportunity cost—instead of investing money into a child’s human capital, the parent could instead use the money to give a bequest to the child that earns an interest of $1 + r$. Equating marginal benefit to marginal cost, we derive a parent’s optimal investment in a child’s human capital.

However, if credit constraints are present, then the parent’s optimization problem must change. Now, the parent no longer has enough capital to give an endowment to his child. Therefore, the value $x$ in the first order condition becomes zero and any investment in the child is only in his human capital. The marginal cost of a parent’s investment in a child’s human capital is therefore also affected by some credit constraint $\lambda$. The new marginal cost is therefore $\frac{\partial w}{\partial x} = 1 + r + \lambda$. The marginal cost curve shifts up and results in a decreased investment in a child’s human capital compared to a situation with no credit constraints. Figure 1 depicts this difference in the two scenarios, with $h$ representing optimal investment without credit constraints and $h'$ representing optimal investment with credit constraints.

**Figure 1: Optimal Allocation of Investment in a Child’s Human Capital**

Solon (2004) presents a theoretical model that solves for a steady state of intergenerational mobility. In this model, a child’s human capital depends on a parent’s investment in the child’s human capital, a public investment in the child’s human capital, and an additional human capital endowment of the child independent of the investment in the child’s human capital. This endowment is the child’s ability and represents traits the child receives either through nature or nurture. Solon gives examples of family culture, family reputation, and inherent ability as potential effects on a child’s human capital endowment. A higher endowment implies the child will have higher human capital. Two important conclusions from the steady state equilibrium in this model are that (1) parents invest more in their child’s human capital if their income is higher and (2) the intergenerational earnings elasticity is lower if a community’s investment in public education is higher.

Corak’s (2013) findings empirically confirm conclusion (1). In the past 50 years, parental investment in enrichment expenditures for education such as computers and private schooling have increased by about $5,000 in 30 years for those in the top income quintile but have stayed about
the same for those in the bottom income quintile. This conclusion is important because a higher private investment in human capital could lead to a higher human capital in children and therefore higher future earnings for children of families from higher income brackets. A remedy for this inequality could be public investment improvements for a child’s education. As conclusion (2) states, higher public investment results in a lower intergenerational earnings elasticity. Parents who otherwise could not afford more investment in their child’s education now receive the benefits of a public investment in their child’s education. Programs such as Head Start seek to improve the suboptimal investment in a child’s human capital through public investment in the child’s human capital, which Becker and Tomes (1979) call a productive activity of the government. This improves the overall stock of human capital for children at the lower end of the income distribution and leads to higher wages. These higher wages then cause a decrease in the intergenerational earnings elasticity.

Early Childhood Intervention’s Promising Effects

Even though investments in a child’s human capital can occur at any point during the child’s youth, Johnson and Jackson (2019) argue through the concept of dynamic complementarity that early educational investments can help break the cycle of poverty. This concept comes from Cunha and Heckman (2007) in their paper “The Technology of Skill Formation.” They state that the skills a child learns early on in development improve later investments in a child’s human capital. This synergy has compounding effects as a child continues to gain skill and improve his human capital. Without early skill building, a child will fall behind and therefore be unable to reap the benefits of later investment in his education. By preparing a child for elementary school education, the Head Start program can provide a direct channel for children to become school-ready when entering the K-12 system.

In addition, spillover effects to non-Head Start participants from children enrolled in Head Start could improve the learning environment for all students. Because children who enroll in Head Start may be in a more socioeconomically segregated area, these positive spillover effects might cause improved educational and earnings outcomes for all students, not just those who participated in Head Start. While my research only focuses on the intergenerational effects of students who participate in Head Start, future analysis could also analyze the intergenerational effect of Head Start on an entire community, not just on individual participants in Head Start. This analysis would then tie back into the earlier discussion of socioeconomic segregation through education; if Head Start had benefits from spillover effects, this could improve education in lower-income neighborhoods and thus yield a decline in income inequality.

Extending this idea of community benefits of early childhood intervention, Art Rolnick and Rob Grunewald (2003) argue in favor of early child development programs and perform a cost-benefit analysis of the Perry Preschool program. This was an early childhood intervention program in Ypsilanti, Michigan that provided high-quality education for children who were three and four. Instead of public investments in economic development, such as subsidizing businesses and improving professional sports stadiums, Rolnick and Grunewald argue that public investment in an early child development program would provide a much higher rate of return. Over the five years of the Perry Preschool program’s existence, they estimate the total overall benefits to participants and the public in Ypsilanti to be $108,002 and the cost to be $12,356 in 1992 dollars. Benefits for the public come in the form of a decrease in crime and welfare payments and in a more efficient K-12 education.
A similar analysis by Barnett (1995) finds comparable benefits to the Perry Preschool program. In addition to his Perry Preschool analysis, Barnett reviews 36 previous studies on early childhood intervention programs to examine their long-term effects on children from low-income families. He finds positive long-term effects in grade retention and special education placement to be among the most significant observations from the studies. Although others have investigated the long-term effects of small, targeted programs, I assess the far-reaching program of Head Start to determine whether positive long-term effects of Head Start on intergenerational mobility exist.

**Head Start’s Long-Run Outcomes**

The literature on Head Start has primarily looked at the program’s effects on outcomes other than intergenerational mobility. Deming (2009) first replicates previous research on initial test score outcomes for Head Start participants, then analyzes Head Start’s impact on six different outcomes including college attendance, crime, and high school graduation. He finds that after applying family fixed effects in the regression, students who participate in Head Start are predicted to score 0.145 standard deviations higher on standardized tests than students who had no preschool at ages 5-6. Deming then goes on to show that although students have improved test scores right after Head Start, the improvements fade out by the time they are in middle school. This pattern is similar to Lee and Loeb’s (1995) previous results.

More central to Deming’s paper, however, are his results on long-term outcomes of participants. When Deming regresses an index of these outcomes—which includes college attendance, crime rates and high school graduation—on participation in Head Start, he finds a statistically significant impact of 0.228 standard deviations for children who participate in Head Start compared to children who do not. This result implies that long-term impacts of Head Start extend beyond just initial test score gains. Deming reports improvements in multiple individual outcomes: the likelihood of high school graduation increases, the likelihood of completing some college increases, and the likelihood of idleness decreases. People were considered idle if they were not enrolled in school and reported zero wages in 2004. These findings suggest long-term benefits of Head Start participants are present and thus prompts further investigation into the program’s future benefits for children.

Garces et al. (2002) perform a similar analysis to Deming’s—they look at four young adult outcomes for white and Black participants using data from the Panel Study of Income Dynamics (PSID). These outcomes are high school graduation rates, some college education, log earnings, and crime rates. When controlling for observable characteristics, white Head Start participants are 20 percentage points more likely to complete high school than their siblings who did not attend Head Start. However, this result does not hold for Black Head Start participants. But Garces et al. do find a decrease in crime rates for Black Head Start participants compared to their siblings who had no early childhood education. This crime reduction suggests a positive social benefit from Head Start that could be measured in a cost-benefit analysis. Most notable is the finding that earnings between the ages of 23 to 25 are significantly higher for children who attended Head Start compared to children that either had no early education or attended preschool. Higher earnings for children from low-income backgrounds could help break the cycle of poverty and provides a good grounding for my research on the intergenerational mobility of children in Head Start.
III. Hypothesis, Data, and Empirical Strategy

Hypothesis

In my research, I aim to continue Deming’s work and explore the implications of the Head Start program on intergenerational income mobility. I hypothesize that the intergenerational earnings elasticity, or IGE, of Head Start participants is smaller than the IGE of non-participants. This hypothesis would suggest the Head Start program improves the mobility of participants and would provide evidence to support improvements in public education aimed to benefit lower-income families if the benefits outweigh the costs of the program.

Econometric Methods

In my regression, I combine the empirical Head Start research done by Deming (2009) with Grawe’s (2010) model on intergenerational mobility. Grawe interacts school quality with average family income instrumented by state to measure the effect of school quality on the IGE. He includes this state instrument in his model because he uses census data and is unable to connect individual sons with their fathers in the sample. Because using NLSY data allows me to directly link children with their mothers, I do not need to include any instrumental variables in my regression. I use a multivariate linear regression to estimate the log of each child’s earnings as a function of the log of family earnings:

\[ y_{ic} = \beta_0 + \beta_1 y_{ip} + \beta_2 HS_i + \beta_3 A + \beta_4 (y_{ip} \times HS_i) \]

where \( y_c \) is the log of the child’s family income, \( y_p \) is the log of family income for the parents of the child, \( HS_i \) is a dummy variable indicating whether the child participated in Head Start, \( A \) is a vector of child and parent ages that I discuss later, and \( i \) indexes the child. The parameters of interest in this model are \( \beta_1 \), the predicted intergenerational elasticity between parents and children, and \( \beta_4 \), the effect on the intergenerational elasticity if a child participates in Head Start. I hypothesize Head Start participation will increase mobility in my sample. If this is true, then I will observe a negative coefficient \( \beta_4 \) in my regression, implying participation in Head Start decreases the IGE and thus increases intergenerational mobility.

Garces et al. (2002) and Deming (2009) include a control for a child’s participation in preschool in their regression. This is because they want to separate the effects of Head Start from other early childhood interventions and compare Head Start to these other interventions. I take a similar approach in my regression, estimating the IGE for Head Start participants, preschool participants, and children with no early childhood intervention.

In a separate regression, I also estimate all of the variables in the above equation with sibling fixed effects. The coefficients on the independent variables in my OLS regression could be capturing other family effects that my model does not specify. The fixed effects model controls for constant characteristics in households, which could include factors such as parental education, the number of members in the household, and other unobserved family background traits. This method does come with limitations. Namely, this method more than halves my sample size because I only include children who have siblings in the sample.

Data

All of my data come from the National Longitudinal Survey (NLS) Investigator in the National Longitudinal Survey of Youth (NLSY79) and Children of the National Longitudinal Survey.
In their research on levels of intergenerational mobility in different geographic locations in the United States, Chetty et al. (2014) use family income observations. I follow their approach, using family income observations for my intergenerational elasticity estimates. I adjust child family income observations in 2009, 2011, and 2013 for inflation with a base year of 2013. Following Lee and Solon (2009), I then average all income observations greater than zero and log their average, excluding any parent or child with a reported income of zero for all three years. I apply the same transformations to parent family income in 1993, 1995, and 1997. A small number of parent income observations include out of range codes close to one million dollars for their 1995 income. I exclude these observations from my data, and this did not change the results. In every survey year, the NLSY top codes some share of the top income observations. I choose to include these observations in my sample. The age variables in my data are from the last year of income responses: For parents, this is 1997, and for children, this is 2013.

The CNLSY asks questions about a child’s Head Start and preschool participation every two years. I take the responses for each child across all years, and if a response ever reports a child attending either Head Start or preschool, then I code the associated variable as “yes.” For scenarios when the response is “yes” for both Head Start and preschool, I recode the preschool variable as “no” and the Head Start variable as “yes.” It is feasible that respondents think Head Start is a form of preschool, so they would respond “yes” to both questions—this is the assumption that I make when creating my variables. My non-categorical variable summary statistics are shown below in Table 1 and my categorical variable summary is shown in Table 2:

**Table 1: Summary Statistics**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>N</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Child Avg Income</td>
<td>1.751</td>
<td>10.799</td>
<td>1.012</td>
<td>5.298</td>
<td>12.724</td>
</tr>
<tr>
<td>Log Parent Avg Income</td>
<td>1.751</td>
<td>10.748</td>
<td>0.803</td>
<td>6.682</td>
<td>12.779</td>
</tr>
<tr>
<td>Child Age</td>
<td>1.751</td>
<td>26.187</td>
<td>3.107</td>
<td>22</td>
<td>33</td>
</tr>
<tr>
<td>Parent Age</td>
<td>1.751</td>
<td>35.605</td>
<td>2.195</td>
<td>32</td>
<td>40</td>
</tr>
</tbody>
</table>

**Table 2: Early Childhood Education**

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head Start</td>
<td>366</td>
<td>1,403</td>
</tr>
<tr>
<td>Preschool</td>
<td>229</td>
<td>1,540</td>
</tr>
</tbody>
</table>

I also run a fixed effects regression using a subset of my original sample. It consists only of CNLSY participants with siblings. This leads to sample attrition for two reasons: Some children in my sample did not have any siblings and some children may have had siblings, but their siblings did not meet my income criteria. In either case, these children are left out of my sample. The summary statistics for my fixed effects sample are shown in Tables 3 and 4:
Table 3: Fixed Effects Summary Statistics

<table>
<thead>
<tr>
<th>Statistic</th>
<th>N</th>
<th>Mean</th>
<th>St. Dev</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Child Avg Income</td>
<td>741</td>
<td>10.763</td>
<td>0.988</td>
<td>5.298</td>
<td>12.724</td>
</tr>
<tr>
<td>Log Parent Avg Income</td>
<td>741</td>
<td>10.725</td>
<td>0.775</td>
<td>7.016</td>
<td>12.427</td>
</tr>
<tr>
<td>Child Age</td>
<td>741</td>
<td>26.193</td>
<td>2.991</td>
<td>22</td>
<td>33</td>
</tr>
<tr>
<td>Parent Age</td>
<td>741</td>
<td>35.564</td>
<td>2.126</td>
<td>32</td>
<td>40</td>
</tr>
</tbody>
</table>

Table 4: Fixed Effects Early Childhood Education

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head Start</td>
<td>175</td>
<td>566</td>
</tr>
<tr>
<td>Preschool</td>
<td>87</td>
<td>654</td>
</tr>
</tbody>
</table>

Table 5 serves as a clarification for the age ranges at which I observe children and their parents. I first restrict my sample to children from ages 22-33 in 2013. The rationale on the lower bound is to avoid measurement error as much as possible—I do not want to observe any incomes from before a child turned 18. This is because the child could mistakenly report income from a part-time job in total family income, even though he is still a dependent. On the upper bound, I want to avoid having to make IGE comparisons across cohorts like Grawe (2010). While analysis about IGE comparisons between cohorts could be fruitful, it is beyond the scope of my current research, which only focuses on IGE comparisons for a particular age group. This age restriction on children also tightens the range of parent ages, as seen below.

Table 5: Age Range Information

<table>
<thead>
<tr>
<th>Year</th>
<th>Observations</th>
<th>Age_ranges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child Generation</td>
<td>2009 - 2013</td>
<td>22-33</td>
</tr>
</tbody>
</table>

Possible Econometric Issues

Even with a solid theoretical base for my regressions, endogeneity and multicollinearity pose a threat to my empirical results. In an ideal data set, I would have access to every individual’s permanent income. However, this is unrealistic because I only average income observations for three years. This leads to life cycle bias, which occurs when the true variance in a population’s lifetime income is either understated or overstated. Figure 2 depicts theoretical earnings profiles for three individuals. Assume age A* is truly representative of an individual’s lifetime earnings. The variance in earnings between individuals at age A1 is much smaller than the variance at A*, and the variance in earnings between individuals at age A3 is much larger than the variance at A*. When we observe income earlier (later) in a person’s life than A*, the predicted mean income of a population understates (overstates) the gap in earnings. For the IGE, this shows that no “true” IGE exists when we observe income at different ages. Instead, the IGE will be different at different ages. Interacting parent income with parent age allows the IGE to vary depending on the parent’s age, controlling for parent earnings at different stages in life. In separate robustness checks, I find that including these terms in my regression does not meaningfully change my results.
Figure 2: Graphical Depiction of Life Cycle Bias

Trends in earnings follow a quadratic path over time. At young ages, individuals have lower earnings because they do not have experience in the workforce. As they gain experience, their income grows. This initial growth is faster when an individual’s lifetime earnings is higher. At some point, people approach retirement age where they work less and therefore earn less than they did at their prime age of earnings. As a result, I include linear and quadratic age parameters for children and parents.

While I can control for many endogeneity issues such as life cycle bias, I will be unable to control for two primary sources of bias: measurement error in my independent variables and top coding in the NLSY. If parents state that their child participated in Head Start due to recall error, but the child participated in a different preschool program, this would move the estimate of $\beta_4$ toward 0. Measurement error will increase the noise of the data and therefore reduce the effect of any relationship that may exist between Head Start participation and income. In addition to misstatement of Head Start participation, classical measurement error on parent income biases the estimate of $\beta_1$ and $\beta_4$ downward.  

Another source of bias in one of my independent variables is top coding in the NLSY. To protect the identities of individuals with high incomes, the NLSY does not provide the reported incomes of individuals with unusually high income values. In 1991 and 1993, the NLSY79 aggregates all values above a cutoff of $100,000, averages them, and replaces them with this average. In 1995, a similar approach is used, but this time with the top two percent of respondents with valid income values. Both cases bias upward the coefficient on parent family income.

In addition to endogeneity, multicollinearity may be present in my regressions. The Head Start program is primarily composed of individuals at the bottom of the income distribution, so parent income could be correlated with Head Start participation. To check for this potential multicollinearity, I analyze the variance inflation factors and correlation matrices of these variables, finding that multicollinearity does not pose a large threat to my model.

IV. Results and Discussion

Table 6 shows results from my OLS, weighted OLS, and fixed effects regressions. Across all three regressions, the IGE is between 0.53 and 0.62. This is in line with other IGE estimates, which range from 0.4 to 0.6. Even though the IGES from my regressions are slightly higher than the approximate IGE of 0.5 in the United States, they would most likely be closer to Corak’s approximation if parent income were not top coded in my sample. The decrease in the IGE from 0.576 in the OLS model to 0.533 in the weighted OLS model suggests that some earnings
persistence was due to the nature of the NLSY. The survey overrepresents minority and low-income populations, so adding weights mutes these effects.

Table 6: Effects of Head Start and Preschool on IGE

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>OLS (1)</th>
<th>Weighted OLS (2)</th>
<th>Fixed Effects (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>3.467</td>
<td>3.442</td>
<td>-7.472</td>
</tr>
<tr>
<td></td>
<td>(6.072)</td>
<td>(5.923)</td>
<td>(10.058)</td>
</tr>
<tr>
<td>Log Parent Avg Income</td>
<td>0.576***</td>
<td>0.533***</td>
<td>0.617***</td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td>(0.033)</td>
<td>(0.054)</td>
</tr>
<tr>
<td>Child Age</td>
<td>-0.207</td>
<td>-0.345***</td>
<td>-0.161</td>
</tr>
<tr>
<td></td>
<td>(0.126)</td>
<td>(0.124)</td>
<td>(0.170)</td>
</tr>
<tr>
<td>Child Age Squared</td>
<td>0.003</td>
<td>0.005**</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Parent Age</td>
<td>0.231</td>
<td>0.378</td>
<td>0.765</td>
</tr>
<tr>
<td></td>
<td>(0.327)</td>
<td>(0.318)</td>
<td>(0.550)</td>
</tr>
<tr>
<td>Parent Age Squared</td>
<td>-0.003</td>
<td>-0.005</td>
<td>-0.010</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.004)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Head Start</td>
<td>2.796***</td>
<td>2.102**</td>
<td>1.884</td>
</tr>
<tr>
<td></td>
<td>(0.752)</td>
<td>(0.912)</td>
<td>(1.265)</td>
</tr>
<tr>
<td>Preschool</td>
<td>0.379</td>
<td>0.042</td>
<td>-0.346</td>
</tr>
<tr>
<td></td>
<td>(1.053)</td>
<td>(1.003)</td>
<td>(1.525)</td>
</tr>
<tr>
<td>Log Parent Avg Income*Head Start</td>
<td>-0.281***</td>
<td>-0.217**</td>
<td>-0.197</td>
</tr>
<tr>
<td></td>
<td>(0.072)</td>
<td>(0.087)</td>
<td>(0.121)</td>
</tr>
<tr>
<td>Log Parent Avg Income*Preschool</td>
<td>-0.028</td>
<td>0.009</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>(0.096)</td>
<td>(0.090)</td>
<td>(0.139)</td>
</tr>
</tbody>
</table>

Observations | 1,751 | 1,751 | 741
Adjusted R²   | 0.222 | 0.233 | 0.233
F Statistic (df = 9; 1741) | 56.361*** | 60.091***

Note: *p<0.1; **p<0.05; ***p<0.01

The coefficient on the interaction between a student’s participation in Head Start and parent income, Log Parent Avg Income*Head Start, is statistically significant in both the OLS and Weighted OLS estimates. For my sample, this implies that children who participate in Head Start...
have a lower IGE than children who do not participate in Head Start. Holding all other variables constant, we find the IGE for a Head Start participant by adding the coefficient on the interaction between Head Start and parent income to the coefficient on Log Parent Avg Income. This results in a predicted IGE of 0.316 for Head Start participants in the weighted OLS model compared to an IGE of 0.533 for children who do not participate in any pre-kindergarten program.

This result matches my hypothesis that participation in Head Start results in a smaller IGE and thus increases mobility. As mentioned by Corak (2013), we must be careful in assigning a particular direction to this economic mobility. The IGE only provides an average measure of the degree of mobility in a given sample and does not say anything about mobility for members of different income brackets. However, because families of Head Start participants are primarily at the bottom of the income distribution, much of the increase in mobility from participation in Head Start could plausibly come from this group. Using this interpretation, my intergenerational mobility model predicts children who participate in Head Start experience more mobility than children who do not participate in any early childhood intervention program. Care must be taken when interpreting these results because all coefficients that include parent income may have upward bias due to top coding. If bias were present, then the IGE effects that I find may not be as strong.

While these results are promising, when I add additional family controls through fixed effects, the coefficient on the interaction between Head Start and a parent’s log average income loses its significance. One possible explanation is that my fixed effects sample size is much smaller than my OLS sample size. This could cause an increase in standard errors that then make the coefficient lose its significance. Another explanation is that the family controls muted the effect of Head Start on intergenerational mobility. Head Start participants could share similar unobserved family characteristics that are also associated with the Head Start program. These family characteristics could be the reason that higher mobility exists for Head Start participants instead of the program itself.

Not surprisingly, the coefficient on the interaction between a student’s participation in preschool and parent income is not statistically significant. Returns to human capital investment and credit constraints have competing effects on this coefficient. Similar to Head Start participants, preschool participants receive positive returns to education. According to Becker and Tomes’s (1979) theory on human capital accumulation, these positive returns lead to higher future earnings. Parents will send their child to preschool if the investment in preschool will pay off for the child. This is more likely to occur if a child’s return on investment in human capital, his ability, is higher. A scenario of optimal investment could happen if government offices such as the Administration for Children & Families provide subsidies or tax credits for low-income families. This could then encourage investment and create a situation where no family has credit constraints. Children with the highest abilities would then attend early education programs regardless of their family’s income. This scenario would imply the interaction between preschool and parent income has a negative coefficient. Depending on the randomly endowed abilities of children, they will move either up or down in the income distribution depending on that ability, not the income of their parents.

However, as discussed in Section II, if government assistance is not enough, then credit constraints force some families to provide suboptimal investment in a child’s education. If only families at the top of the income distribution can afford to send their children to preschool, then only the children at the top of the income distribution will receive human capital investments. Therefore, mobility will be lower than if no credit constraints exist. Children at the lower end of
the income distribution will not get the necessary investments in education to earn more when they are adults. This would imply the interaction between preschool and parent income has a positive coefficient. Because this coefficient is not statistically different from 0 in my results, I cannot say which of the two effects dominates.

The overall effects of preschool and Head Start on a child’s future income are also notable. While we would expect preschool to have a strong positive effect on child earnings later on in life, the coefficient on this variable is not statistically significant in all three models. Therefore, children who participate in preschool have similar predicted earnings to children who do not participate in preschool. This is a strange result—if we follow Johnson and Jackson’s (2019) theory of dynamic complementarity, then children who participate in early childhood intervention should have higher predicted earnings than those who do not. One potential explanation for this oddity is that different preschools have varying school quality. If some children who attend preschool do not accumulate very much human capital, then they will also have lower future earnings. We would then predict this through small (and sometimes negative) coefficients on preschool in our regression results.

Unlike with the coefficient on preschool, the coefficient on Head Start is statistically significant in both of my OLS regressions and is positive for all child income observations one standard deviation from the predicted coefficient in my fixed effect regression. This implies that, holding all other variables constant, children who participate in Head Start have higher predicted earnings than children who do not participate in Head Start. These higher predicted earnings suggest that the Head Start program has a long-term positive effect on its participants. Along with my result about increased mobility for Head Start participants, this suggests that federal early intervention programs such as Head Start provide opportunities for children who come from low-income families to have higher long-term earnings, ultimately escaping the poverty trap.

One drawback of using the IGE as my primary mobility measure is that I cannot compare the IGE on any population subgroup to the entire distribution. Bhattacharya and Mazumder (2011) explain that not many studies examine racial differences in intergenerational mobility. For example, if they find the IGE of a sample of a Black population, they would not be able to compare that IGE to the IGE of the entire population distribution because it only describes the rate at which earnings of Black children regress to the mean earnings of Black parents.

Similarly, I cannot compare the IGE of Head Start participants to the overall IGE of my sample. However, I can compare the IGE to non-Head Start participants and preschool participants. In addition, Head Start is a choice for families rather than an inherent characteristic, such as race. Therefore, it is still valuable to look at the differences in mobility between Head Start and other forms of childhood education.

Figure 3 provides a visualization of the relationship between parent and child earnings. As parent earnings start to increase, child earnings also increase for both Head Start and preschool participants. The coefficient on log parent earnings reflects this fact. However, for Head Start participants, this increase happens at a slower rate because of the negative coefficient on the interaction between Head Start and parent income. A positive coefficient on the Head Start dummy variable implies that the intercept of child earnings for participants of Head Start is higher than the intercept of child earnings for preschool participants. We therefore observe some point y* where a child’s earnings for Head Start and preschool participants is equal. This represents the point at which it no longer pays off for parents to send their child to Head Start and should instead choose to send their child to preschool.
 Checks for Robustness

In addition to the regression results reported here, I also perform regressions that include interactions between parent age and family income to control for potential life cycle bias of parent income. After calculating the IGE for the mean age of parents and children, my results are similar to those reported above. Tables 7-9 show the results of several tests for covariance and inclusion. The Variance Inflation Factors for each variable detail the percentage that the variance is inflated for each coefficient compared to if there was no multicollinearity. Normally, an inflation factor above five is cause for concern, but I include dummy variables and squared terms in my model. Because the inflation factor on parent log income is low and the rest of my variables are theoretically important, I choose to keep all of these variables. I also perform an F-test of inclusion for my Head Start variable, with the results of this test shown in Table 8. Because the p-value on Head Start and the interaction between Head Start and parent income is significant at the one percent level, I choose to keep these variables. No variable has a high covariance with other variables in Table 9, so from these three tests, I conclude multicollinearity is not a problem in my regressions. I also perform a Shapiro-Wilk test for normality and a Breusch-Pagan test for heteroscedasticity. While the Shapiro-Wilk test finds that my residuals are not normal, and this could be a result of the top-coding of high incomes. Additionally, all variables in my model are theoretically significant, so I choose to keep them. After performing the Breusch-Pagan test, I conclude that heteroscedasticity is not a major concern in my model.
Extensions

While my analysis on the IGE of Head Start participants is a strong foundation for looking at differences in intergenerational mobility, a few extensions of my work could provide a more complete picture of my results. For example, I control for family effects in my regression using sibling fixed effects, but my regressions do not specify what family characteristics affect my results. Additional research could include regressions with variables for inherent family characteristics, such as whether or not a child’s father lives in the household or how many earners live in the household. Finding variables that control for these family characteristics would provide more precise estimates of the coefficients in my model because they would keep the same sample size as my OLS regressions, unlike my fixed effects model.

The IGE is useful for measuring the overall mobility of a given population, which is why I use it as my primary mobility predictor. However, upward mobility is another useful measure that calculates the probability an individual will surpass his parent in income. Researchers
display these results in a rank mobility matrix detailing probabilities of mobility across the income distribution. These matrices also allow researchers to compare population subgroups to the entire population. As previously mentioned, this comparison is not possible using the IGE because splitting up a sample will only show regression to the subgroup mean, not the population mean. Bhattacharya and Mazumder advocate for the use of upward mobility because it allows computations of intergenerational mobility at the margin. They find that for children with parents at the bottom quartile of the income distribution, white children are more likely to move to a higher income bracket than Black children. An extension of my work on the mobility of Head Start participants could use upward mobility to assess the probability of upward mobility for children of families at specific income brackets. This would give a more complete picture of how mobility changes across the parent’s income distribution as a result of Head Start.

While my research details some of the potential benefits of Head Start, I never assess the program using a cost-benefit framework. Deming (2009) notes that a full analysis of the Head Start program would look at the future earnings of Head Start participants and compare them to the costs of the program. Future research could further investigate Head Start’s overall effects using my framework and compare them to the average cost of the program per child.

V. Conclusion

Early childhood intervention programs represent promising opportunities to improve the outcomes of children who would not otherwise be able to learn at a young age. Head Start in particular is a large, national program that emphasizes a holistic education, including traditional learning activities as well as meal programs, health programs, and parental involvement. With more participants reaching adulthood, continued research can assess the program’s benefits on participant earnings. With more research about Head Start’s effects, knowledge of the benefits of early childhood intervention programs and public investments in education will improve.
Endnotes


7 Ibid.


10 Ibid.

11 Ibid.

12 Ibid.

13 Ibid.

14 Ibid.

15 Ibid.

16 Ibid.


18 Ibid.


20 Ibid.

21 Ibid.

22 Ibid.

23 Ibid.

24 Ibid.

25 Ibid.

26 Ibid.

27 Ibid.


29 This assumption comes from the cultural transmission model (Bisin and Verdier 2001). Children become educated either because of family education or because of high quality school education. Depending on which effect is dominant, this leads to either complementarity or substitutability between investment and school quality. Arenas and Hindriks assume complementarity in their main paper and provide a proof that their results also hold when substitutability exists.


36 Ibid.
37 Ibid.
39 Ibid.
40 Transfer payments are examples of non-productive government activities. Becker and Tomes (1979) present a model that shows inequality can still persist even with government transfers.
44 Ibid.
45 Ibid.
46 Ibid.
47 Ibid.
48 Ibid.
50 Ibid.
51 Ibid.
53 Ibid.
54 Ibid.
57 Ibid.
59 Ibid.
60 Ibid.
61 Ibid.
62 Ibid.
63 Ibid.
64 Ibid.


Data available at https://www.nlsinfo.org/investigator/pages/login


Found on NLS information page “Income | National Longitudinal Survey” at https://nlsinfo.org/content/cohorts/nlsy79/topical-guide/income/income


