NBER WORKING PAPER SERIES

# PRESCHOOL LOTTERY ADMISSIONS AND ITS EFFECTS ON LONG-RUN EARNINGS AND OUTCOMES

Randall Akee
Leah R. Clark

Working Paper 32570
http://www.nber.org/papers/w32570

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2024

No funding sources or financial relationships to declare. Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau or the National Bureau of Economic Research. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release. DRB Approval Numbers: CBDRB-FY2022-CES010-022; CBDRB-FY2023-CES005-008; CBDRB-FY24-0038; CBDRB-FY24-0208.

Preschool Lottery Admissions and Its Effects on Long-Run Earnings and Outcomes
Randall Akee and Leah R. Clark
NBER Working Paper No. 32570
June 2024
JEL No. I20,I21,I24,I26,J31

## ABSTRACT

We use an admissions lottery to estimate the effect of a non-means tested preschool program on students' long-run earnings, employment, family income, household formation, and geographic mobility. We observe long-run outcomes by linking both admitted and non-admitted individuals to confidential administrative data including tax records. Funding for this preschool program comes from an Indigenous organization, which grants Indigenous students admissions preference and free tuition. We find treated children have between 5 to 6 percent higher earnings as young adults. The results are quite large for young women, especially those from the lower half of the initial parental household income distribution. There is also some evidence that children, regardless of gender, from households with below median parental incomes realize the largest average increases in earnings in adulthood. Finally, we find that increased earnings start at ages 21 and older for the treated students. Likely mechanisms include high-quality teachers and curriculum.

Randall Akee
Department of Public Policy
University of California, Los Angeles
3250 Public Policy Building
Los Angeles, CA 90095
and NBER
rakee@ucla.edu

Leah R. Clark
U.S. Census Bureau
809 S 10th St
Lafayette, IN 47905
leah.lrc@gmail.com

# 1 Introduction

Despite growing demand for universal preschool,[1] little evidence exists on the effects of preschool on adult labor market outcomes—one of the primary targets of American education. Assessing the impact of preschool programs is often difficult due to the small size of treated cohorts, non-random implementation and take-up, and the lack of detailed data on initial familial conditions (Weiland et al., 2023). When preschools are operated by state and local educational agencies, it is possible to follow students throughout their formal education, but it is much harder to observe employment and earnings outcomes. While a few studies have shown that preschool programs far exceed their investment costs (Bartik and Hershbein, 2017; Cascio, 2017; Heckman et al., 2010), other studies have shown that the cognitive gains of preschool fade out (Bailey et al., 2020; Currie and Thomas, 2000; Kline and Walters, 2016; Weiland et al., 2020). Recent research of Boston's universal preschool program finds that, despite having no effect on test scores in high school, preschool attendance boosts high school graduation and college attendance (Gray-Lobe et al., 2023). If preschool improves college-going it should also improve employment and wages, yet no studies in the U.S. setting have evaluated these outcomes.

Our study overcomes many challenges to research on the long-term effects of preschool by leveraging an admissions lottery for an oversubscribed preschool program, and by using linked confidential administrative data for the children in adulthood. We examine long-run outcomes for children who applied to a high-quality state-wide preschool program over five entry cohorts. Our analysis links admitted and non-admitted individuals from the preschool program to their adult tax records. As a result, we are able to directly examine the program's effects on individual earnings, family income, employment, household formation, and state of residence in young adulthood.

The school system is operated by an Indigenous organization and provides dozens of classrooms, teachers and instructors aides throughout the state free of charge for Indigenous students and some non-Indigenous students. Children from a range of economic backgrounds apply for the program. The demand for these preschool services exceeds the total number of spaces available in every year the program has operated. Thus, the school system instituted a lottery system of admissions for the preschool programs within the districts throughout the state.

We instrument for attending the preschool program with the lottery offer of admissions in order to causally identify the effect of this preschool program on the long-run employment and earnings outcomes of the treated children. Our results indicate that the preschool program increases average annual earnings by 5.1 to 5.4 percent (or 5.03 to 5.33 log points). Subgroup analyses indicate that females that entered the preschool program realize the largest gains in earnings in young adulthood, on average. This suggests that preschool attendance may reduce earnings inequality in adulthood. There is also some evidence that children, regardless of gender, from households with below median parental incomes realize the largest average increases in earnings in adulthood. We also find evidence that the increased average earnings begins at ages

[1] A call for universal and free preschool was included in recent presidential administration goals for all three and four year olds in the U.S. See https://www.whitehouse.gov/briefing-room/statements-releases/2021/04/28/fact-sheet-the-american-families-plan/

21 and older for the treated children.

We examine other adult outcomes (for the treated and non-treated children) derived from tax return records, including state of residence, marital status, and whether individuals claim dependent children. We find no effects on these household formation outcomes in the combined analysis, but the subgroup analyses suggest that treated women, in particular, are delaying marriage and remaining in-state. For Indigenous Peoples, the ability to reside in their home territories and homelands as adults has often been difficult due to the lack of employment and/or the relatively high cost of housing in many cases; thus, this finding is unique in the literature and indicates a potentially positive outcome given the accompanying increase in individual earnings.

Many existing studies focus on the effect of an additional year of early childhood education on students' school readiness through socialization and cognitive gains. Some of the strongest research findings come from the Perry Preschool Program and establish the importance of preschool on cognitive and non-cognitive (personality traits) skills for disadvantaged children (Heckman et al., 2013, 2010). Studies of city- and state-administered programs generally find gains in cognitive skills, which sometimes persist into elementary and middle school, but sometimes fade out (Bailey et al., 2020; Cascio, 2017; Cascio and Schanzenbach, 2013; Currie and Thomas, 2000; Dodge et al., 2017; Gormley Jr et al., 2005, 2018; Harden et al., 2023; Kline and Walters, 2016; Lipsey et al., 2018; Weiland et al., 2020).

As a result, subsequent research has examined longer-term outcomes such as high school graduation, college attendance and behavioral outcomes. A study in Montreal, Canada found that preschool programs result in improvement in non-cognitive skills that persist throughout the life-course resulting in more employment, higher earnings, and reduced criminality for the treated boys (Algan et al., 2022). Rossin-Slater and Wüst (2020) examines the long-run effect of Danish preschool programs and find that targeted programs increase children's educational attainment, lifespan and earnings in adulthood. Reexamining the Boston Preschool program, Gray-Lobe et al. (2023) identify that individuals who attend the preschool program are more likely to graduate high school and attend college, and experience a lower probability of incarceration for boys, in particular. Generally, these research efforts identify several positive long-run outcomes and suggest that even when fadeout occurs during middle school or high school years, the positive effects of early childhood education may rebound in young adulthood.

These preschool programs have the biggest potential benefit in disadvantaged communities, even when they are universal in nature (Bartik and Hershbein, 2017). Programs targeting low-income children, including the Perry Preschool program, show gains in health-supporting behaviors in young adulthood (Campbell et al., 2014). Recent research examining the long-run effects of the Perry Preschool program has shown that there is an intergenerational effect on the children of the treated individuals; their children are in better health, more likely to be employed and experience less criminality (García et al., 2021). The roll out of Head Start improved outcomes for children, including educational attainment and economic outcomes; more recent work finds positive spillover effects for the subsequent generation of children as well (Bailey et al., 2021; Barr and Gibbs, 2022; Ludwig and Miller, 2007). The program we analyze offers a bridge between the universal and targeted progams, since it focuses on Indigenous children, but requires no means-testing.

Our first contribution to the literature is that we provide a precisely estimated coefficient for the return to education for this year of early childhood education. Prior research for the U.S. does not have well-measured earnings or income data for treated populations; outcomes have focused on educational achievement, criminality, and postsecondary educational attainment. Additionally, we improve upon measures of socioeconomic status in childhood relative to existing work, which typically relies on proxy measures such as free or reduced price lunch eligibility. We observe tax-reported family income for program applicants during their childhood years. We find that earnings are higher, on average by about five percent for individuals who were awarded lottery admission to the preschool program. In fact, the results are largest for the girls and the effect on their earnings approaches 9 percent. Previous researchers have found some evidence that girls have long-term educational attainment gains from preschool (Anderson, 2008). However, Deming (2009) finds that boys in Head Start programs gain the most in achievement and Gray-Lobe et al. (2023) find larger college enrollment effects for boys. In a meta analysis, Magnuson et al. (2016) finds very little differences across gender in a study of over 100 studies published between 1960 and 2007; however, they do note that for the handful of long-run (into adulthood) studies the effect sizes for women are large but not statistically significant.

Our final contribution is the extension of research to an understudied population - Indigenous children in the U.S. While there are broadly similar socioeconomic characteristics among Indigenous Peoples and other minority race and ethnic groups in the U.S., these groups also have a long and distinct political relationship and identity within this country. As a result, it is not immediately clear that the same types of interventions will function similarly in these populations (Faircloth, 2015; Romero-Little, 2010; Yazzie-Mintz, 2011). However, the incentives for young adults to attend college and find steady employment are the same for this population as any other in the U.S.

The rest of the paper is organized as follows. Section 2 provides an overview of the preschool program setting. We then, in Section 3, describe the data set and the sample selection. In Section 4 we detail the empirical strategy. In Section 5 we provide the main results and heterogeneity analysis and we discuss likely mechanisms for our observed results in Section 6. We conclude in Section 7.

## 2 Description of the Preschool Program

The particular setting is fairly unique: the program under study provides free preschool for a particular community of Indigenous children at age 4 who reside throughout the state.[2] However, as in the case of other programs, there is higher demand for places in these preschools than can be accommodated and as a result admissions are determined by a lottery annually. There are approximately a dozen districts that operate these schools and the lottery admission is conducted within each of those districts. The program was quite strict in that households had to show their residency within the local region for the district site that they applied for;

[2] Note that under our data use agreement with this educational institution, we have agreed to keep identifying information confidential.

proof of residence was required by showing a utility bill with the parents' names and address.

The preschool program began in the mid 1990's and has continued ever since with a new incoming cohort of students each year. As such, the youngest admitted cohorts are 26-30 in 2022 (our last year of wage data) and it is possible to examine their adult earnings, employment and household formation outcomes.

The preschool program was open to all Indigenous residents in the community and it was not restricted to low income households or other family characteristics. Non-Indigenous children were eligible to submit their application to the preschool program but preference was given to Indigenous children. In a few cases, the leased spaces for these preschool programs were part of the public school system and non-Indigenous children attended those schools as well. These programs were a year in length.

The preschool program required all teachers to have at least a Bachelor's degree in early childhood education; the teaching assistants (there was one in each classroom) were required to have a child development associate's degree at a minimum. The teachers were also required to take 16 hours of professional development training each year as well. In later years, the program reimbursed part of the teacher's tuition in pursuing their Masters degrees. The data also indicate that on average classroom sizes were kept relatively small ranging between 19 and 20 students per classroom over all five years in our data. Additionally, the students were often, though not always, taught by Indigenous teachers and helpers in the classroom.

In this state, when the preschool program began there were very few preschool programs available for this Indigenous population. Private preschools existed, but they were often out of reach for many, but not all, in this population due to cost. Head Start served less than 4 percent of the state-wide population of preschool-aged children. This preschool program operated along normal school hours from 8 am to approximately 2 pm daily; other options existed but were primarily day care options. Thus, this preschool program provided an opportunity for school preparedness that many of these students would otherwise not have had access to.

# 3 Data Set Description and Sample Selection

## 3.1 Data Linking and Sample Selection

In order to examine the long-run effect of attending a preschool program, we link across multiple administrative datasets. The educational institution provided identifying information on individuals who participated in the preschool admissions lottery, with indicators for which individuals were admitted and, subsequently, enrolled. After receiving these data securely, the U.S. Census Bureau ran them through the Person Information Validation System (PVS), which assigns each individual a Protected Information Key (PIK) using the personally identifiable information provided.[3] To conduct the analysis presented here, we use a version of the preschool administrative data that includes PIKs, but no other identifying information. Other datasets we use – including IRS Form 1040 and W-2 records – similarly undergo PVS and receive PIKs, permitting us to link anonymous individuals across various data sources in the Census Bureau's secure research environment.

[3]For more information on PVS, see Layne et al. (2014).

**Table 1:** Descriptive Statistics for the Full and PIK-Assigned Data

| | Average characteristics | |
|---|---|---|
| | All applicants | PIK'd applicants |
| | (1) | (2) |
| A. Sample description | | |
| Admitted | 0.715 | 0.721 |
| Attended | 0.682 | 0.690 |
| Has PIK | 0.766 | 1.000 |
| B. Applicant demographics | | |
| Female | 0.486 | 0.476 |
| Age (end of application year) | 3.50 | 3.51 |
| | (0.295) | (0.296) |
| C. Parent demographics | | |
| Parent Mean AGI | | 35,600 |
| | | (30,100) |
| Parent Age at Birth | | 29.9 |
| | | (8.28) |
| Married Parent | | 0.623 |
| Parent Born in State | | 0.908 |

Note: Full Sample observations = 6,700; PIK'd Sample observations = 5,100. Standard deviations for the reported means are provided in parentheses; we do not report the standard deviation for binary measures. All numbers have been rounded according to the U.S. Census Bureau Disclosure Review Board specifications for this table. These figures have been approved for disclosure. DRB Approval Number: CBDRB-FY24-0038; CBDRB-FY24-0208.

In Table 1 we show descriptive characteristics based on the raw data and the subsample of individuals assigned a PIK. In the first column of the table, we show the average characteristics for the variables contained in the raw data: whether each child is admitted to the preschool program, attends the program, is assigned a PIK, is female, and child age at the end of the application year. The second column shows the corresponding values for the subset that were assigned a PIK; there are additional variables reported in this column under parent demographics that are derived from the PIK-linked administrative data such as IRS Form 1040 or Census Numident records.

Table 1 reveals that the sample with PIKs is similar in terms of observable characteristics to the full sample. In particular, we find that the proportions of individuals who were admitted to or who attended the preschool program are similar across the two samples. Between 71-72% of individuals were admitted to the preschool program in the data over all years. Approximately 69% of individuals attended the program, indicating a compliance rate of over 95% in both samples. We are able to assign about 77% of our observations a unique PIK; this proportion is similar to match rates found for other young and minority populations in the U.S. The proportion of female applicants in our data is fairly close between the full sample and the PIK-assigned sample ranging between 48-49% of the full data. We also find that the age of the child (by the end of the calendar year) is very similar across the two samples at 3.5 years. Overall, we take this as evidence that the PIK assignment process does not significantly affect the average characteristics of our data.

For the sample with PIKs, we are able to show selected parental characteristics since we are able to link these individuals to various administrative data sources. IRS Form 1040 data provide a particularly useful source of information on individuals' circumstances in childhood, and are available for tax years 1994, 1995, and 1998 forward. Since children can be claimed on tax returns by different adults at different points in time (due to marriage, divorce, and living with grandparents, among other reasons), our first task is to identify the "parent" PIKs most relevant to our analysis.

Specifically, we define individuals' parents as the adults who claim them on a tax return in the same year of the application, or, if no tax return is available that year, the most recent prior year in which they were claimed. Then, for those not claimed prior to applying, we take the adult PIKs from the first tax record on which they were claimed after the application year. If there are two parent PIKs available in the tax record we privilege according to the afforementioned steps, we designate them as having married parents; otherwise, we classify them as having a single parent. This process enables us to identify parental PIKs for virtually all child applicants, so to avoid small changes to sample composition (and related disclosure concern) our primary PIK'd sample is comprised of applicants who both have PIKs themselves and parent PIKs.

Having identified either one or two parent PIKs for each applicant child, we infer parental income based on the adjusted gross income reported in tax filings by these parent PIKs (regardless of whether the applicant child is claimed) during and just prior to the year of application. These years are most indicative of family economic circumstances at the time of preschool application. Specifically, for applicants in 1995, 1996, and 1997, we construct parental mean

**Table 2:** Balancing Table for Admissions Offers by Full and PIK-Assigned Data

| | Admissions Offer Differentials | |
|---|---|---|
| | All applicants | PIK'd applicants |
| | (1) | (2) |
| A. Sample Description | | |
| Has PIK | 0.0177 | N/A |
| | (0.0114) | |
| B. Applicant demographics | | |
| Female | 0.001 | 0.006 |
| | (0.014) | (0.016) |
| Age (end of application year) | 0.021*** | 0.023*** |
| | (0.008) | (0.009) |
| C. Parent demographics | | |
| Parent Mean AGI | | -1,440 |
| | | (939) |
| Log Parent Mean AGI | | -0.0574* |
| | | (0.0305) |
| Parent Age at Birth | | -0.322 |
| | | (0.259) |
| Married Parent | | 0.001 |
| | | (0.015) |
| Parent Born in State | | 0.045*** |
| | | (0.009) |

Note: Full Sample observations = 6,700; PIK'd Sample observations = 5,100. These are estimated regression coefficients where the applicant or parental characteristics were regressed on an admissions indicator variable to test whether the lottery was balanced. Standard errors are shown in parentheses for each estimated coefficient in the table. All numbers have been rounded according to the U.S. Census Bureau Disclosure Review Board specifications for this table. These figures have been approved for disclosure. *** indicates statistical significance at the 1% level, ** indicates statistical significance at the 5% level, * indicates statistical significance at the 10% level. DRB Approval Number: CBDRB-FY24-0038; CBDRB-FY24-0208.

adjusted gross income (AGI) as the average of AGI in 1994 and 1995; for 1998 applicants, we average 1995 and 1998; and for 1999 applicants, we average 1998 and 1999. All dollar amounts are inflation-adjusted to 1999 dollars prior to averaging. If we observe only one record we use that as our parental mean AGI measure; if we observe no records, we use the first observed tax record from 2000 or later.

We measure parental age and state of nativity from Census Numident records, which draw information from Social Security Administration records. These variables have universal coverage in the PIK'd sample by construction: the Numident forms the backbone of the PVS process. If a child is linked to two parents, we take the age of the primary filer parent, and flag them as having a parent born in state if either parent was born in state.

Parental mean AGI is $35,600 in 1999 dollars. The average parental age at birth for the child applying is nearly 30 years old and about 62% of parents are married. Finally, approximately 91% of parents are born in the state in which the preschool program operates.

Table 2 provides the estimated coefficients from regressions of the stated characteristic regressed on whether an individual was offered admissions to the preschool program; these estimated coefficients provide some insight into the "fairness" of the admissions lottery. Overall, a picture emerges that is consistent with our understanding of the program context: the program favored individuals who could document their Indigenous heritage, and this Indigenous community, on average, experiences economic disadvantage.

We first show that having a PIK is not a statistically significant predictor of admission; that is, the process of linking applicants to other administrative data does not systematically favor admitted individuals. Next, we examine characteristics available in both the full and PIK'd sample: gender and age. We find no statistically significant difference in whether female applicants were admitted to the preschool program in either sample; the estimated coefficients on admissions offer are small and statistically insignificant. We do find, however, some evidence that slightly older children may have been favored in the admissions process. In both samples, coefficients on age are positive and statistically significant at the 1% level but small in magnitude – equivalent to about one week difference in age, on average. Overall, these findings reinforce our contention from Table 1 that the sample with PIKs appears observably similar to the overall pool of applicants.

Next, we test whether there the admissions offer differs by parental characteristics for the PIK-assigned sample. We are unable to test this for the full sample as we are only able to link parental characteristics (using administrative records) for observations with a PIK. We find that the level and log mean of parental adjusted gross income is negative for those applicants that were offered admissions. The estimated coefficients are not statistically significant at conventional levels for the level version of the variable, but it is statistically significant at the 10% level for the log version of the variable. These estimated coefficient indicate, on average, that parents of children who were offered admissions to the preschool program had lower AGI than parents of those who were not offered admissions. We control for parental log income in our regression analysis to account for potential differences in income composition of treatment and control groups. We find no evidence of difference in admissions offers based on parental age or marital status.

Finally, we show that for applicants whose parents were born in the state, they are about four percentage points more likely to be offered admissions to the preschool program. We believe that this measure is a proxy for a parent's Native or Indigenous status, for which we have no direct measures. The program implemented a system to determine whether applicants met standards for preferential access based on documentation of ancestry and tribal enrollment status. Since the program was understood to target this population, the vast majority of applicants likely met these standards, but, unfortunately, we have no indicator for whether they did. Notably, this requirement to provide documentation may contribute further to the discrepancy in whether parents were born in state, because Indigenous parents born out of state may have had more difficulty producing the required documents.

Ideally, we would be able to adjust for the Indigeneity determination made by the program, but in the absence of that, we control for whether parents were born in state in all specifications. Note, however, that we have no reason to suspect that an admission process favoring Indigenous children is one that favors children with better long-run outcomes in the absence of the intervention. Table 3 shows that self-identified Indigenous individuals in the state generally have worse economic circumstances, according to 2000 U.S Census data.

**Table 3:** Indigenous Income and Poverty Relative to the Statewide Totals

| Measure | Percent |
|---|---|
| Indigenous Median Household Income as a Percent of Statewide Median HH Income | 91 % |
| Indigenous Per Capita Income as a Percent of Statewide Per Capita Income | 66 % |
| Indigenous Poverty Rate as a Percent of Statewide Poverty Rate | 183 % |

Source: 2000 US Decennial Census Public Use data.

### 3.2 Key Variables

The preschool administrative data were collected in the late 1990s for individuals born between 1992-1996. These individuals turned 18 years old between 2010-2014. Thus, to observe their adult outcomes, we rely on IRS data from 2011-2022. We can only measure these outcomes for individuals with PIKs; thus, all our outcome analyses are restricted to the sample with PIKs.

Table 4 provides the means of the main outcome variables for the admitted and non-admitted children in adulthood. These outcome variables are constructed from two IRS data sources: Forms W-2 and 1040. We assess these outcomes in the years in which individuals turn 19 or older. Form W-2 data is reported by employers to the IRS, while Form 1040 data is reported by the individual tax unit. Sometimes, an individual receives W-2 wages but does not file a tax return, though the vast majority of W-2 recipients in the sample also file a tax return. Informal earnings that are not reported to the IRS cannot be observed. We describe the outcome variables used in the analysis to follow in detail below.

*Wages.* Wages are estimated by totaling the wages and tips in all the W-2 records for an individual in each year. When we observe no W-2 records for an individual within a year, we assume they had no wages. Technically this is an assumption of no informal wages.

Specific wage outcome variables average across multiple years of wages. Our main measure is mean wages across all adult years observed 2011-2019. For the 1992 birth cohort, this includes ages 19-27; for the 1996 birth cohort, this includes ages 19-23. As shown in Table 4, average wages are approximately \$15,900 for the non-admitted children in adulthood while it is about \$16,500 for the admitted children in adulthood. We also run regressions using the log transformation of this variable which effectively assumes zero wages are missing, and leads to the exclusion of approximately 2% of the sample who never have W-2 income in the years observed. We analyze other mean wage and log mean wage outcomes, including ages 21 and above (2013-2019), pandemic years (2020-2022, ages 24-30) and the following age subsets available for all cohorts: ages 19-20, ages 21-23, and ages 24-26 (which includes pandemic years for 1997-1999 cohorts).

*Adjusted Gross Income (AGI).* AGI is measured from the Form 1040 (tax return) record in which the individual appears as a primary or secondary filer in a given year. This reflects an individual's income if not married filing jointly, or a couple's income if married filing jointly. We define mean AGI outcomes for the same year and age configurations as mean wages, except our 1040 data ends in 2021. In Table 4, we report the mean AGI for ages 19 and above (2011-2019), which is about \$20,800 for non-admitted individuals and \$21,500 for the admitted individuals. Not all W-2 wage earners file a 1040, so this outcome loses approximately 6% of individuals.

*Employment.* We designate an individual as employed in a given year if they have any W-2 wages, and assume a lack of W-2 wages indicates non-employment. We note that our analysis can not rule out subsistence and non-market labor activities. Our outcomes average employment indicators for the age/year configurations described for wages. Approximately 84% of the not-admitted and admitted populations are employed on average at ages 19 and above (2011-2019).

*Married.* We designate individuals as married if they file a married tax return (jointly or separately) as the primary or secondary filer. There are incentives to file taxes when married, so we suspect this will capture most – but maybe not all – married individuals in our data. Table 4 documents marital status as of the last 1040 we observe for individuals in our data. Just 19% of non-admitted and 20.5% of admitted individuals are married at that point. We also analyze outcomes for whether an individual is married at any point (2011-2021) or whether an individual appears as married by age 26. We do not separate pandemic years for this outcome.

*Claims children.* We designate individuals as claiming children if they report any child dependents on their tax return. Note that the rate of claiming children exceeds the rate of married filing in both samples. This highlights an issue with this variable, which is an imperfect measure of fertility: for unmarried co-residing parents, only one parent may claim a given child, and for non-custodial parents, a child can only be claimed if they are providing more than one-half of the child's financial support. Thus, there are likely some individuals with children in our data whom we never see claim children, particularly if they are not living with and providing the bulk of financial support for the child.

*Resides in state.* We designate individuals as residing in their home state if they report a mailing address in the same state as the preschool program on their tax return. As with the other Form 1040-derived outcomes, we cannot observe this outcome for never-filers (approximately 6% of our sample). As of the last 1040 observed for each individual, approximately 75% of not-admitted individuals reside in-state and 76.5% of admitted individuals do.

**Table 4:** Table of Means for Outcome Variables for Children in Adulthood by Admission Category

| | Not Admitted Mean | Admitted Mean |
|---|---|---|
| Mean Wages, Age 19+ | 15,900 | 16,500 |
| | (11,300) | (14,200) |
| Mean Wages Excluding Zeros, Age 19+ | 18,000 | 18,600 |
| | (11,300) | (16,700) |
| Mean AGI, Age 19+ | 20,800 | 21,500 |
| | (13,600) | (20,200) |
| Mean Employment, Age 19+ | 0.840 | 0.843 |
| | (0.24) | (0.24) |
| Married ‡ | 0.191 | 0.205 |
| Claims children ‡ | 0.251 | 0.295 |
| Resides in State ‡ | 0.747 | 0.765 |
| N | 1,400 | 3,700 |

Note: ‡˙ indicates the last (most recent) occurrence in the 1040 data. Standard deviations for the reported means are provided in parentheses; we do not report the standard deviation for binary measures. All numbers have been rounded according to the U.S. Census Bureau Disclosure Review Board specifications for this table. These figures have been approved for disclosure. DRB Approval Number: CBDRB-FY24-0038.

# 4 Empirical Strategy

Our causal identification strategy relies on the fact that individual students were admitted to the preschool program based not on their individual or familial characteristics (other than the preference for Indigenous or Native American status) but by a lottery admissions process. We use a two-stage-least-squares approach to estimate the local average treatment effect of program attendance, as well as a reduced-form model to estimate the intent-to-treat effect of program admission. Due to extremely high take-up of the admissions offer, these two treatment effects should be quite similar.

To estimate the local average treatment effect, we use the offer of admissions to instrument for attending the preschool program. In the second stage analysis, we use instrumented attendance to identify the effect of attendance on outcomes including wages, employment, and household formation.

The first equation below indicates that the admission to the preschool program determines attendance in the school. The outcome variable is an indicator variable for whether the child i attended the preschool program. The variable admitted is an indicator variable for whether the child was offered admission to the preschool program. The vector $X$ includes control variables for gender, parental income at or around the time of application, parental marital status, and an indicator for whether parents were born in state. We conduct this via ordinary least squares estimation in a two-stage least squares analysis. The predicted values for attended are then included in the second regression and we estimate the coefficient of interest $\beta$ which explains the effect of attending the preschool program on long-run adult outcomes $Y_i$; these outcomes are wages, total income, marital status, fertility, and residing in the home state. We pool data for the five admitted cohorts for our analysis and report their outcomes in young adulthood.

$$Attended_i = \gamma \times Admitted_i + X_i^{'}\mu + \delta_i + \epsilon_i \tag{1}$$

$$Y_i = \beta \times Attended_i + X_i^{'}\theta + \delta_i + \epsilon_i \tag{2}$$

Our empirical specification also includes a year of application and district fixed effect as well as their interactions and indicated by $\delta_i$.

We use the lottery based admissions to instrument for whether a person attends the preschool program or not. As is standard, we are concerned that individual children who attend these relatively novel preschool programs differ along unobservable family or individual characteristics. The two-stage least squares method and the lottery assignment for admissions should be unrelated to these unobserved characteristics.

The reduced form model is equivalent to Equation 2, except the admitted variable takes the place of the attended variable.

There is a second potential source of selection in our data: whether individuals receive PIKs and, thus, have observable long-run outcomes. To adjust for potential selection into having a PIK, we weight individuals in the reduced form model by the inverse probability they receive a PIK, modeled as a function of their birth year, gender, and residential zip code (the only characteristics we observe for all applicants). We calculate these probabilities by estimating this model via logit. The resulting intent-to-treat effects upweight individuals from districts with lower PIK rates, and assume that individuals from these districts with PIKs are similar to those without PIKs.

# 5 Results

We examine long-run student outcomes via three different methods: a two-stage least squares analysis, ordinary least squares (OLS) analysis, and an OLS analysis using inverse probability weights. The first stage coefficient for the two-stage least squares analysis is shown at the bottom of Table 5 and indicates that winning the preschool lottery (admission) is a strong predictor of a child attending the preschool. Consistent with the implied attendance rate in Table 1, the estimated coefficient of 0.96 indicates almost certain attendance if a child is granted admission. Clearly the program was in high demand. For our main analyses, we show both the IV and reduced-form results, but the findings across our specifications are virtually identical as a result of the very high compliance rate. Thus for heterogeneity analyses, we use the reduced-form specification.

## 5.1 Effects on Income

We first examine the effects of preschool on labor market outcomes for applicants in young adulthood, shown in Table 5. Each estimated coefficient comes from a different regression and the outcome variables are listed in the leftmost column. These regressions all include the following control variables: child's gender, parental income and marital status, whether the parent was born in the state and application year by district fixed-effects.

The first two rows of Table 5 provide the analysis for the mean of wages (at ages 19 and older and at ages 21 and older, respectively) regressed on attending the preschool program. When we average wages earned at ages 19 and above, the estimated effect is around $700 but is not statistically significant at conventional levels, except in the specification that reweights observations by the probability of having a PIK (column 3). The coefficients are larger and significant at the 10% level when we restrict wages to ages 21 and above. Attending the preschool program increases annual wages by more than $900, on average.

In the next two rows, we report a similar analysis for the log mean wages. We find that children who attend the preschool program have approximately 5% higher average earnings over ages 19 and older. These results are consistent in the OLS analysis shown in the next two columns. If we, instead, restrict the outcome variable to examine the effect of preschool attendance on earnings at ages 21 and older, the estimated coefficients are slightly larger in magnitude.

In the fifth row, we examine the highest wage for a single year that individuals experience at age 19 or older, and find that people who attended the preschool program have about a 6% higher maximum wage than a person who did not attend the program preschool in all three regression analyses. Note that these earnings measures are derived from IRS W-2 data.

Next we evaluate effects on adjusted gross income at ages 19 and older or ages 21 and older, respectively. The estimated coefficients in these regressions are positive but quite small in magnitude, and they are not statistically significant. Recall that AGI is impacted by several sources of selection: namely, the choice to file taxes and the choice to get married. If the program causes individuals to slightly delay marriage, for example, then we might be comparing more married couples' combined income in the control group to more single filers in the treatment group, washing out the individual wage gains we observe. Ultimately, AGI might be a better outcome to evaluate when individuals are in their thirties and forties and their career and family formation decisions are more established. Note that these adjusted gross income measures are derived from IRS 1040 data.

In the bottom two rows of Table 5, we examine mean employment over ages 19 or 21. The coefficients are positive but not statistically significant, suggesting that the preschool program does not cause substantial gains in employment on average.

Overall, our results indicate that attending the preschool program results in higher earnings at ages 19 and above and even more so at ages 21 and above. There does not appear to be any overall differences in attending the preschool program on adjusted gross income or on average employment levels.

In Appendix Table A1 we repeat the regression analyses from Table 5 and include the additional years impacted by the COVID-19 pandemic (2020-2022). We show that the both the mean wages and log of mean wages are still positive, but they are smaller in magnitude and they do not attain statistical significance at conventional levels. We also show that there does not appear to be any statistically significant differences in average employment levels for adults who attended the preschool program to those who did not. An interesting question that we lack the data to adequately examine is whether exposure to lock down – due to different occupational or locational environments – varied between the treatment and control groups.

In Table 6, we show the estimated coefficient on being offered admissions to the preschool program on average earnings in adulthood at ages 19-20, 21-23 and 24-26. The analysis shows that the small but not significant wage effect at ages 19-20 transitions into a sizeable wage gain of more than $1,100, on average, at ages 21-23. At ages 24-26, the magnitude of the wage gain remains positive but is no longer statistically significant. This is not clear evidence of fade-out, however, since for three of the five cohorts studied, ages 24-26 coincided with the labor market shocks of the COVID-19 pandemic. Whether wage effects fade-out as individuals age will be better studied in future years with additional data.

Disaggregating employment outcomes out by age reveals positive and statistically significant employment effects of preschool admission at ages 19-20. Admitted individuals are 2.6% more likely to be working in formal employment at these ages than non-admitted individuals. This suggests that the preschool program may have contributed to a stronger start in the labor market for attendees.

**Table 5:** Effects of Preschool Attendance on Labor Market Outcomes, 2011-2019

| | 2SLS | OLS | OLS with IPWs | N |
|---|---|---|---|---|
| | (1) | (2) | (3) | |
| Mean wages, ages 19+ | 748 | 717 | 719* | 5100 |
| | (461) | (444) | (406) | |
| Mean wages, ages 21+ | 965* | 925* | 934* | 5100 |
| | (577) | (556) | (507) | |
| Log mean wages, ages 19+ | 0.053* | 0.051* | 0.050* | 5000 |
| | (0.028) | (0.027) | (0.028) | |
| Log mean wages, ages 21+ | 0.055* | 0.053* | 0.054* | 4900 |
| | (0.031) | (0.030) | (0.030) | |
| Log max wage | 0.063** | 0.060** | 0.060** | 5000 |
| | (0.029) | (0.028) | (0.028) | |
| Log mean AGI, ages 19+ | 0.006 | 0.005 | 0.01 | 4800 |
| | (0.042) | (0.041) | (0.042) | |
| Log mean AGI, ages 21+ | 0.015 | 0.014 | 0.016 | 4600 |
| | (0.051) | (0.049) | (0.049) | |
| Mean employment, ages 19+ | 0.012 | 0.011 | 0.012 | 5100 |
| | (0.008) | (0.008) | (0.008) | |
| Mean employment, ages 21+ | 0.006 | 0.006 | 0.006 | 5100 |
| | (0.009) | (0.008) | (0.009) | |

First stage estimated coefficient

| | Attended | | | N |
|---|---|---|---|---|
| Admitted | 0.959*** | | | 5100 |
| | (0.006) | | | |

Note: All numbers have been rounded according to the U.S. Census Bureau Disclosure Review Board specifications for this table. These figures have been approved for disclosure. DRB Approval Number: CBDRB-FY24-0038; CBDRB-FY24-0208. All specifications include gender, parental characteristics, and application-year-by-district fixed effects. In this analysis, the data series includes and ends in 2019. *** indicates statistical significance at the 1% level, ** indicates statistical significance at the 5% level, * indicates statistical significance at the 10% level.

**Table 6:** Mean Wage and Employment effects by age group, 2011-2022

| | Ages 19-20 | Ages 21-23 | Ages 24-26 |
|---|---|---|---|
| | (1) | (2) | (3) |
| Mean wages, specified ages | 310 | 1,111* | 819 |
| | (272) | (572) | (744) |
| Mean employment, specified ages | 0.026** | 0.012 | 0.003 |
| | (0.012) | (0.009) | (0.010) |

Note: All numbers have been rounded according to the U.S. Census Bureau Disclosure Review Board specifications for this table. These figures have been approved for disclosure. *** indicates statistical significance at the 1% level, ** indicates statistical significance at the 5% level, * indicates statistical significance at the 10% level. DRB Approval Number: CBDRB-FY24-0038.

### 5.2 Household Outcomes for the Children in Adulthood

In Table 7 we show the OLS analysis of being admitted to a preschool program on household characteristics in adulthood.[4] The first row of Table 7 provides the estimated coefficients on being admitted to the preschool program on three different measurements of the child's marital status derived from tax return data from 2011-2022 - at any time age 19 or older (column 1), by age 26 (column 2) or the last year observed in our data (column 3). The effect of being admitted to the preschool program does not appear to have a statistically significant effect on marital status in young adulthood. The next row provides a similar analysis for whether the individual claims any children on their 1040 tax forms. Once again, there does not appear to be a statistically significant effect of being admitted to the preschool program on claiming children in any of the three different points in time (across columns 1-3). Finally, there is little evidence that being admitted to the preschool program results in a differential effect of living in-state in adulthood on average.

**Table 7:** Effect of Being Admitted to the Preschool Program on Household Formation using Tax Return Data, 2011-2021

| | Any time ages 19+<br>(1) | By age 26<br>(2) | Last observed<br>(3) |
|---|---|---|---|
| Married | -0.005 | -0.003 | 0.0003 |
| | (0.014) | (0.014) | (0.014) |
| Claims children | 0.022 | 0.016 | 0.016 |
| | (0.016) | (0.015) | (0.015) |
| Resides in-state | 0.014 | 0.013 | 0.014 |
| | (0.012) | (0.013) | (0.015) |
| N | 4,800 | 4,800 | 4,900 |

Note: All numbers have been rounded according to the U.S. Census Bureau Disclosure Review Board specifications for this table. These figures have been approved for disclosure. DRB Approval Number: CBDRB-FY24-0038. All specifications include gender, parental characteristics, and application-year-by-district fixed effects. Intention to Treat Analysis. *** indicates statistical significance at the 1% level, ** indicates statistical significance at the 5% level, * indicates statistical significance at the 10% level.

### 5.3 Heterogeneity Analysis by Parental Median Earnings and Child Gender

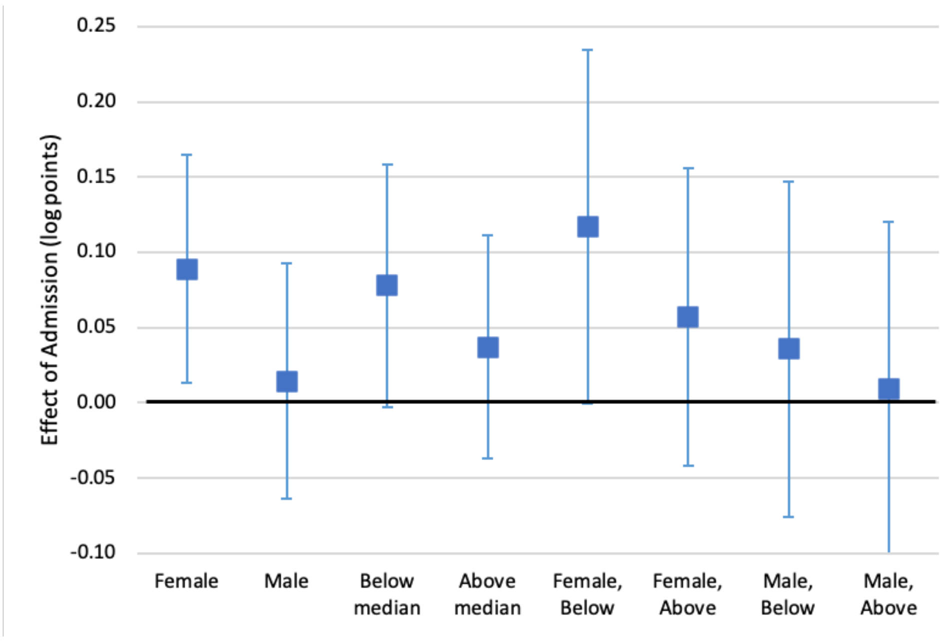
Our primary results indicate that the admissions to the preschool program has a positive and statistically significant effect on the adult earnings of the treated children. There is little evidence for an effect on employment, AGI or other household formation characteristics on average. In this next section, we take advantage of the richness of the data in order to conduct heterogeneity analysis by gender and parent mean AGI (near the time of preschool application). Specifically, we split the sample by gender or whether parent mean AGI is above or below the median, and re-run our main outcome analyses on these split samples. Then, we split the sample by the interaction of these two characteristics, and repeat the analyses. The full set of

[4] We provide the full set of analyses for the two-stage least squares and for the OLS with inverse probability weighting table in Appendix Table A4.

outcomes are provided in Appendix Tables A5 and A6. We provide graphical depictions of the results below.

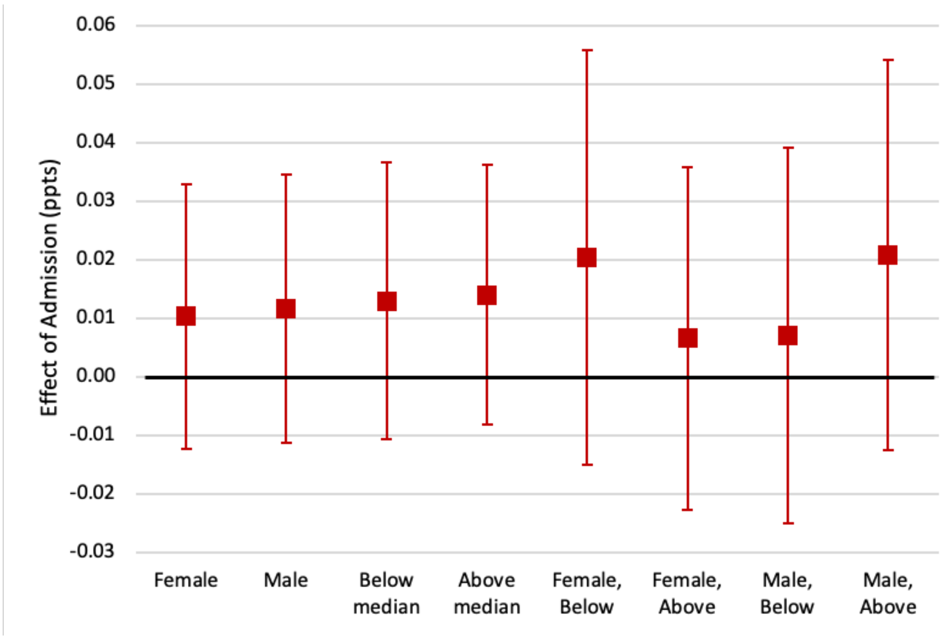
In Figure 1 we chart the estimated coefficients from the OLS regressions of average wages from ages 19 onward on being admitted to the preschool program by different subgroups. The first two estimated coefficients in the graph show the effects for women and men; the estimate is positive and statistically significant for women, while it is close to zero and not statistically significant for men. However, we do fail to reject a null hypothesis that these two estimated coefficients are the same at conventional levels; the p-value is 0.174. The next two estimates suggest that individuals below the median parental income realize larger proportional gains in wages than individuals above; however, we are not able to reject the null hypothesis that these estimated coefficients are the same. The last four estimates show the interaction of gender and above or below median of initial parental income. Women with below-median parental income realize the largest wage gains, statistically significant at the 10% level. However, none of the point estimates are statistically significantly different from one another. It is important to note that we also fail to reject the null hypothesis that the two most extreme coefficient estimates (Female, Below and Male, Above) are statistically significantly different from one another at conventional levels; p-value is 0.188.

**Figure 1:** Wages by Gender and Initial Household Income



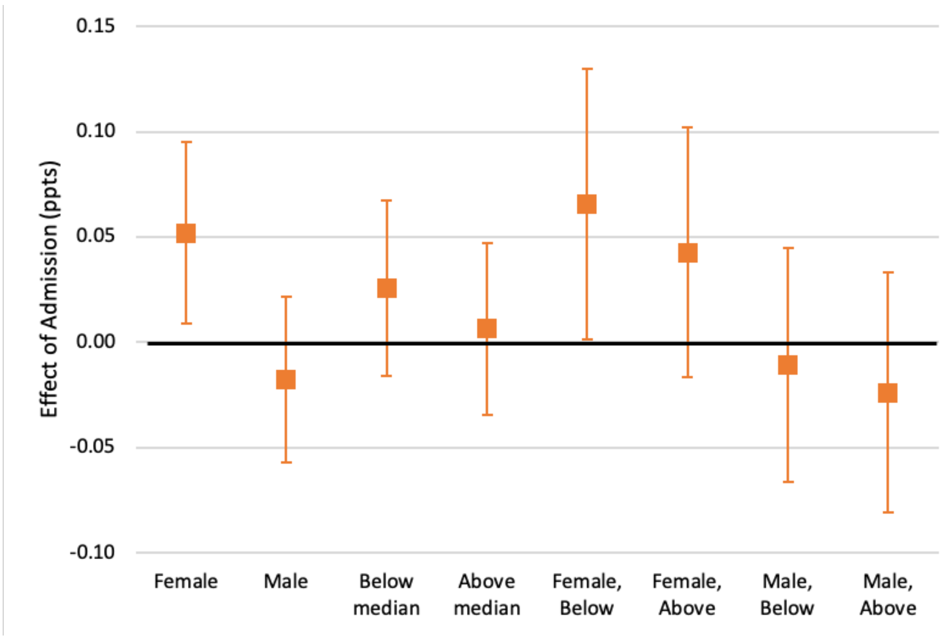

In Figure 2 we show the average employment probability for ages 19 and older for the same set of disaggregated groups from the prior analysis in Figure 1. Overall, the estimated coefficients are positive, however, none of them attain statistical significance at conventional levels. Additionally, none of the estimated coefficients are statistically significantly different from one another either.

**Figure 2:** Employment by Gender and Initial Household Income




In Figure 3 we display the same set of results for a different outcome variable - resides in the same state as the preschool program in adulthood. We show that women, on average, and women from households that were below the initial parental median income are more likely to reside in the home state in adulthood. The estimated coefficients for women and men are statistically different at the 2 percent level; p-value is 0.0169. We do not observe statistically significant differences between individuals from initially below or above median incomes with regard to residing in-state in adulthood. However, there is a statistically significant difference in estimated coefficients between Female, Below and Male, Below and Male, Above. The respective p-values are 0.068 and 0.033.

**Figure 3:** Resides In State by Gender and Initial Household Income



While women, and women from initially low-income households, in particular, are more likely to reside in state, we also see evidence that they are less likely to be married. Figure 4 indicates that treated women, on average, are 4 percentage points less likely to file taxes as a married couple in young adulthood than untreated women; this difference is statistically significant from zero at the 10% level. We also note that this estimated coefficient is different from that of men; the p-value is 0.0196. This suggests that the preschool program may cause young women to delay marriage. As with the in-state outcome, the effects for initially low-income women are statistically significant and statistically distinct from the effects for both low-income and high-income men at the 10% level; the p-values are 0.0613 and 0.0575, respectively.

**Figure 4:** Marital Status on Tax Return by Gender and Initial Household Income

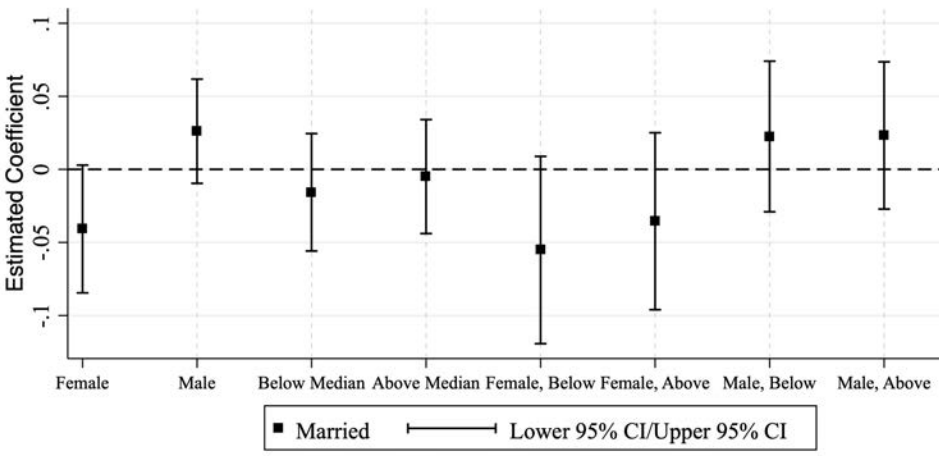
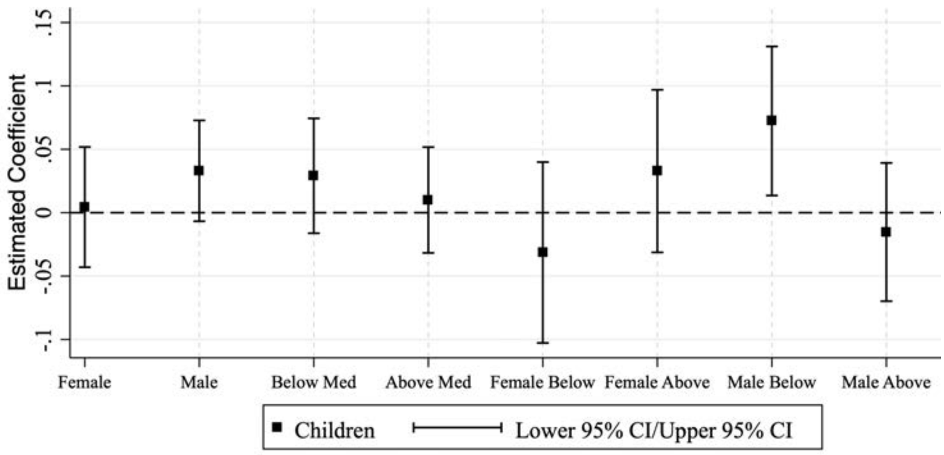



Figure 5 indicates that treated men with below median initial parental income are more likely to claim dependent children in young adulthood; these results are statistically significant at the 5% level. We find little evidence for statistically significant effects for the other subgroups. Note, however, that this is a suboptimal measure of fertility in men, in particular, due to low marriage rates at these young ages and the tendency of children to reside with their mothers. Thus, this could reflect treated low-income men having more children, or it could reflect that they are more likely to be providing financial support and claiming their children. These estimated coefficients for Male, Below are statistically significantly different from Female, Below and Male, Above; p-values 0.0265 and 0.0304, respectively.

**Figure 5:** Claims Child on Tax Return by Gender and Initial Household Income




## 6 Discussion of Potential Mechanisms

While the analysis has identified the long-run effect of preschool attendance on earnings, income, geographic mobility and household formation, it is useful to discuss the potential mechanisms responsible for the observed results. Other researchers have pointed to the impact that early childhood education interventions have on a children's cognitive and non-cognitive skills. Researchers have shown that educational achievement in test scores often fade out in many of the Head Start evaluations (Bailey et al., 2020; Currie and Thomas, 2000; Kline and Walters, 2016). Heckman et al. (2013) points out that improvement in non-cognitive skills (personality traits) may play an important role in student long-term success. In fact, this may help to explain why many of these early childhood education interventions display long-run effects. While the shorter term cognitive skills may moderate over time, the social skills learned early in preschool persist into young adulthood and may play an ever increasing role in educational and labor market success (Deming, 2017). We find that preschool attendance improves earnings, on average, for our treatment population; these effects increase at ages 21-23. Treatment also increases likelihood of employment at ages 19-20. These findings indicate that the students attending the preschool are getting a labor market boost at a critical age, potentially due to gains in both cognitive and non-cognitive skills.

One additional component of skill development that this preschool provided was the inclusion of cultural content and curriculum into the learning environment. Speaking with the teachers, they conveyed that the curriculum and methods of teaching were informed and based upon various aspects of this Indigenous People's history, traditions and language. Often, the teachers and their classroom aides were from the same Indigenous community as the students; this has often been shown to assist in student outcomes in other settings (Gershenson et al., 2022). The preschool program also provided enrichment activities such as field trips and ex-

**Table 8:** Classroom Description and English Language Achievement Scores

| School Year | No. of Classrooms | No. Enrolled | Avg Classroom Size | Avg PPVT-R Scores Start %ile | End %ile |
|---|---|---|---|---|---|
| 1995-1996 | 39 | 780 | 20 | 16 | 34 |
| 1996-1997† | 48 | 921 | 19 | 30 | 42 |
| 1997-1998 | 47 | 940 | 20 | 21 | 39 |
| 1998-1999‡ | 48 | 960 | 20 | 34 | 53 |
| 1999-2000 | 49 | 942 | 19 | 37 | 53 |

Note: †This measure is interpolated as the data was provided within three different score range categories 1-25, 26-75, 76-99. ‡The PPVT Changed in 1998-1999 and percentiles are not comparable to previous years. Source: Annual Preschool program reports 1995-2000 school years.

cursions. Along with labor market gains, the potential for increased community attachment may help to explain why girls in the program are more likely to reside in-state as young adults, a generally positive outcome for Indigenous communities. It is not clear, however, why such a mechanism would work on girls more than boys. We are not the first study to find some heterogeneous impacts of preschool by gender (Anderson, 2008; Deming, 2009; Magnuson et al., 2016).

Cognitive gains could also drive effects. In addition to emphasizing the Indigenous community, the preschool curriculum was firmly focused on standard achievement outcomes. Aggregate data for the total preschool Peabody Picture Vocabulary Tests scores indicate that there were significant improvements over the course of the school year. Annual reports for the preschool program for various years indicate that there were improvements in the PPVT scores from the beginning to the end of the year by approximately 15 to 20 percentage points (see Table 8). As mentioned previously, the preschool program required all teachers to have at least a Bachelor's degree in early childhood education; the teaching assistants (there was one in each classroom) were required to have a child development associate's degree at a minimum. Therefore, there were relatively high educational credentials required for both teachers and their assistants in these classrooms. Continuing education and skill development were required as well including 16 hours of professional development training annually. The data also indicate that on average classroom sizes were kept relatively small ranging between 19 and 20 students per classroom over all five years in our data. Other researchers studying the randomization of classroom size in Tennessee have shown that assignment to smaller classroom sizes in kindergarten led to beneficial long-run outcomes (Chetty et al., 2011).

### 6.1 Evidence on College Attendance

Researchers have shown that attending preschool programs leads to an increase in educational attainment and college completion (Bailey et al., 2021; Gray-Lobe et al., 2023). This is a likely intermediate mechanism that may explain the higher earnings that we observe in our study, particularly those at ages 21-23. While we observe the long-term earnings in early adulthood for the children who applied to the preschool program, we do not have information on their college enrollment or postsecondary educational attainment. Currently, we do not have permission to link our data to additional sources such as the National Student Clearinghouse data.

However, we do have data on the years in which the children's parents claimed them as dependents in the IRS Form 1040 data. This is potentially informative about college-going, because parents can continue to claim children ages 19-23 as dependents if they provide more than half of their child's support and their child attends school full-time. However, among young people not in college, being claimed by a parent may be an adverse outcome, as it implies they are not financially independent. Further complicating this indicator, low-income students, in particular, may financially support themselves while in college, and not be claimed by anyone else. Being claimed at ages 24 and above suggests the child is heavily financially dependent on the parent and/or has a permanent disability. Thus we separate three patterns of dependent claiming: being claimed at ages 19-23, being claimed at ages 24 and above, and being claimed at ages 19-23 but not 24 and above. The latter would arguably best isolate typical college student claiming patterns.

In Table 9 we provide the estimated treatment effects of preschool admission on later dependent claiming activity. None of the coefficients are statistically significant, but the sign and magnitudes suggest that admitted students may be less likely to be claimed as a dependent during college ages or after compared to non-admitted students, and equally likely to exhibit the hypothetically ideal pattern of being claimed at ages 19-23 but not after. Ultimately, we have no evidence to suggest that the preschool program is driving early labor market gains by discouraging college attendance, but we also have no evidence to confirm improved college-going rates for this program.

We contend that labor market gains in young adulthood are a positive result irrespective of college-going, since, especially for disadvantaged communities, the two may go hand-in-hand. It is estimated that 70-80 percent of college students work (Carnevale et al., 2015). In the Boston context, Gray-Lobe et al. (2023) find that roughly one-half of preschool applicants enroll in college on time, in a city with an abundance of colleges. The enrollment rate in the population under study is likely lower. Thus, there is plenty of capacity for early labor market gains regardless of whether individuals attend college, and it should not be assumed that early labor market gains we document come at the expense of college attendance.

## 7 Conclusion

Our analysis focuses on a universal preschool program with lottery admissions for a population of Indigenous children throughout a single U.S. state. In contrast to prior work, we are able to link individuals who were and were not offered admissions via the lottery to their reported earnings and income data in their IRS Form W-2 and 1040 records. Additionally, we are able to link individuals back to their parents' 1040 tax forms to construct measures of childhood family income. As a result, we provide the first estimates of the return to education for an early childhood education intervention in the U.S., including documenting heterogeneity in returns by family income and gender.

We find that attending a preschool program for one year increases average adult earnings by approximately 5 to 6 percent. These results are robust to different measures of parental income or earnings. We find particularly large effects for women and for those from the bottom

**Table 9:** Parents' Claiming Child as Dependent in Early Adulthood

| | Child Claimed as a Dependent |
|---|---|
| Ages 19-23 | -0.015 |
| | (0.016) |
| Age 24 or later | -0.012 |
| | (0.010) |
| Ages 19-23, but not after | -0.004 |
| | (0.017) |

Note: All numbers have been rounded according to the U.S. Census Bureau Disclosure Review Board specifications for this table. These figures have been approved for disclosure. DRB Approval Number: CBDRB-FY24-0038. All specifications include gender, parental characteristics, and application-year-by-district fixed effects. Intention to Treat Analysis. *** indicates statistical significance at the 1% level, ** indicates statistical significance at the 5% level, * indicates statistical significance at the 10% level.

half of the parental income distribution at the time of preschool application. There is suggestive evidence that the increased earnings observed for women is related to residing in the home state and, to lesser extent, due to a reduction in marriage. Further work will need to clarify if these characteristics play an important role in earnings determination or if they are merely a byproduct of increased college attendance for this population. Overall, we find evidence that high quality preschool programs may play an important role in reducing inequality in adulthood for the most disadvantaged subgroups within a population.

## References


Algan, Y., E. Beasley, S. Côté, J. Park, R. E. Tremblay, and F. Vitaro (2022). The impact of childhood social skills and self-control training on economic and noneconomic outcomes: Evidence from a randomized experiment using administrative data. *American Economic Review 112*(8), 2553–79.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association 103*(484), 1481–1495.

Bailey, D. H., G. J. Duncan, F. Cunha, B. R. Foorman, and D. S. Yeager (2020). Persistence and fade-out of educational-intervention effects: Mechanisms and potential solutions. *Psychological Science in the Public Interest 21*(2), 55–97.

Bailey, M. J., S. Sun, and B. Timpe (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.

Barr, A. and C. R. Gibbs (2022). Breaking the cycle? intergenerational effects of an antipoverty program in early childhood. *Journal of Political Economy 130*(12), 3253–3285.

Bartik, T. J. and B. Hershbein (2017). Pre-k in the public schools: Evidence from within all states. In *Summer Research Workshop, Institute for Research on Poverty, University of Wisconsin–Madison. Madison (June 19–22, 2017).*

Campbell, F., G. Conti, J. J. Heckman, S. H. Moon, R. Pinto, E. Pungello, and Y. Pan (2014). Early childhood investments substantially boost adult health. *Science 343*(6178), 1478–1485.

Carnevale, A., N. Smith, M. Melton, and E. Price (2015). Learning while earning: The new normal. *Georgetown University Center on Education and the Workforce*.

Cascio, E. U. (2017). Does universal preschool hit the target? program access and preschool impacts. Technical report, National Bureau of Economic Research.

Cascio, E. U. and D. W. Schanzenbach (2013). The impacts of expanding access to high-quality preschool education. *Brookings papers on economic activity* (2), 1–54.

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan (2011). How does your kindergarten classroom affect your earnings? evidence from project star. *The Quarterly journal of economics 126*(4), 1593–1660.

Currie, J. and D. Thomas (2000). Eschool quality and the longer-term effects of head start. *Journal of Human Resources 35*(4), 755–774.

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from head start. *American Economic Journal: Applied Economics 1*(3), 111–134.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics 132*(4), 1593–1640.

Dodge, K. A., Y. Bai, H. F. Ladd, and C. G. Muschkin (2017). Impact of north carolina's early childhood programs and policies on educational outcomes in elementary school. *Child development 88*(3), 996–1014.

Faircloth, S. C. (2015). The early childhood education of american indian and alaska native children: State of the research. *Journal of American Indian Education 54*(1), 99–126.

García, J. L., J. J. Heckman, and V. Ronda (2021). The lasting effects of early childhood education on promoting the skills and social mobility of disadvantaged african americans. Technical report, National Bureau of Economic Research.

Gershenson, S., C. M. Hart, J. Hyman, C. A. Lindsay, and N. W. Papageorge (2022). The long-run impacts of same-race teachers. *American Economic Journal: Economic Policy 14*(4), 300–342.

Gormley Jr, W. T., T. Gayer, D. Phillips, and B. Dawson (2005). The effects of universal pre-k on cognitive development. *Developmental psychology 41*(6), 872.

Gormley Jr, W. T., D. Phillips, and S. Anderson (2018). The effects of tulsa's pre-k program on middle school student performance. *Journal of Policy Analysis and Management 37*(1), 63–87.

Gray-Lobe, G., P. A. Pathak, and C. R. Walters (2023). The long-term effects of universal preschool in boston. *The Quarterly Journal of Economics 138*(1), 363–411.

Harden, B. J., B. E. Brett, J. T. Gross, C. Weiland, J. Berne, E. L. Klein, and C. Tirrell-Corbin (2023). Benefits of pre-kindergarten for children in baltimore, md. *Early Childhood Research Quarterly 64*, 1–12.

Heckman, J., R. Pinto, and P. Savelyev (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review 103*(6), 2052–86.

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Yavitz (2010). The rate of return to the highscope perry preschool program. *Journal of public Economics 94*(1-2), 114–128.

Kline, P. and C. R. Walters (2016). Evaluating public programs with close substitutes: The case of head start. *The Quarterly Journal of Economics 131*(4), 1795–1848.

Layne, M., D. Wagner, C. Rothhaas, et al. (2014). Estimating record linkage false match rate for the person identification validation system. *Center for Administrative Records Research and Applications Working Paper 2.*

Lipsey, M. W., D. C. Farran, and K. Durkin (2018). Effects of the tennessee prekindergarten program on children's achievement and behavior through third grade. *Early Childhood Research Quarterly 45*, 155–176.

Ludwig, J. and D. L. Miller (2007). Does head start improve children's life chances? evidence from a regression discontinuity design. *The Quarterly journal of economics 122*(1), 159–208.

Magnuson, K. A., R. Kelchen, G. J. Duncan, H. S. Schindler, H. Shager, and H. Yoshikawa (2016). Do the effects of early childhood education programs differ by gender? a meta-analysis. *Early childhood research quarterly 36*, 521–536.

Romero-Little, M. E. (2010). How should young indigenous children be prepared for learning? a vision of early childhood education for indigenous children. *Journal of American Indian Education*, 7–27.

Rossin-Slater, M. and M. Wüst (2020). What is the added value of preschool for poor children? long-term and intergenerational impacts and interactions with an infant health intervention. *American Economic Journal: Applied Economics 12*(3), 255–286.

Weiland, C., R. Unterman, S. Dynarski, R. Abenavoli, H. Bloom, B. Braga, A.-M. Faria, E. H. Greenberg, B. Jacob, J. A. Lincove, et al. (2023). Lottery-based evaluations of early education programs: Opportunities and challenges for building the next generation of evidence.

Weiland, C., R. Unterman, A. Shapiro, S. Staszak, S. Rochester, and E. Martin (2020). The effects of enrolling in oversubscribed prekindergarten programs through third grade. *Child Development 91*(5), 1401–1422.

Yazzie-Mintz, T. (2011). Native teachers' beliefs and practices: Choosing language and cultural revitalization over uniformity and standardization. *Contemporary Issues in Early Childhood 12*(4), 315–326.

# A Appendix Tables

**Table A1:** Regression Results for Pandemic Years Included

**Panel A.IV**

| | (1) Log Mean Wages Pandemic Alone | (2) Log Mean Wages All 19+ Incl Pan | (3) Log Mean Wages All 21+ Incl Pan | (4) Log Max Wages All 19+ Incl Pan | (5) Average Employ Pandemic Years | (6) Mean Employ All 19+ Incl Pan | (7) Mean Employ All 21+ Incl Pan |
|---|---|---|---|---|---|---|---|
| Attended | 0.0459 | 0.0296 | 0.0315 | 0.0321 | 0.00344 | 0.00955 | 0.00561 |
| | (0.0365) | (0.0271) | (0.0287) | (0.0277) | (0.0112) | (0.00806) | (0.00854) |
| Observations | 4,600 | 5,000 | 5,000 | 5,000 | 5,100 | 5,100 | 5,100 |
| R2 | 0.0468 | 0.0393 | 0.0362 | 0.0383 | 0.0195 | 0.0159 | 0.0153 |

**Panel B. OLS**

| | (1) Log Mean Wages Pandemic Alone | (2) Log Mean Wages All 19+ Incl Pan | (3) Log Mean Wages All 21+ Incl Pan | (4) Log Max Wages All 19+ Incl Pan | (5) Average Employ Pandemic Years | (6) Mean Employ All 19+ Incl Pan | (7) Mean Employ All 21+ Incl Pan |
|---|---|---|---|---|---|---|---|
| Admitted | 0.0441 | 0.0284 | 0.0302 | 0.0308 | 0.00330 | 0.00915 | 0.00538 |
| | (0.0352) | (0.0261) | (0.0276) | (0.0267) | (0.0108) | (0.00776) | (0.00822) |

**Panel C. OLS with IPW**

| | (1) Log Mean Wages Pandemic Alone | (2) Log Mean Wages All 19+ Incl Pan | (3) Log Mean Wages All 21+ Incl Pan | (4) Log Max Wages All 19+ Incl Pan | (5) Average Employ Pandemic Years | (6) Mean Employ All 19+ Incl Pan | (7) Mean Employ All 21+ Incl Pan |
|---|---|---|---|---|---|---|---|
| Admitted | 0.0449 | 0.0265 | 0.0291 | 0.0284 | 0.00489 | 0.0100 | 0.00625 |
| | (0.0368) | (0.0261) | (0.0277) | (0.0263) | (0.0108) | (0.00788) | (0.00827) |

Note: All numbers have been rounded according to the U.S. Census Bureau Disclosure Review Board specifications for this table. These figures have been approved for disclosure. DRB Approval Number: CBDRB-FY24-0038.

**Table A2:** Lee Bounds for Main Regression Results

| | Mean Wages Age 19+ | Mean Wages Age 21+ | Log Mean Wages Age 19+ | Log Mean Wages Age 21+ | Log Max Wage | Log Adjusted Gross Income Age 19+ | Log Adjusted Gross Income Age 21+ | Mean Employ 19+ | Mean Employ 21+ |
|---|---|---|---|---|---|---|---|---|---|
| Admitted | 717.1 | 924.8* | 0.0511* | 0.0532* | 0.0603** | 0.00545 | 0.0141 | 0.0111 | 0.00593 |
| | (444) | (555.6) | (0.0272) | (0.0296) | (0.0276) | (0.0405) | (0.0494) | (0.00807) | (0.00843) |
| Lower Bounds | | | | | | | | | |
| | -458.2 | -665.7 | 0.00577 | 0.00488 | 0.0119 | -0.0445 | -0.0560 | 0.000567 | -0.00405 |
| | (572.6) | (707.9) | (0.0326) | (0.0350) | (0.0333) | (0.0426) | (0.0506) | (0.00803) | (0.00818) |
| Upper Bounds | | | | | | | | | |
| | 1033** | 1174** | 0.107*** | 0.117*** | 0.119*** | 0.135*** | 0.176*** | 0.0246* | 0.0200 |
| | (461) | (567.6) | (0.0413) | (0.0452) | (0.0433) | (0.0483) | (0.0572) | (0.0138) | (0.0157) |

Note: These Lee Bounds have been calculated for the specified outcome variables at the upper and lower levels for standard errors. All numbers have been rounded according to the U.S. Census Bureau Disclosure Review Board specifications for this table. These figures have been approved for disclosure. DRB Approval Number: CBDRB-FY24-0038.

**Table A3:** Lee Bounds for Family Regression Results

| | Married Age 19+ | Claims Children Age 19+ | Resides in State At Last Observed |
|---|---|---|---|
| Admitted | -0.00543 | 0.0229 | 0.0145 |
| | (0.0147) | (0.0162) | (0.0119) |
| | | | |
| Lower Bounds | | | |
| | -0.0113 | 0.0320* | 0.00977 |
| | (0.0185) | (0.0184) | (0.0143) |
| | | | |
| Upper Bounds | | | |
| | 0.0147 | 0.0580*** | 0.0359* |
| | (0.0139) | (0.0158) | (0.0186) |

Note: These Lee Bounds have been calculated for the specified outcome variables at the upper and lower levels for standard errors. All numbers have been rounded according to the U.S. Census Bureau Disclosure Review Board specifications for this table. These figures have been approved for disclosure. DRB Approval Number: CBDRB-FY24-0038.

**Table A4:** Marital Status, Number of Children, and In State

A. IV

| | (1) Married 19+ | (2) Married Oldest Age | (3) Children 19+ | (4) Children Oldest Age | (5) In State 19+ | (6) In State Oldest Age |
|---|---|---|---|---|---|---|
| Attended | -0.00543 | 0.000267 | 0.0229 | 0.0165 | 0.0145 | 0.0149 |
| | (0.0147) | (0.0140) | (0.0162) | (0.0155) | (0.0119) | (0.0150) |
| | | | | | | |
| Observations | 4,800 | 4,900 | 4,800 | 4,900 | 4,800 | 4,900 |
| R2 | 0.0376 | 0.0313 | 0.0777 | 0.0679 | 0.0279 | 0.0247 |
| IV Coefficient | 0.96 | 0.96 | 0.96 | 0.96 | 0.96 | 0.96 |
| Std Error | 0.006 | 0.006 | 0.006 | 0.006 | 0.006 | 0.006 |
| F Stat | 743 | 754 | 743 | 754 | 743 | 757 |

B. OLS

| | (1) Married 19+ | (2) Married Oldest Age | (3) Children 19+ | (4) Children Oldest Age | (5) In State 19+ | (6) In State Oldest Age |
|---|---|---|---|---|---|---|
| Admitted | -0.00521 | 0.000256 | 0.0219 | 0.0158 | 0.0139 | 0.0143 |
| | (0.0142) | (0.0135) | (0.0156) | (0.0149) | (0.0115) | (0.0145) |

C. OLS with IPW

| | (1) Married 19+ | (2) Married Oldest Age | (3) Children 19+ | (4) Children Oldest Age | (5) In State 19+ | (6) In State Oldest Age |
|---|---|---|---|---|---|---|
| Admitted | -0.00656 | -0.000707 | 0.0218 | 0.0155 | 0.0156 | 0.0146 |
| | (0.0143) | (0.0136) | (0.0157) | (0.0149) | (0.0118) | (0.0146) |

Note: All numbers have been rounded according to the U.S. Census Bureau Disclosure Review Board specifications for this table. These figures have been approved for disclosure. DRB Approval Number: CBDRB-FY24-0038.

**Table A5:** OLS Regressions by Gender and Initial Median Household Income

A. Women

| | (1)<br>Log Mean Wages<br>Age 19+ | (2)<br>Mean Emp<br>Age 19+ | (3)<br>Married at<br>Age 19+ | (4)<br>Any Children<br>Age 19+ | (5)<br>In State at<br>Oldest Age |
|---|---|---|---|---|---|
| Admitted | 0.0891** | 0.0103 | -0.0408* | 0.00440 | 0.0520** |
| | (0.0377) | (0.0113) | (0.0223) | (0.0242) | (0.0216) |
| Observations | 2,400 | 2,400 | 2,300 | 2,300 | 2,300 |

B. Men

| | (1)<br>Log Mean Wages<br>Age 19+ | (2)<br>Mean Emp<br>Age 19+ | (3)<br>Married at<br>Age 19+ | (4)<br>Any Children<br>Age 19+ | (5)<br>In State at<br>Oldest Age |
|---|---|---|---|---|---|
| Admitted | 0.0145 | 0.0116 | 0.0261 | 0.0330 | -0.0176 |
| | (0.0392) | (0.0115) | (0.0182) | (0.0203) | (0.0197) |
| Observations | 2,600 | 2,700 | 2,500 | 2,500 | 2,600 |

C. Below Median

| | (1)<br>Log Mean Wages<br>Age 19+ | (2)<br>Mean Emp<br>Age 19+ | (3)<br>Married at<br>Age 19+ | (4)<br>Any Children<br>Age 19+ | (5)<br>In State at<br>Oldest Age |
|---|---|---|---|---|---|
| Admitted | 0.0779* | 0.0129 | -0.0157 | 0.0291 | 0.0255 |
| | (0.0403) | (0.0118) | (0.0205) | (0.0231) | (0.0208) |
| Observations | 2,500 | 2,600 | 2,400 | 2,400 | 2,400 |

D. Above Median

| | (1)<br>Log Mean Wages<br>Age 19+ | (2)<br>Mean Emp<br>Age 19+ | (3)<br>Married at<br>Age 19+ | (4)<br>Any Children<br>Age 19+ | (5)<br>In State at<br>Oldest Age |
|---|---|---|---|---|---|
| Admitted | 0.0370 | 0.0139 | -0.00494 | 0.010 | 0.00654 |
| | (0.0370) | (0.0111) | (0.0199) | (0.0213) | (0.0204) |
| Observations | 2,500 | 2,600 | 2,400 | 2,400 | 2,500 |

Note: All numbers have been rounded according to the U.S. Census Bureau Disclosure Review Board specifications for this table. These figures have been approved for disclosure. DRB Approval Number: CBDRB-FY24-0038.

**Table A6:** OLS Regressions in Subgroups by Gender and Initial Median Household Income

A. Women Below Median

| | (1) Log Mean Wages Age 19+ | (2) Mean Emp Age 19+ | (3) Married at Age 19+ | (4) Any Children Age 19+ | (5) In State at Oldest Age |
|---|---|---|---|---|---|
| Admitted | 0.117** | 0.0203 | -0.0552* | -0.0314 | 0.0657** |
| | (0.0586) | (0.0177) | (0.0327) | (0.0364) | (0.0322) |
| Observations | 1,200 | 1,200 | 1,100 | 1,100 | 1,100 |

B. Women Above Median

| | (1) Log Mean Wages Age 19+ | (2) Mean Emp Age 19+ | (3) Married at Age 19+ | (4) Any Children Age 19+ | (5) In State at Oldest Age |
|---|---|---|---|---|---|
| Admitted | 0.0573 | 0.00652 | -0.0355 | 0.0328 | 0.0427 |
| | (0.0495) | (0.0146) | (0.0309) | (0.0327) | (0.0296) |
| Observations | 1,200 | 1,200 | 1,200 | 1,200 | 1,200 |

C. Men Below Median

| | (1) Log Mean Wages Age 19+ | (2) Mean Emp Age 19+ | (3) Married at Age 19+ | (4) Any Children Age 19+ | (5) In State at Oldest Age |
|---|---|---|---|---|---|
| Admitted | 0.0358 | 0.00700 | 0.0225 | 0.0723** | -0.0108 |
| | (0.0558) | (0.0160) | (0.0263) | (0.0300) | (0.0277) |
| Observations | 1,300 | 1,400 | 1,300 | 1,300 | 1,300 |

D. Men Above Median

| | (1) Log Mean Wages Age 19+ | (2) Mean Emp Age 19+ | (3) Married at Age 19+ | (4) Any Children Age 19+ | (5) In State at Oldest Age |
|---|---|---|---|---|---|
| Admitted | 0.00919 | 0.0208 | 0.0232 | -0.0153 | -0.0239 |
| | (0.0556) | (0.0167) | (0.0257) | (0.0278) | (0.0285) |
| Observations | 1,300 | 1,300 | 1,200 | 1,200 | 1,300 |

Note: All numbers have been rounded according to the U.S. Census Bureau Disclosure Review Board specifications for this table. These figures have been approved for disclosure. DRB Approval Number: CBDRB-FY24-0038.