Parents' Earnings and the Returns to Universal Pre-Kindergarten*

John Eric Humphries, *Yale University & NBER*

Christopher Neilson, *Yale University & NBER*

Xiaoyang Ye, *Brown University*

Seth Zimmerman, *Yale SOM & NBER*†

April 15, 2025

Abstract

This paper asks whether universal pre-kindergarten (UPK) programs can increase parental earnings and, if so, how much these gains affect the economic returns to UPK. Using admissions lotteries for an extended-day UPK program in New Haven, Connecticut, we find that UPK enrollment increases childcare coverage to span the workday and raises parents' earnings by 21.7% during pre-kindergarten. Gains persist for at least six years. We find little evidence of effects on children's academic and behavioral outcomes during elementary and middle school. Combining these results, we demonstrate that tax revenues from parents' earnings gains reduce the net government costs of UPK by 90% relative to estimates that ignore gains for parents. Overall, we estimate that each dollar spent yields $10 in benefits. Our findings demonstrate the potential of UPK programs that combine quality education with full-day childcare and underscore the importance of thinking about parents when designing and evaluating early-childhood policies.

*We thank Michele Bonanno, Marquelle Middleton, and the New Haven Public Schools for assistance in accessing school assignment data. We thank Katie Breslin, Scott Gaul, Ajit Gopalakrishnan, Xian Guan, Liam McGucken, Rachel Leventhal-Weiner, and other staff at the Connecticut State Department of Education, Department of Labor, and Office of Early Childhood for assistance in accessing state data. We thank Dave Wilkinson and Nat McLaughlin for their assistance with cross-governmental coordination. We thank Sergio Arango, Manuela Cardona Badillo, Florian Caro, Clara Chambers, and Pablo Uribe for their excellent research assistance. We thank Manasi Deshpande, Paul Goldsmith-Pinkham, Henrik Kleven, Magne Mogstad, and numerous seminar participants for their valuable comments. This research received financial support from the Tobin Center for Economic Policy. Neilson discloses that he is the CEO of and has a financial stake in TetherEd, a firm that provides school choice support services to school systems, including to the New Haven public schools. Zimmerman discloses that he holds a volunteer position as a member of the Connecticut State Board of Education and that he has a financial stake in TetherEd. All errors are our own. This work would not be possible without data provided by the P20 WIN Participating Agencies. The findings do not necessarily reflect the opinions of the State of Connecticut or the organizations and agencies contributing data.

†Humphries: johneric.humphries@yale.edu. Neilson: christopher.neilson@yale.edu. Ye: xiaoyang.ye26@gmail.com. Zimmerman: seth.zimmerman@yale.edu.

1 Introduction

Over the past thirty years, many US states and cities have implemented free, large-scale, non-means-tested pre-kindergarten programs—i.e., universal pre-kindergarten (UPK).[1] At the federal level, expanded investment in UPK has been a topic of substantial debate, with the Biden administration's American Families Plan calling for $200 billion in UPK funding (White House, 2021). The policy logic underlying UPK expansion is that parents may lack access to or underinvest in pre-kindergarten childcare and that the educational benefits for children and expanded labor market opportunities for parents combine to outweigh the costs of public provision.

Despite the central role of labor market opportunity for parents in the case for UPK, evidence on how UPK affects parents' earnings is limited. Existing work relies on non-randomized research designs to evaluate programs that do not provide childcare for the full workday. This research typically cannot rule out null effects, though, as we discuss below, the standard errors are large (Fitzpatrick, 2010; Cascio and Schanzenbach, 2013; Cascio, 2021). As a result, the cost-benefit proposition posed by UPK programs—and in particular by a growing set of city-level programs that combine high-quality educational inputs with childcare coverage that spans the workday—remains unclear. Understanding whether such programs improve labor market outcomes for parents is crucial for program design and evaluation.

A central challenge facing researchers is that it is hard to link random variation in children's access to UPK to outcomes for parents. In this paper, we address that challenge head-on. We use data from admissions lotteries in a large extended-day UPK program in New Haven, Connecticut to provide a randomized evaluation of the effects of UPK enrollment on parents' labor market outcomes. We combine our estimates of effects for parents with estimates of educational effects for kids and net program costs to conduct cost-benefit analyses that incorporate outcomes for both parents and children. Our key finding is that earnings gains for parents are large and, because of this, the returns to UPK investment are high.

The New Haven Public Schools (NHPS) have offered extended-day public pre-kindergarten to three- and four-year-olds since the late 1990s. These programs enroll students from New Haven and the surrounding suburbs and are universal in that they are not means-tested. In practice, however, they are often slot-constrained, with spots allocated through a centralized assignment mechanism that combines coarse priority groups with lottery tiebreakers. These lotteries provide the random variation in program access necessary for our empirical design.

[1] Boston, New York, and Washington, DC have each adopted a policy of this type at the municipality level. Friedman-Krauss et al. (2023) review the policy environment and report that six states and Washington, DC have "mostly achieved" universal pre-kindergarten.

The NHPS UPK programs share key features with programs in other cities, such as Boston, New York, and Washington, DC. One such feature is long hours. New Haven supplements a 6.5 hour academic day with before- and after-care for a total of ten hours of care during most of our study period. This is similar to the nine-hour day offered at Boston's early learning centers (BPS, 2024). Another is high educational quality. Standards for curricula and teacher qualifications in New Haven parallel those in public schools. A third is the presence of binding capacity constraints within the UPK system. Boston, New York, and Washington, DC each also use lottery tiebreakers to allocate slots in UPK programs. Finally, in each of these cities, UPK programs coexist with a wide variety of other pre-kindergarten options, including public means-tested programs such as Head Start and private providers that offer subsidized and unsubsidized slots.

Our analysis relies on a set of linked datasets that together help us capture the effects of enrollment in universal pre-kindergarten on parent, child, and cost outcomes. We start with records of NHPS pre-kindergarten admissions lotteries from 2003 through 2022. We link children in this dataset to state data on school enrollment and achievement. In addition, we link the *parents* listed on the admissions forms to earnings and employment data from state Unemployment Insurance (UI) records. Finally, we survey parents of past applicants about pre-kindergarten enrollment at public and private programs, childcare costs, and hours worked.

We begin our analysis by describing the childcare arrangements from which children substitute when they enroll in UPK. We find that nearly all students substitute away from some other form of childcare outside the home, so UPK enrollment does not raise overall childcare usage. Despite the absence of *enrollment* effects, we find large effects on *hours* of childcare. Due to the long UPK school day, children enrolling in UPK get an average of 11.3 more hours of childcare coverage per week.

We next consider how UPK enrollment affects children's educational outcomes. We find little evidence that UPK affects children's outcomes in the medium run. Effects on test scores, attendance, and grade retention in kindergarten through eighth grade do not statistically differ from zero and are generally small in economic terms. These findings mirror results from recent randomized evaluations of similar US UPK programs (Lipsey et al., 2018; Durkin et al., 2022; Gray-Lobe et al., 2023).

We then turn to our main analysis, which examines the impact of UPK enrollment on parents' labor market outcomes. We find evidence of large and sustained earnings gains. Enrolling a child in a UPK program raises parent earnings by 21.7% (SE=6.6%), or $5,464 per year, during the one- or two-year period when the child is of pre-kindergarten age. Earnings effects persist for at least six years after pre-kindergarten ends. The average earnings effect over this period is 20.9% (SE=7.9%), or $6,474 per year.

Earnings gains appear to arise from a combination of modestly increased labor supply and reduced career disruption. On the labor supply side, we find that labor force participation during the pre-kindergarten years rises by 5.7 percentage points (SE=2.6), or 7%, on a base of 82.4%. This effect persists at a roughly similar magnitude for the six years following pre-kindergarten. Turning to job disruption, we find that children's UPK enrollment reduces the likelihood that parents experience career gaps or switch industries during the pre-kindergarten period, while increasing the chances that parents hold a single high-earning job. For example, UPK enrollment reduces the share of individuals who switch industry during the pre-kindergarten years by 34%. Qualitative reports lend further support to this idea. Survey respondents described how UPK helped them maintain their career paths, keep normal work hours, or work more effectively.

How are the benefits of UPK distributed? To answer this question, we split our analysis by terciles of pre-enrollment neighborhood income. We first show that, because the UPK program serves a relatively low-income population, the top tercile of the applicant population consists mostly of people with earnings near the population mean. Turning to treatment effects, we find that parents' earnings gains are large in the top two terciles while we cannot rule out a null effect in the bottom tercile. This finding is consistent with the observation that the career returns to longer work hours are high for more skilled workers (Kuhn and Lozano, 2008; Cortés and Pan, 2019). We conclude that benefits mostly accrue to middle-income families.

We bring together our analyses of effects on parents, children, and program costs to construct cost-benefit calculations for the program. We apply the MVPF framework of Hendren and Sprung-Keyser (2020). The MVPF is the ratio of recipients' after-tax willingness to pay (WTP) for the program to net program costs.

We find that the overall MVPF of the UPK program is high, and that earnings gains for parents are the crucial reason why. On the cost side, tax revenues generated by parents' earnings reduce net program costs by around 90% compared to calculations that omit this revenue source. Starting with upfront costs of $24,200, substitution away from other programs saves $8,800, and tax revenue from children's projected earnings saves $430 more. Taxes from parents' income then reduce net costs from $15,400 to $1,500. Thus, for any given value of WTP, net cost reductions from parents' earnings gains multiply the MVPF by about 10.

We next consider willingness to pay. A challenge here is that families' willingness to pay for the program depends on the extent to which the welfare benefits of increased earnings are offset by the disutility of work, which in turn depends on the availability of substitutes for extended-day care in the private market and how parents weigh benefits for children. We provide a range of estimates, with the goal of bracketing the true effect. Our benchmark approach, in which families value UPK at the net cost of provision, yields an MVPF of 10.1: each dollar of spending on UPK generates \$10.10 in benefits. Alternate calculations based on different assumptions about how families value the earnings gains from UPK yield MVPF values ranging from 4.5 to 39.8.

How does the cost-benefit proposition posed by UPK compare to other programs? From the parent side, UPK performs well. Focusing on our benchmark approach, a "no-kid" evaluation of UPK that excludes gains for children yields an MVPF of 7.9. This value is high relative to other policies that aim to promote adult labor market activity. For example, Hendren and Sprung-Keyser (2020) estimate MVPFs for EITC policies of around 1. In contrast, our benchmark approach yields a "no-parent" MVPF (i.e., excluding earnings gains for parents) of 1.03. This estimate is well below estimates reported in Hendren and Sprung-Keyser (2020) for a variety of policy interventions targeted at children, such as increased school spending, expanded health insurance, and expanded college access.

We contribute to several strands of literature. Most directly, we build on research that evaluates the labor market effects of other US UPK programs (Fitzpatrick, 2010; Cascio and Schanzenbach, 2013). This work uses non-randomized designs to evaluate state-level policies in Georgia and Oklahoma, which took effect in the 1990s and mandated between 2.5 and 6.5 hours of childcare per day—less than a typical workday. The stylized conclusion from this line of work is that there is little evidence that UPK affects mothers' labor supply (Cascio, 2021). Our findings provide evidence that the extended-day UPK model employed in several major cities can have large effects on parents' earnings, and that these effects are crucial for cost-benefit evaluation.

Our work also builds on papers studying the labor supply effects of means-tested and demonstration pre-kindergarten programs, public kindergarten programs, and universal pre-kindergarten programs outside the US (Gelbach, 2002; Cascio, 2009; Havnes and Mogstad, 2011; Fitzpatrick, 2012; Sabol and Chase-Lansdale, 2015; Chaparro et al., 2020; García et al., 2020; Schiman, 2022; Attanasio et al., 2024; Wikle and Wilson, 2023). These settings differ from the contemporary US UPK context. For example, means-tested programs target low-income groups for whom labor market opportunities are, by definition, more limited. Evaluations of demonstration programs, kindergarten programs, and universal programs in developed countries are based on data from the 1970s through the early 2000s, and do not capture the effects of the massive increase in childcare costs in the US that began in the late 1990s and continues through the present (Swenson and Burgess Simms, 2021). In contrast, our evaluation of a contemporary UPK program speaks directly to the ongoing debate over UPK policy in the US.

Finally, we contribute to the literature on the educational effects of UPK. Our findings on test scores echo results from randomized evaluations of programs in Boston, Tennessee, and Georgia (Lipsey et al., 2018; Weiland et al., 2020; Gray-Lobe et al., 2023;

Woodyard et al., 2023; Bruhn and Emick, 2023). Our findings demonstrate that UPK programs have the potential to yield high economic returns even if gains for children are modest and underscore the importance of considering outcomes for parents in both design and evaluation.

2 Institutions

2.1 Program attributes

Our study focuses on free, non-means-tested pre-kindergarten programs in the New Haven Public Schools. New Haven is a low-income, majority-minority school district. In 2022-23, 83% of students in NHPS were Black or Hispanic and 66% qualified for free or reduced price lunch, compared to 43% and 42% statewide (see Appendix Table A.1).

New Haven's UPK programs serve three- and four-year-olds in grades PK3 and PK4, respectively. There are nine total sites, each of which is part of an elementary school that runs through eighth grade. Eight sites operate as part of NHPS' interdistrict magnet program and are open to any student in Connecticut. The ninth is a charter school that opened in 2014 and is only open to New Haven residents. As reported in Panel (a) of Figure 1, UPK enrollment grew during the 2000s before stabilizing at around 700 students annually by 2014.

The UPK programs offer high-quality academic programming. Each of the eight magnet-affiliated UPK programs provides 6.5 hours of educational curriculum with a certified teacher during the school day (Bonanno, 2023). Maintaining a high-quality curriculum is one of the requirements of the granting process through which the magnet schools are established (State of Connecticut, 1999). The charter school follows a Montessori curriculum, also with a 6.5 hour academic day (Elm City Montessori, 2024).

The UPK programs also provide extended-day childcare coverage. Prior to the 2021-22 academic year, each magnet school offered before-care starting at 7:30am and after-care until 5:30pm, for a total of 10 hours of free childcare each day.[2] The charter also offers extended-day care starting at 7:30am and ending at 5:30pm. Unlike the magnet programs, the charter charges a modest fee for wraparound care.

NHPS' UPK programs are typically oversubscribed. Panel (b) of Figure 1 reports the ratio of open slots to the number of applicants separately for PK3 and PK4 from 2003 to 2021. Before 2021, there were between 1.5 and 4 slots available for every 10 PK3 applicants in each application cycle. Oversubscription is even more pronounced in PK4 because most PK4 slots are filled by rising PK3 students admitted the previous year. Relative demand slackened for both grades in 2021, when NHPS schools were still

[2]Several schools reduced before- and after-care offerings starting with the 2021-22 school year.

in the process of reopening after the Covid-19 pandemic.

NHPS resolves excess demand using a centralized assignment mechanism that combines coarse neighborhood, zip code, and sibling priorities with random lottery tiebreakers. Students typically apply in February and learn about their placement outcomes in early spring. The details of this process have changed over time (Peak, 2019; Akbarpour et al., 2022), but the basic structure has been in place since the beginning of the interdistrict magnet system in the late 1990s (Davidoff-Gore, 2017, p. 48). The lottery tiebreakers generate the exogenous variation in school assignments that we use to evaluate the effects of UPK access. See Online Appendix B for details.

New Haven is one of several cities that in recent years have developed and scaled up UPK programs characterized by high educational quality, extended childcare coverage, and binding system-wide or location-specific capacity constraints. Boston is a leading example here. During the 2000s, Boston scaled up a free, full-day, untargeted public pre-kindergarten program with high-quality educational inputs (Duncan and Murnane, 2014). The program is capacity constrained, with excess demand resolved through a centralized assignment system (Gray-Lobe et al., 2023). Early learning center sites offer free before- and after-care that extend childcare coverage from 7:30am to 4:35pm (BPS, 2024). Washington, DC and New York City scaled up untargeted pre-kindergarten programs in the late 2000s and 2010s, respectively. These programs enforce educational standards similar to those in public schools, impose minimum day lengths of 6.5 hours, with extended childcare coverage available for some students, and often face capacity constraints of different types.[3]

The modern *city*-level programs in New Haven and elsewhere contrast in important ways with *state*-level UPK policies. Foremost among these is the length of the school day. Many states with UPK policies on the books require only a 2- to 3-hour school day (Friedman-Krauss et al., 2023). These states include Oklahoma, the focus (along with Georgia, which requires a 6.5 hour day) of major existing evaluations of UPK's labor supply effects (Fitzpatrick, 2010; Cascio and Schanzenbach, 2013). The crucial question is whether the benefits of programs that offer both high-quality educational inputs and extended-day childcare coverage are enough to justify the higher costs.

2.2 Market context

Many different pre-kindergarten options are available to families in New Haven. However, the UPK programs we study are the only choice that offers free extended-day care

[3]See NYCPS (2023) and Shapiro (2023) for a discussion of the assignment mechanism in New York, NYCPS (2024) for a discussion of childcare coverage, and Brown (2018) for a discussion of program quality. See DCPS (2024) for a discussion of childcare coverage in DC and Malik (2018) for a discussion of other program attributes.

without means testing. Other subsidized options for three- and four-year-olds in the area include federal Head Start programs, two major state programs (School Readiness and Care 4 Kids), and a set of smaller state-level programs. Head Start programs are means-tested and often have limited hours during the period we study. The state programs are all either means-tested, not extended-day, not free, or some combination thereof. Families may also enroll their children in private pre-kindergarten programs without any subsidy. To describe the market for pre-kindergarten, we assemble data on the subsidized pre-kindergarten programs from a variety of sources. Online Appendix C details the different program types and the data available for each.

Panel (c) of Figure 1 reports enrollment in subsidized programs by program type for PK3 and PK4 students in New Haven between 2006 and 2021. The UPK programs account for 22% of subsidized enrollment during the 2013-2018 period, the years for which our administrative records are richest. For comparison, Head Start accounts for 34% of enrollment and Care 4 Kids accounts for 28% of enrollment.

Market conditions in New Haven are similar to those in other cities. As in New Haven, UPK programs in New York and Boston share a market with private and means-tested options. The average weekly price for private center-based care in the New Haven area was $255 in 2018, similar to Los Angeles and Chicago but below an average price of $291 in the Boston area. As in other cities, the price of center-based care in New Haven rose steadily over the 2000s. See Online Appendix D for details.

3 Data

3.1 NHPS admissions and enrollment data

Application and admissions records for the NHPS UPK programs form the center of our analysis. These records span the years 2003 through 2022. For each year in this range, we observe students' full choice applications, the administrative rules and student priorities used to process the applications, and school placement outcomes. Applications also include demographic information such as race/ethnicity, gender, and age. We supplement the admissions records with NHPS data on enrollment in district schools, including the UPK programs. NHPS UPK programs began in the late 1990s, so the program was running at scale several years before our first cohort of applicants.

The first column of Panel A of Table 1 describes the sample of UPK applicants. We observe 18,795 applications from 16,037 individuals. 41.8% of applicants are Black and 28.5% are Hispanic. 52% apply to grade PK3 and 48% to PK4. On average, applicants reside in Census block groups where the median household income is $59,713.[4] 25.1%

[4]We use 2019 ACS 5-year data and convert to 2015 dollars, as we do for dollar values throughout.

of applicants are assigned to a UPK program, and 26.5% go on to enroll in a program that year. Our UPK assignment variable measures "initial offers" (De Chaisemartin and Behaghel, 2020), so we expect two-sided noncompliance, with some non-offered students attending UPK programs and some offered students declining spots.

3.2 School enrollment and achievement

We worked with the Connecticut State Department of Education (CSDE) and NHPS to link UPK applicant records to statewide data on student enrollment and academic achievement using name and date of birth. This link, conducted by state personnel, allows us to observe outcomes for NHPS UPK applicants who enroll in any Connecticut public school from 2006-2022, allowing us to follow many applicants through middle school, and the oldest to roughly age 23. We exclude the 2005, 2008, and 2022 UPK application cohorts from our analysis of linked state records because we did not have these data when the merge was conducted. The 2022 application process took place after the merge, and we did not recover historical records of the 2005 and 2008 application processes until after the merge.

The CSDE data include scores on mandated state assessments conducted in kindergarten and grades 3 through 8. Kindergarten measures come from the Kindergarten Entrance Inventory (KEI), which assesses literacy, numeracy, social skills, physical skills, and creativity. Test scores in grades 3-8 are from statewide accountability exams in math and reading. We average across standardized subscores for our primary achievement measures. Online Appendix E describes the achievement measures we use.

Our state records also include a comparison set of all students enrolled in public pre-kindergarten programs in New Haven County during 2003-2021. These students provide the reference population for standardizing achievement measures. Appendix Table A.2 presents descriptive statistics for this population.

Column 2 of Table 1 describes the sample of applications we match to state data. Conditional on attempting a merge (i.e., excluding the 2005, 2008, and 2022 application years), we match 91% of observations, for a total of 16,485 matched applications and 13,917 matched individuals. Demographics, assignment rates, and enrollment rates for this group are essentially identical to those for the full set of UPK applicants. UPK applicants score 0.123σ above the New Haven County average on their KEI. This gap diminishes to 0.069σ and 0.038σ by grades 3 and 8, respectively.

In addition to our link to the CSDE public school records, we link our UPK applicant data to Connecticut Office of Early Childhood (OEC) records on subsidized private early childhood education, allowing us to observe children using state-provided subsidies in private programs. The time span of coverage varies by program type. We observe enrollment data for all major OEC-funded programs from 2016-2018, with more limited data from 2013-15 and 2019. As reported in Panel (c) of Figure 1, these programs account for about 22% of observed enrollment for lottery applicants over the 2016-18 period.

The administrative early childhood records have two limitations: they exclude private pre-kindergarten without state subsidies, and they miss some major Head Start providers that operate as private organizations with federal funding rather than through state channels. One of the reasons we conduct a parent survey is to address these issues.

3.3 Parent earnings data

We worked with NHPS and the Connecticut Department of Labor (DOL) to link parent records from NHPS UPK applications to Connecticut Unemployment Insurance records for 1999-2022. This match was conducted by state personnel using name and address information. Address information on the state side comes from the Department of Motor Vehicles. We keep only unique matches and conduct extensive checks of the quality and representativeness of matched data. We deflate earnings to 2015 dollars and topcode at the 99th percentile within bins defined by the year of application and the year-by-quarter in which earnings are reported.[5]

Panel D of Table 1 describes the parent data. All applicants must list at least one parent or guardian. In 2013 and earlier, the application process was done mainly on paper and applications rarely listed more than one contact. In 2014 and later, the process played out primarily online, and applicants were prompted to list multiple contacts and describe the relationship between the contact and the student. In the full dataset, 33% of applicants list two parents on their application. After 2013, that share rises to 56%, which is similar to the 63% of children five and under in New Haven County who are observed living in two-parent households in the 2019 ACS. Where one contact is listed, that contact is usually the mother. In the post-2013 data, 69% of listed contacts are mothers. Of the 16,037 UPK applicants, we match 9,162, or 57%, to earnings records for at least one parent. As shown in Appendix Figure A.1, match rates to earnings are higher for more recent cohorts.

Column 3 of Table 1 reports descriptive statistics for applicants and parents in the parent earnings sample. This sample is similar to the full sample in terms of child age, gender, test scores, and block-group income, though with somewhat fewer Hispanic students (21.9% vs. 28.5% in the full sample). Baseline labor force participation is high, with 79% of parents reporting positive annual income in the two academic years prior to application. Mean baseline individual earnings is $25,157.

Online Appendix F provides additional evidence that our merge procedure generates accurate matches. For example, we show that earnings for mothers drops relative to

[5]Omitting the topcoding step does not change our findings.

earnings for fathers following the birth of a child, as in (Kleven et al., 2019). We would expect this test to fail if we had matched parents to the wrong earnings records. Other than the accuracy of successful matches, the main concern related to the parent merge is that matches may be imbalanced with respect to UPK assignment. We discuss this concern in Section 4 and find no evidence that matching to the earnings sample is related to placement in a UPK program.

3.4 Parent surveys

We supplement the administrative data with a survey of parents of past applicants to capture outcomes not in our administrative data. Working with NHPS and NORC, we fielded the survey between May and November of 2023. Using contact information provided on the application form, we emailed the parents of all past applicants, with phone follow-up for non-completers or those without email addresses. To maximize the statistical power of experimental analyses using survey data subject to our budget constraint, we focused phone follow-up on parents whose children faced an interior probability of admission based on their application and priorities. Online Appendix F describes survey procedures in detail.

Out of all the parents we contacted, 5.7% responded to our survey, yielding 966 survey responses. Excluding applicants to Covid-affected cohorts for whom in-person UPK was not available and late applicants not subject to randomization (see Section 4.1 below) leaves us with 840 responses in our analysis sample and a response rate of 5.3%. This overall rate masks a sharp increase over time. As shown in Appendix Figure A.2, response rates rise for recent cohorts, reaching 23% for 2022 applicants. Rates are likely low for early cohorts because contact information provided on old applications is no longer current or because past applicants become less likely to respond to school-choice-related communications as they move farther from the application date.

As reported in column 4 of Table 1, survey respondents tend to come from higher SES backgrounds than the applicant population as a whole. They are less likely to be Black (30.5% vs. 41.8% in the full sample), more likely to be White (31.4% vs. 21.7% in the full sample), and live in Census block groups where median household income is about 10% higher than in the full sample. They were also more likely to be assigned to a school (41.9% vs. 25.1%) and to enroll (54.6% vs. 26.5%). As we show below, none of the survey findings we report are affected by reweighting the survey sample so that it matches the observable characteristics of the full sample. Most respondents (89.7%) are women, with 86% reporting labor force participation, which is nearly identical to what we observe in the baseline administrative records. Respondents who work report spending an average of 33.5 hours per week on the job.

Survey results appear to be reliable. For example, enrollment reports from survey data are consistent with the enrollment choices we observe in administrative records. Further, there is a strong positive relationship between survey reports of household income and the median household income in the block group where the respondent lived at the time of their UPK application. See Online Appendix F for details.

The survey provides richer data on non-UPK childcare usage than is available in the administrative data alone. We observe that program types and out-of-pocket (OOP) costs vary across the income distribution. Panel (a) of Figure 2 reports the kinds of pre-kindergarten programs that children not enrolling in UPK programs attend, split by quintile of block group median household income. The Head Start share of non-UPK enrollees declines from 35.3% in the bottom quintile to 7.4% at the top, while the private pre-kindergarten share grows from 39.2% to 80.9%. The share of children not in a formal childcare or pre-kindergarten program also declines, from 19.6% in the bottom quintile to 3.2% in the top quintile. Panel (b) of Figure 2 shows that out-of-pocket costs grow with block-group income, rising from $286 per month in the bottom quintile to $942 per month in the top quintile.

As was the case with parent earnings records, a possible concern about the survey data beyond the quality of observed responses is that survey response may be correlated with placement in a UPK program. We discuss this in Section 4.

4 Results

4.1 Empirical design

We estimate instrumental variables (IV) specifications of the form

$$Y_i = \beta D_i + \sum_p \alpha_p 1[P_i = p] + X_i'\Gamma + \epsilon_i$$
$$D_i = \delta Z_i + \sum_p \rho_p 1[P_i = p] + X_i'\pi + \eta_i. \tag{1}$$

Here, Y_i is an outcome of interest. D_i, the endogenous regressor, is an indicator equal to one if a child enrolls in a UPK program. We instrument for D_i using Z_i, an indicator equal to one if an applicant is assigned to a UPK program in the choice process.

The key control variables here are indicators for each value of the assignment propensity P_i (in 0.01 bins), interacted with application year and grade indicators. These blocks identify students facing the same assignment propensity within a given grade-level and year, and controlling for them isolates the random variation in treatment assignment (Abdulkadiroglu et al., 2005; Rosenbaum and Rubin, 1983). We obtain assignment propensities through simulation, re-running the assignment algorithm in each grade and year with new random lottery draws. We also include controls for race/ethnicity, gender, child age in September following the application year, and measures of neighborhood block-group attributes from the ACS, denoted X_i.[6]

We modify this approach across outcome types and datasets to accommodate differences in sample size, data structure, and data availability. For parent earnings outcomes, we use panel data identified by parent, application, and academic year. We add interactions between the P_i, as defined above, and indicators for years elapsed since the time of application. We also include controls for baseline income and two-way cluster standard errors at the application (to account for possible correlations between parents of the same child over time) and parent (to account for cases where the same parent shows up in multiple child records) levels. We estimate specifications using data for various time windows following the application year.

For outcomes from the survey data, we use a recentered instrument approach (Borusyak and Hull, 2023), dropping the propensity indicators and instrumenting with $\tilde{Z}_i = Z_i - P_i$. Like controlling directly for the assignment propensity, recentering isolates random variation in treatment assignment. The benefit of recentering in survey specifications is that we do not need to estimate many fixed effects in the much smaller survey sample.

Two sample restrictions are important to highlight. First, we drop late applications because they are not included in the lottery process. Second, we restrict our sample to limit the effects of Covid-19 school closures on our analysis. NHPS closed its schools in March 2020 and did not reopen elementary schools for hybrid instruction until January 2021. This re-opening provided pre-kindergarten students with at most four days of in-person learning per week (Hays, 2021). District schools resumed their normal schedule in Fall 2021. We therefore exclude all 2020 applicants and 2019 PK3 applicants from our parental labor market specifications.[7]

4.2 Balance and first stage

We begin by validating the randomized research design. Panel A of Table 2 tests if predetermined student attributes are balanced with respect to lottery assignment. We estimate reduced-form versions of Equation 1 with lottery assignment as the independent variable of interest and covariates as outcomes. Columns 1 and 2 report control complier and full control group means for the dependent variables.

A straightforward comparison of assigned and unassigned students reveals substantial demographic differences. Column 3 reports results from a regression *excluding* the

[6] Appendix Table A.3 provides a list of these covariates.

[7] The 2020 application cohort faced either fully remote or part-time instruction during their first UPK year. The 2019 PK3 application cohort was exposed to the shutdown in the second half of the first year, and then to remote or hybrid school for all of their second year.

controls for propensity scores. Students who receive UPK offers are less likely to be female. They are more likely to be White and to live in neighborhoods with higher median household income. Appendix Table A.3 reports balance test results for additional neighborhood characteristics. A joint test rejects the null that the coefficient on UPK assignment is zero in each of these specifications ($p < 0.001$).

As expected, adding the propensity score indicators as controls eliminates these differences. Column 4 reports these results. Differences in race and neighborhood income drop sharply and become statistically insignificant at conventional levels. A modest difference in female share remains, but we fail to reject the joint null that all coefficients are zero ($p = 0.526$). We regard this as strong evidence that, as expected, our approach succeeds in isolating the random component of UPK assignment.

We find little evidence that school assignments affect downstream match rates to other data sources. Panel B of Table 2 reports estimates from reduced-form specifications in which the outcomes are indicators for successful matches between the lottery data and each of these datasets. 89.5% of control compliers match to state student data, 75.7% match to parent earnings data, and 5.3% match to survey data. As reported in column 4, there is no evidence of differential match for assigned students into the state enrollment or parents' earnings datasets once we control for propensity scores. There is some evidence of differential response in the survey, with parents of placed students being 1.2 percentage points more likely to respond ($p = 0.055$), but we cannot reject the joint null of no differential selection into any of the state, earnings or survey datasets ($p = 0.133$).

We see no evidence of imbalance on predetermined covariates within our matched samples. Columns 5-7 of Panel A report balance tests within the state, parent earnings, and survey samples, respectively. In each case, we fail to reject the null that coefficients from all specifications are zero. For the parent earnings sample we additionally test for balance on the parent's pre-assignment income. We find no evidence of imbalance on income in dollars or an indicator for any positive income.

The first stage effects of assignment on UPK enrollment are strong. As reported in Panel C of Table 2, assignment raises the share of students enrolling in a UPK program by 0.390 (SE=0.013) in the full sample, conditional on propensity score. The first-stage F-statistic is 844.1. First stage effects are similar in the matched samples. The first stage in the smaller survey sample is 0.409 (SE=0.053), with an F-statistic of 63.0.

UPK assignment increases the number of years students attend a UPK program by 0.576 years in the full sample. This coefficient is 48% higher than the first-stage enrollment effect, consistent with the observation from Table 1 that roughly half of the applications are to PK3 programs that give students the option to stay in UPK for two years once they enroll.

4.3 The effects of UPK on childcare

Our administrative and survey data let us describe how UPK enrollment shifts children across different childcare options. This helps us interpret the UPK treatment. Our approach is to run IV regressions of the form shown in Equation 1, with various childcare outcomes as dependent variables. We present these results in Table 3.

We first consider the substitution patterns observed in administrative data sources. Column 3 of Table 3 reports results for all application cohorts matched to state records. We find that enrolling in UPK reduces the share of students enrolling in state-administered Head Start by 13.9 percentage points, the share enrolling in School Readiness by 7.6 percentage points, the share enrolling in Care 4 Kids by 5 percentage points, and the share enrolling in any other pre-kindergarten program appearing in SDE or OEC records by 17 percentage points. Overall, UPK enrollment raises the rate at which students enroll in any observed program by 39.3 percentage points. 62.4% of compliers would otherwise attend some form of public or subsidized care.[8]

Administrative data miss private non-subsidized enrollment and also omit major Head Start programs in some years. Survey data helps us fill this gap. As reported in column 4, enrolling in UPK reduces the share of students who enroll in a private or paid pre-kindergarten program by 62.3 percentage points. Combined with substitution from Head Start and other public programs, we find that UPK does not increase enrollment in pre-kindergarten or other childcare outside the home. Essentially all UPK students would otherwise attend another program.

Our finding that nearly all UPK students substitute away from other forms of childcare echoes Weiland et al. (2020)'s finding that nearly all applicants to Boston's UPK program have other center-based care as their outside option. It is also consistent with the observation that most four-year-olds in the US were enrolled in school of some kind over the period we study (Cascio, 2021). The implications for cost-benefit analysis are potentially large. For comparison, Kline and Walters (2016) finds that about a third of households offered enrollment in Head Start substitute away from other center-based care, and shows that accounting for this substitution is crucial to the overall cost-benefit evaluation of the Head Start program.

Though UPK enrollment does not affect *whether* children enroll in childcare, it does increase the hours of childcare coverage that parents can access. To see this, we compile center- and year-specific reports of daily childcare schedules, and place measures of available hours at the center a child attends on the left side of our IV specifications. Online Appendix G describes this process in detail.

As reported in Panel B of Table 3, UPK enrollment raises weekly childcare coverage

[8]Children may enroll in multiple pre-kindergarten programs, so the sum of program-specific substitution effects need not equal the rate of substitution away from any program.

by 11.3 hours. The weekly total equates to 2.26 additional hours per weekday, or roughly 570 hours when aggregated across the school year. The large treatment effects are due to the ten hours of coverage available each day at UPK programs, more than most alternative programs.[9] Findings are similar across a wide variety of approaches to constructing childcare schedule data (see Online Appendix G). We note that, because the survey sample skews towards recent applicants, and because some UPK sites reduced hours after 2021, these estimates may understate the gains in childcare coverage for parents across most of our sample period.

In addition to expanding childcare hours, UPK enrollment reduces families' out-of-pocket childcare costs. The last row of the last column of Table 3 reports the effects of UPK enrollment on monthly OOP costs. A caveat in this analysis is that, due to an error in survey logic, we did not ask the parents of students enrolling in magnet programs about their out-of-pocket costs. These costs are likely fairly low, since these students receive ten hours of free childcare each school day. In Table 3, we report results that assign students enrolling in UPK programs the average value of costs observed among Head Start students. We view this as likely overstating the costs for UPK students (and understating the effects of treatment), given that they receive more hours of free childcare. UPK reduces parents' childcare costs by \$375 per month, 77% of the control complier mean.

As reported in Appendix Table F.1, reweighting the survey sample to match the demographic characteristics of the full sample does not change our findings.

4.4 Children's educational outcomes

We find little evidence that UPK affects children's academic performance in the medium run. Figure 3 presents IV estimates for key educational outcomes.

Panel (a) shows effects on standardized assessments in kindergarten and grades 3-8. We cannot reject the null that effects are zero for any grade level. The point estimate for kindergarten scores is modestly positive (0.062σ; SE=0.072), while estimates for grades 3-8 are negative. Appendix Table A.4 confirms that there are no significant effects on any KEI subscores, including personal/social skills. Similarly, Panel (b) of Figure 3 shows no significant effects on chronic absenteeism (defined as missing ten percent or more of total school days in a year), and Appendix Figure A.3 shows null effects on grade retention (defined as a cumulative indicator for ever being retained).

These findings align with other randomized UPK evaluations that report null or negative effects in grades K-8 (Durkin et al., 2022; Gray-Lobe et al., 2023). However, they do not preclude positive effects in the long run. The effects of early childhood

[9] Note that the estimated complier control mean of 51.24 has a large standard error (13.0) and is far above the control mean of 28.9.

interventions often fade out after pre-kindergarten before returning in high school or later (Heckman et al., 2013), and previous randomized evaluations of UPK programs have demonstrated both the fade-out of large short-run effects (Lipsey et al., 2018) and long-run gains following medium-run null effects (Gray-Lobe et al., 2023). Alternatively, the null effects may simply indicate that NHPS programs offer similar educational quality to the alternative programs from which children substitute.

4.5 Parents' labor market outcomes

4.5.1 Earnings and labor supply

We now turn to our main results. Panels A and B of Table 4 report estimates of the IV specification in Equation 1 that take administrative records of parent earnings as the dependent variables of interest. We focus on earnings measured in dollars (including zero values) and an indicator for positive annual income. To provide a percentage interpretation for effects in dollars, we follow Chen and Roth (2024) and report results from a Poisson IV specification, which we compute using the control function approach outlined in Lin and Wooldridge (2019). This is our preferred specification. See Online Appendix H for details. All specifications control for demographics reported at the time of UPK application. Each specification also controls for the baseline value of the dependent variable, measured in the two academic years prior to the year of application. These controls are linear except in the Poisson regression, which includes indicators for each decile of the baseline earnings distribution.

Panel A of Table 4 reports results that use all available earnings data. The first row shows our findings for the years when children are of pre-kindergarten age. For PK4 applicants, these run from Q4 in the year of application, when the student would enroll for the first time, through Q3 of the next calendar year. For PK3 applicants, the pre-kindergarten years extend through Q3 two calendar years after the application.

We find that UPK enrollment raises parent income by \$5,464 (SE=\$1,718) while the child is of pre-kindergarten age. This is 15.9% of the control complier mean of roughly \$34,000. The Poisson IV specification indicates a 21.7% (SE=6.6%) earnings gain during this period. We find evidence of modest extensive-margin gains, with rates of positive earnings rising 5.7 percentage points (SE=2.6) on a base of 82.4 percent.

To understand how earnings effects evolve as children age out of pre-kindergarten, we compute separate estimates of the effect of UPK on enrollment for each two-year interval following on-schedule pre-kindergarten completion. For children following the standard grade progression, these will correspond to kindergarten and first grade, second and third grade, and so on. We report these findings in the lower rows of Panel A of Table 4. Figure 4 plots estimated percentage effects from the Poisson IV specification, pooling data from 9 or more years after pre-kindergarten. We use all available data for each regression, so sample sizes decline as we push farther from pre-kindergarten completion.

Earnings gains persist after children age out of UPK. Our preferred Poisson IV estimates are stable through six years after kindergarten, with earnings effects between 19% and 23% in each two-year interval. A pooled specification combining years 1-6 after pre-kindergarten yields an estimated earnings gain of 20.9% (SE=7.9%) over the full period.

After six years, standard errors grow. Pooling years seven and later in our Poisson specification yields estimated earnings gains of 9.4%, with a standard error of 21.3%. We cannot rule out null effects in years seven and later, but we also cannot rule out gains similar to what we observed in the first six years after pre-kindergarten.

Estimates from our other specifications follow similar patterns. Pooling over the first six years after pre-kindergarten, the earnings effect of UPK enrollment is $6,474 (SE=$2,260) in dollar terms. Rates of labor force participation rise by 4.8 percentage points (SE=0.028), 6.0% of the control complier mean. Similar to our Poisson specifications, standard errors increase substantially when we look beyond six years.

4.5.2 Alternate samples and specifications

A possible concern about specifications that use all available data is that they may conflate heterogeneity in treatment effects by time relative to pre-kindergarten with heterogeneity in treatment effects across treatment cohorts. Long lags relative to treatment are only observed for early cohorts. Panel B of Table 4 addresses this issue by reporting results for the pre-kindergarten years and the first four years afterward for a balanced panel of individuals we observe over this whole period. This sample excludes application cohorts after 2017 for PK3 and 2018 for PK4, reducing the sample size by 14% relative to our main specification.

Effects in the balanced panel are slightly smaller than in the full sample. For example, during the pre-kindergarten period, the Poisson effect is 15.3% (SE=6.8%) in the balanced sample. As in our main specifications, we see no evidence of fade-out in the four years following pre-kindergarten. In our Poisson specification, we find an effect of 16.4% (SE=7.3%) in the first two years after pre-kindergarten and 18.1% (SE=8.6%) in the third and fourth years.

Our findings are also robust to alternate IV approaches and alternate control sets. Appendix Table A.5 reports results that (a) use the recentered instrument approach from Borusyak and Hull (2023) rather than controlling for propensity scores, or (b) drop all demographic controls. We find the same pattern of large and sustained earnings gains in both cases; point estimates are slightly smaller under approach (a) and slightly larger under approach (b).

4.5.3 Mechanisms

What drives the large and sustained earnings effects of UPK? One hypothesis is sustained increases in labor supply. While extensive-margin labor supply effects observed in administrative data are substantially smaller than earnings gains in percentage terms, labor supply may rise on the intensive margin as well. Panel C of Table 4 presents results that take survey reports of weekly work hours in the year following the child's UPK application (i.e., the first year of UPK for enrolling students) as well as reports for the second year following application for PK3 applicants. Weekly hours rise by 12.80 (SE=4.25) on a control complier base of 27.9. We cannot reject the null hypothesis that this value is equal to the 11.3 hours of additional weekly childcare that we estimate the UPK programs provide, as reported in Table 3. Labor supply effects decline after the pre-kindergarten years, though estimates are noisy enough that we cannot rule out sustained gains. The estimated effect of UPK enrollment on work hours in the years after pre-kindergarten is 1.48 (SE=3.91). Appendix Table F.1 shows that reweighting the survey sample to match the population does not affect these findings.

Our interpretation is that there is strong evidence of labor supply gains *during* pre-kindergarten. In contrast, while some increase in labor supply may persist *after* pre-kindergarten, there is little to suggest that long-run increases in labor supply are sufficient to explain the long-run earnings effects we observe.

Our earnings estimates are also larger than might be expected based on a pure returns-to-experience model in which labor supply increases during pre-kindergarten pay off in subsequent years. For example, early-career estimates of the returns to a year of work experience in the NLSY from Deming (2023) range from 2.5% to about 10%, depending on worker skill.

Another hypothesis is that earnings gains may arise because UPK reduces career disruption. We explore career disruption effects using administrative data. We do not observe employer identifiers, so we cannot identify job spells. However, we do observe job-specific earnings and industry codes in each quarter, so we can measure the consistency of labor force attachment within the academic year, the degree to which individual earnings come from one job as opposed to multiple jobs, and changes in industry. We use these variables as outcomes in Equation 1.

Table 5 reports our results. The first column of Table 5 takes an indicator for switching main industry since the prior academic year as the outcome of interest, which we interpret as a simple measure of career disruption. Main industry is defined as the industry in which the individual has the most earnings. We find that UPK enrollment reduces the share of individuals who switch industry during the pre-kindergarten years by 7.2 percentage points (SE=2.7), 34% of the control complier mean of 0.21. Industry switch effects fade to zero immediately following the completion of pre-kindergarten.

Column 2 of Table 5 examines the effects of UPK enrollment on the number of quarters per year in which individuals hold one "main job," defined as having exactly one job and that pays at least $4,000 each quarter, roughly equivalent to a full-time minimum wage job. This is a simple measure of job attachment. During the pre-kindergarten years, UPK enrollment raises the number of quarters individuals hold a single main job by 0.300 (SE=0.108), 15% of the control complier mean. As with industry-switching, this main job effect falls when pre-kindergarten ends.

Columns 3 and 4 of Table 5 consider the effects of UPK enrollment on the count of low-earning quarters per year (column 3) and the total number of low-earning quarters up to and including the current academic year (column 4). We define a low-earning quarter as one where the person earns less than $4,000. The goal here is to measure how career disruptions accumulate over time. We find that UPK enrollment reduces the number of low-earning quarters by 0.206 per year (13.8% of the control complier mean) in the pre-kindergarten years, with the cumulative reduction in low-income quarters reaching 0.76 between three and four years after pre-kindergarten before declining back towards zero in years five and six.[10]

The large, sustained earnings effects we find are similar to estimates of the effects of career disruptions reported in prior studies, but with the opposite sign. For example, Bertrand et al. (2010) report that MBA graduates with at least six months of non-work at any point between their MBA year and the year in which earnings are reported earn 17-29% less than other observably comparable students. Studying a very different population, Jacobson et al. (1993) find that high-tenure workers displaced from distressed firms see their earnings fall by about 25% over the long run.

Our findings are consistent with qualitative reports that UPK reduced career disruptions. For example, one parent in our survey who was not selected for a UPK spot described how with UPK, "I could have worked a different full time schedule—more normal hours. My husband and I ended up doing what felt like shift work [so that] I could pick up [my child] at 3:00 when the private preschool program ended." See Online Appendix I for more discussion.

4.6 Heterogeneous effects

4.6.1 Neighborhood income

Middle-income families appear to be the biggest beneficiaries of the UPK program. To see this, we split our sample into three groups based on terciles of median household income in the Census block group where the child lived at the time of application.

[10]Because the pre-kindergarten enrollment period can be either one or two years and because the other blocks consist of two-year periods, the cumulative effect coefficients do not equal the sum of the contemporaneous effect coefficients.

We use this measure rather than individual income because mother's income between childbirth and pre-kindergarten may not be a useful proxy for household income in the long run. We estimate Equation 1 within each group for childcare, education, and earnings outcomes. We report our findings in Table 6.

A preliminary point is that the UPK program draws students mainly from low- and middle-income families. Even the higher-income families in our analysis fall in the middle of the population distribution. As reported in Table 6, the mean individual income for control compliers in the second tercile during pre-kindergarten is $34,324 per year. For control compliers in the third tercile, it is $43,153. Median individual income for workers in the US was about $39,000 (after converting to 2015 dollars).

Panel A of Table 6 reports first-stage effects by income tercile. We observe a strong first-stage effect in each tercile. Point estimates are somewhat smaller in the third tercile (0.355, compared to 0.419 in the first tercile) but F-statistics are uniformly high.

Panel B of Table 6 reports how UPK enrollment affects childcare usage in each income tercile. Focusing first on hours, we find that point estimates for weekly gains are larger in the middle and top terciles (11 and 12 hours) than in the bottom tercile (6 hours). Though standard errors in the split survey sample are fairly large, these estimates suggest that gains in access to childcare coverage persist through relatively high income levels.

Turning to substitution across programs, we find that all three income groups substitute away from other forms of outside-the-home childcare when they enroll in UPK, with no evidence of effects on the extensive margin for any income tercile. We do see differences in the kinds of programs from which students in different terciles substitute. Bottom-tercile students enrolling in UPK are more likely to substitute away from Head Start according to both survey and administrative records. Middle- and top-tercile students mainly substitute away from other paid options. Consistent with these findings, UPK reduces childcare costs more for higher-income applicants.

Differential substitution patterns translate to differences in educational treatment effects. Panel C of Table 6 reports these results. UPK enrollment raises test scores for middle-tercile students by 0.29σ. Effects are negative (though noisily estimated) in the bottom tercile and close to zero in the top tercile. Within terciles, effects are similar across the different KEI subscores. See Appendix Table A.4.

Following students past Kindergarten, we observe a classic fade-out/fade-in pattern for the middle tercile, with score effects going to zero in elementary grades and then rising again in grades seven and eight. For the upper- and lower-tercile groups, we see no evidence of score gains in any grade. By seventh and eighth grade, point estimates for upper- and lower-tercile students are negative and economically large, though also statistically noisy. We cannot rule out null effects on absenteeism and grade retention for any tercile in any grade. See Appendix Figure A.5.

One story consistent with our findings on substitution patterns and score gains is that there is a "doughnut hole" in access to high-quality pre-kindergarten at middle income levels. Low-income students have access to subsidized care, and higher-income students can access high quality private programming. In the middle of the income distribution, subsidized care is limited and high-quality private options are out of reach.

Parents' earnings effects are also heterogeneous across the income distribution. Panel D of Table 6 reports these findings for our Poisson IV estimates. We find earnings gains in the neighborhood of 25% for middle- and top-tercile families during pre-kindergarten. Earnings gains for bottom-tercile families have a point estimate of 10.6%, and we cannot rule out a null of no effect at conventional levels. After pre-kindergarten, effects remain high for top- and middle-tercile families for at least six years, although standard errors are, in some cases, large. Effects for the bottom tercile are again zero. Our results are consistent with findings that gains from expanded work hours are larger for (relatively) higher-skilled workers, who may be in better position to climb the career ladder (Kuhn and Lozano, 2008; Cortés and Pan, 2019).

4.6.2 Other demographic categories

Table 7 reports the effects of UPK on parents' earnings by race, relationship to child, and family structure using our preferred Poisson specifications. Columns report results for different samples in the time period listed in the horizontal panel.

The first three columns report results by race and ethnicity. We find limited evidence of earnings gains for parents of Black students and large effects for parents of White students. Parents of Hispanic students fall in between. These results mirror our finding that earnings effects are larger at higher income levels: complier means are highest for parents of White students, lowest for parents of Black students, and in the middle for parents of Hispanic students.

Columns 4-9 of Table 7 restrict the sample to applications after 2013, when the application form encouraged listing multiple adults and recorded the relationship between the adult and the child. Column 4 shows that estimates from the post-2013 sample are very similar to estimates from the full sample. Splitting by the relationship between the adult and the child (columns 5 and 6), we find evidence of sustained gains for both mothers and fathers. We cannot reject the null that effects for mothers and fathers are equal during ($p = 0.53$) and after pre-kindergarten ($p = 0.77$).

The finding that UPK causes fathers to earn more may be surprising given that past research on the labor market effects of childcare focuses almost exclusively on mothers. Important context for this result is that the UPK applicant pool draws disproportionately from families where both parents work and where gender gaps are small relative to previous studies and to the local population. For example, Gelbach (2002) reports a 51% labor force participation rate for married mothers and 62% rate for single mothers in his sample, drawn from the 1980 Census.[11] Using CPS data from 1999-2022, Gibbs et al. (2024) reports a labor force participation rate of 67% for mothers, compared to 88% for fathers, and reports that fathers earn 157% more than mothers on average. By contrast, the labor force participation rate for mothers in the UPK applicant pool is 80%, nearly identical to the 81% for fathers, and the gap in average earnings is much smaller, at 59%. The story is similar when we compare two-parent families in our applicant pool to the population of two-parent families with young children in New Haven County: earnings gaps are much smaller in our sample. See Online Appendix J for details. A hypothesis consistent with both the observed treatment effects and these descriptive data is that the UPK applicant pool consists largely of families where both parents participate in work and childcare.

In the 7th and 8th columns, we split by observed family structure: specifically, whether one or two parents are listed on the application form. We cannot reject the null that effects for the two groups are equal during pre-kindergarten ($p = 0.63$). After pre-kindergarten, gains for one-parent families are typically larger, though again we cannot reject the null that effects are equal ($p = 0.66$). Lastly, the 9th column reports impacts on household income for the subset of data where all listed parents are matched to the earnings records. The estimated increases in household income are similar to our main estimates in percentage terms.

5 Cost-benefit analysis

Our results show that UPK programs increase parents' earnings and reduce out-of-pocket costs for families. However, UPK is costly. We compare the costs and benefits of the program using the marginal value of public funds (MVPF) framework from Hendren and Sprung-Keyser (2020). The MVPF equals the ratio of beneficiaries' willingness to pay to net program costs ($\frac{\Delta W}{\Delta E - \Delta C}$), where ΔW represents beneficiaries' willingness to pay for the program, ΔE measures upfront program costs, and ΔC measures reductions in government expenditures or increases in government revenue due to the long-run effects of the policy. We compute ΔW, ΔE and ΔC using estimates from our randomized design. See Online Appendix K for details on the calculations.

[11] Gelbach (2002) reports rates separately for mothers whose youngest child is five and mothers whose youngest child is under 5; these statistics are population-weighted averages over the two groups.

5.1 Government expenditures

We begin by examining how UPK affects net government costs, $\Delta E - \Delta C$. The gross cost of the program, ΔE, is approximately \$24,000 per child, which is the yearly per-pupil expenditure (PPE) of \$15,500 times the average of 1.56 years enrolled.[12]

Gross costs to the government are offset by two kinds of cost reductions. The first type is substitution away from other subsidized programs. As shown in Table 3, many applicants substitute away from other publicly funded programs. To capture how this affects public costs, we estimate the change in years of enrollment in other publicly funded programs, and scale these changes by their per pupil expenditures.[13] We estimate that program substitution reduces net costs by \$8,800.

The second type of cost reduction comes from increased tax revenue from parents' and children's earnings gains. For parents, we calculate the gains in discounted after-tax income using the impacts on individual wage income reported in Table 4. We conservatively assume no impact beyond six years after the UPK program ends, and scale individual earnings effects by 1.56, the average number of adults per family in the years where multiple members are systematically observed (see Table 1). Assuming an effective tax rate of 0.20, this generates \$13,400 in additional tax revenue.[14] For children, we project earnings gains based on kindergarten test scores following Chetty et al. (2011), yielding \$430 in additional tax revenue.

Putting these numbers together shows that tax revenues from increased parent earnings dramatically reduce the net costs of UPK. Panel (a) of Figure 5 illustrates this point. Starting from \$24,200 in upfront costs, accounting for program substitution reduces net costs to about \$15,400, and accounting for increased future tax revenue from children further reduces costs to about \$15,000. Tax revenue from parents' increased earnings then reduces net costs by \$13,400, down to \$1,500, or about 6% of upfront costs.

Put another way, accounting for the fiscal externality from parents' additional earnings reduces the net government costs of UPK by 90% relative to estimates that ignore earnings gains for parents. It therefore raises the MVPF by a factor of 10 for any fixed value of willingness to pay. This finding is even more pronounced under alternate assumptions about children's earnings gains drawn from the literature. For example,

[12] The PPE value comes from the average PPE among NHPS elementary schools during the 2018-2019 school year (in 2015 dollars). For comparison, the PPE for Head Start in Connecticut is \$10,100 and the average PPE for all public pre-k programs in the state is \$9,600 (Friedman-Krauss et al., 2022, 2023).

[13] We focus on Head Start, School Readiness, Care 4 Kids, and other public state programs for the 2015-2017 application cohorts. Per-pupil expenditures for these programs are from Friedman-Krauss et al. (2022) and Friedman-Krauss et al. (2023). See Online Appendix K for details.

[14] We follow Hendren and Sprung-Keyser (2020), who use calculations from Congressional Budget Office (2016) and find that the effective tax rate is close to 0.2 for those between 100 and 400 percent of the poverty line.

taking Lipsey et al. (2018)'s estimates of a 0.4σ score gain from Tennessee's UPK program yields a projection of $2,800 rather than $400 in additional tax revenue.[15] Alternatively, using Gray-Lobe et al. (2023)'s estimate of UPK's effect on college enrollment combined with returns estimates from Zimmerman (2014) yields $2,500 in additional tax revenue. Under either approach, including parent income in the cost calculation leads to negative net government costs (see Online Appendix K.2 for details). As reported in Table 8, the result is an infinite MVPF regardless of willingness to pay.

5.2 Willingness to pay

We next consider ΔW, the willingness to pay for the benefits that accrue to parents and children. Because UPK is an in-kind transfer, computing willingness to pay presents a conceptual challenge: to what extent are the welfare benefits of increased earnings offset by the disutility of work? Under the envelope theorem, changes in optimized variables are not welfare-relevant, so the key question is which outcomes parents take into account and what constraints they face when making childcare decisions. We consider several approaches for computing WTP, which we believe bracket the true effect.

Our benchmark approach is to assume families value UPK at the net cost of provision, paralleling procedures for calculating MVPFs of pecuniary transfers (Hendren and Sprung-Keyser, 2020). An attractive feature of this strategy is that the MVPF would be equal to one in the absence of any fiscal externalities. The results therefore illustrate how reductions in net costs due to increased tax revenue shape MVPF calculations for a given value of WTP.

The alternative is to take a hedonic approach that attempts to quantify the value of different elements of the UPK bundle to families. We consider three channels through which families may benefit. The first is changes in out-of-pocket childcare costs. We assume that families value cost reductions dollar-for-dollar. This assumption likely understates the value of UPK to parents since it offers services, particularly extended hours, that other options may not fully replace. Second are children's after-tax future earnings, which past work includes in WTP calculations (Cascio, 2023; Hendren and Sprung-Keyser, 2020). As we show in Online Appendix K, children's after-tax earnings should be included in WTP calculations if parents either cannot purchase substitutes for full-day UPK on the private market or do not incorporate future earnings gains for children when making choices about childcare and labor supply.

The final channel we consider is parents' after-tax earnings. As with earnings gains for children, whether one should include parents' earnings gains in WTP depends on the underlying economic model. If parents can access private-market substitutes for full-

[15] This is similar to Cascio (2023)'s approach to computing the MVPF of UPK, but uses score estimates based on Lipsey et al. (2018)'s randomized design.

day UPK and make labor supply choices optimally, earnings gains will be fully offset by the disutility of work and should be excluded from WTP. However, if parents lack access to extended-day care in the private market, they may value additional earnings up to dollar-for-dollar. A third possibility is that parents make optimal short-run choices without fully internalizing returns to job continuity or full-time work (Blesch et al., 2023; Costa-Ramón et al., 2024), suggesting we should include future but not contemporaneous earnings gains. See Online Appendix K for details.

The data provide some guidance on how parents' earnings gains should enter WTP. In particular, there is evidence that parents face constraints on access to full-day care. For example, the hours of care available at non-UPK childcare centers are limited (see Section 4.3 and Online Appendix G), and parents who are not enrolled in UPK report challenges balancing work and childcare (see Online Appendix I). In addition, childcare is challenging, and the disutility of market work should likely be measured relative to this benchmark rather than leisure in the population we study.

Because we regard these points as suggestive rather than definitive, we present WTP estimates based on three different approaches: including all parent earnings, excluding all parent earnings, and including only earnings realized after pre-K.

Measurement proceeds as follows. To compute the reduction in out-of-pocket costs, we take our cost results from Table 3 and assume that this reduction applies for nine months per year. For parents' earnings and children's earnings we use the procedures described above in our discussion of net costs, now focusing on after-tax income. Panel (b) of Figure 5 reports the value of each component of willingness to pay discussed here.

5.3 MVPF results

Table 8 reports our estimates of the MVPF of UPK. Our first finding is that the MVPF is high regardless of how we compute willingness to pay. The upper-left cell reports our benchmark estimate, based on the assumption that parents value UPK at cost. We find that the MVPF of UPK is 10.1 (90% CI=$[1.87, \infty]$): the policy generates ten dollars in benefits per dollar of government expenditure.

Moving to the right along the first row, we present MVPF calculations under different assumptions about how families value the program. Even our most conservative approach, which excludes parent earnings from WTP and values UPK as the sum of cost reductions and children's earnings gains, yields an MVPF of 4.55 (90% CI=$[0.58, \infty]$). Adding parents' after-tax earnings raises the MVPF to 39.81 (90% CI=$[4.03, \infty]$). Including only the share of parents' earnings gains realized after pre-kindergarten yields an MVPF of 32.88 (90% CI=$[3.16, \infty]$).

Our second finding is that if we incorporate more optimistic assumptions about gains for children, based on Lipsey et al. (2018) or Gray-Lobe et al. (2023), MVPFs rise further and in fact become infinite regardless of how we compute willingness to pay. This is because larger gains for children push our estimates of net government costs below zero. We report these findings in the second panel of Table 8.

Our third finding is that measuring earnings gains for parents is crucial to understanding the return on UPK investments. As reported in the bottom panel of Table 8, when we exclude parent earnings from both WTP and net costs (following most previous analyses of UPK) we find an MVPF of 1.03 using our cost-based approach, and 0.46 using our hedonic approaches.[16] In contrast, earnings gains for parents alone yield high MVPF values, even if one ignores gains for children. Using cost-based WTP, the 'no-kids' MVPF is 7.87 (90% CI=$[1.82, \infty]$). Hedonic values range from 2.67 under the conservative OOP-only WTP to 30.14 when all parent income is included in WTP.

5.4 Comparison to other programs

From a policy perspective, the crucial question is how gains from UPK compare to other potential uses of the funds. Figure 6 reproduces a key plot from Hendren and Sprung-Keyser (2020), adding points for the "no-parent" and "no-kid" MVPF estimates for New Haven's UPK program and Cascio (2023)'s UPK MVPF estimate, which excludes parents.[17] The horizontal axis is the age of the beneficiary, and the vertical axis is the value of the MVPF. The point that Hendren and Sprung-Keyser (2020) make using their version of this plot is that there are many child-focused programs that yield high and even infinite MVPFs. These include increases in school spending, health insurance for children, and expansions in college access, among others. In contrast, relatively few programs targeted at adults yield high MVPFs. In particular, many labor market programs for adults, including work incentives such as the EITC and recent increases in top tax rates, have MVPFs near one.

Placing the UPK MVPFs on this graph makes two points. The first is that an evaluation of UPK that excludes parents' earnings does not offer a high MVPF compared to other child-focused policies. The benefits for children from UPK are relatively modest regardless of how they are calculated, and other child-focused options offer high returns.

The second is that, even excluding gains for children, UPK offers high returns compared to other policies targeting adults. The "no-kid" UPK MVPF (i.e., excluding gains for children) is higher than the MVPF for all the job training policies, all the unemployment insurance policies, and all the cash transfer policies (the category which includes

[16]This result holds even under optimistic assumptions about children's gains based on Lipsey et al. (2018) or Gray-Lobe et al. (2023). While these alternate scenarios raise MVPFs slightly, they remain close to one when excluding parent earnings. See Appendix Table K.4.

[17]Cascio (2023) reports MVPFs under various assumptions about program costs. The estimate we plot assumes that UPK program outlays are equal to K-12 spending on a per-pupil basis because that assumption most closely matches our analysis of the NHPS program.

EITC, Paycheck plus, and Negative Income Tax programs) evaluated in Hendren and Sprung-Keyser (2020). Evaluated as an active labor-market program, extended-day UPK performs well.

6 Discussion

Credit constraints and the market for childcare We show that parents' earnings returns from UPK substantially exceed the costs of childcare provision. This presents a puzzle: why can't parents randomized out of NHPS UPK purchase a similar product on the private market? In principle, it would seem possible for private providers to offer comparable programs at prices between the cost of provision and parents' earnings returns.

One plausible answer to this question is that parents face credit constraints. As discussed in Cameron and Taber (2004) in the context of higher education, credit-constrained individuals may forgo profitable human capital investments if returns accrue over time but require upfront financing. We find that investments in childcare have this structure: most of the earnings payoff for parents is realized after pre-kindergarten is over. In terms of credit supply, we would expect constraints to bind more tightly in childcare markets than in higher education markets. Banks do not offer loans for childcare expenses (US Department of the Treasury, 2021).

Our findings are quantitatively consistent with a role for credit constraints. We estimate that enrolling in UPK increases after-tax family income by \$10,549 during pre-kindergarten, nearly identical to the \$10,174 in additional childcare costs. Given these averages, many families likely face childcare costs exceeding their contemporaneous earnings gains.

To sum up, our finding that UPK access pays off over the long run—not just while children are enrolled in pre-kindergarten—helps rationalize the absence of privately-provided substitutes. If credit constraints do bind, they also provide a rationale for transfers to parents of young children that is not captured by our MVPF analysis above—namely, that the marginal value of consumption for this group is high. See Online Appendix K for details.

Policy design and the quality-quantity tradeoff States seeking to expand UPK at a given budget face a tradeoff between program quality and hours of coverage (Povich, 2024). Our findings indicate that UPK programs that span the work day can have large economic returns. For quality-focused investments to generate comparable returns, they would need implausibly large effects on children: UPK would need to raise test scores by 1.9σ—about 2.5 times the IQ gains from Perry Preschool (Heckman et al., 2013) and triple the largest estimates of UPK effects in the literature (Cascio, 2023). Alternatively, quality-focused investments would need to substantially increase other productive traits, such as non-cognitive skills.

Comparison to other childcare interventions We find large earnings effects for parents. This contrasts with prior studies of state-level UPK programs, which typically cannot rule out null effects. While our findings are consistent with the hypothesis that the extended-day program we study has larger impacts than state programs offering fewer hours, lack of statistical precision in prior studies makes it hard to rule out the possibility that those programs also had large effects. For example, Fitzpatrick (2010), cited by Cascio (2021) as the most convincing evidence, estimates an intent-to-treat effect of UPK eligibility on mothers' annual earnings of \$332 (SE=\$578). With a first-stage effect of 0.072 and converted to 2015 dollars, this implies an IV estimate of \$6,531 (SE≈\$11,400). This point estimate is larger than ours, and its 95% CI includes large positive and negative values.

The contemporaneous earnings effects we estimate for parents are similar to those attributed to universal childcare in Quebec, but children in our setting do better. Lefebvre and Merrigan (2008) and Lefebvre et al. (2009) study the expansion of childcare programming in Quebec and find that childcare access raises mother's earnings by \$3,000 to \$6,000 annually on a base of \$30,000 to \$40,000 and that gains persist after children age out of the program. The major downside of the Quebec program, which also offered workday-spanning childcare coverage, was that it substantially reduced long-run well-being for children (Baker et al., 2008, 2019). Our findings show that combining high-quality educational inputs with extended childcare coverage can yield high economic returns by improving outcomes for parents without compromising outcomes for children.

Evaluations of kindergarten expansion also generally find sizable labor market effects for mothers. Gelbach (2002) studies the expansion of kindergarten programming in the US and finds that kindergarten enrollment raises contemporaneous earnings for mothers by 24% Gibbs et al. (2024) studies the impacts on maternal labor supply from increasing kindergarten hours, using variation across states and over time from 1992-2022. They find that moving from part- to full-day kindergarten increases maternal employment by 4.5 percentage points, with both increases in part-time and full-time work.

7 Conclusion

This paper examines how randomized assignment to a high-quality, extended-day UPK program in New Haven, Connecticut affects outcomes for parents and children.

While effects on children are limited, we find that UPK enrollment raises parent earnings by 21.7% during the pre-kindergarten years, and that these gains persist for at least six years following pre-kindergarten completion before fading out.

A cost-benefit calculation incorporating parents' earnings, academic gains for children, and cost offsets from substitution away from alternate pre-kindergarten programs shows that the return to UPK enrollment is high, and that this is mostly due to increases in parents' earnings. Evaluated as an active labor-market policy, the UPK program performs well. We conclude that UPK programs combining high-quality educational inputs with childcare coverage that spans the workday can yield high economic returns and that it is crucial to think about parents when designing and evaluating early-childhood policies.

References

Abdulkadiroglu, A., P. A. Pathak, and A. E. Roth (2005). The New York City High School Match. American Economic Review 95(2), 364–367.

Akbarpour, M., A. Kapor, C. Neilson, W. Van Dijk, and S. Zimmerman (2022). Centralized School Choice with Unequal Outside Options. Journal of Public Economics 210, 104644.

Attanasio, O., R. Paes de Barros, P. Carneiro, D. K. Evans, L. Lima, P. Olinto, and N. Schady (2024, June). Public childcare, labor market outcomes of caregivers, and child development: Experimental evidence from brazil. Working Paper 30653, National Bureau of Economic Research.

Baker, M., J. Gruber, and K. Milligan (2008). Universal Child Care, Maternal Labor Supply, and Family Well-being. Journal of Political Economy 116(4), 709–745.

Baker, M., J. Gruber, and K. Milligan (2019). The Long-run Impacts of a Universal Child Care Program. American Economic Journal: Economic Policy 11(3), 1–26.

Bertrand, M., C. Goldin, and L. F. Katz (2010). Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors. American Economic Journal: Applied Economics 2(3), 228–255.

Blesch, M., P. Eisenhauer, P. Haan, B. Ilieva, A. Schrenker, and G. Weizsäcker (2023). Biased Wage Expectations and Female Labor Supply. Technical report, Ludwig Maximilian University of Munich.

Bonanno, M. (2023, October). Personal communication.

Borusyak, K. and P. Hull (2023). Nonrandom Exposure to Exogenous Shocks. Econometrica 91(6), 2155–2185.

Boston Public Schools (2024). Universal Pre-Kindergarten. Available at https://www.bostonpublicschools.org/page/6521. Accessed March 7th, 2024.

Brown, J. H. (2018). Does public pre-k have unintended consequences on the child care market for infants and toddlers? Technical report, Princeton University, Industrial Relations Section.

Bruhn, J. and E. Emick (2023). Lottery Evidence on the Impact of Preschool in the United States: A Review and Meta-Analysis. Technical Report 2023.20, Blueprint Labs.

Cameron, S. V. and C. Taber (2004). Estimation of Educational Borrowing Constraints using Returns to Schooling. Journal of Political Economy 112(1), 132–182.

Cascio, E. U. (2009). Maternal Labor Supply and the Introduction of Kindergartens into American Public Schools. The Journal of Human Resources 44(1), 140–170.

Cascio, E. U. (2021). Early Childhood Education in the United States: What, When, Where, Who, How, and Why. Working Paper 28722, National Bureau of Economic Research.

Cascio, E. U. (2023). Does Universal Preschool Hit the Target?: Program Access and Preschool Impacts. Journal of Human Resources 58(1), 1–42.

Cascio, E. U. and D. W. Schanzenbach (2013). The Impacts of Expanding Access to High-Quality Preschool Education. Technical Report w19735, National Bureau of Economic Research.

Chaparro, J., A. Sojourner, and M. J. Wiswall (2020). Early Childhood Care and Cognitive Development. Working Paper 26813, National Bureau of Economic Research.

Chen, J. and J. Roth (2024). Logs with Zeros? Some Problems and Solutions. The Quarterly Journal of Economics 139(2), 891–936.

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.

Congressional Budget Office (2016). Effective Marginal Tax Rates for Low- and Moderate-Income Workers in 2016. https://www.cbo.gov/publication/50923. Accessed: 1/23/2024.

Cortés, P. and J. Pan (2019). When Time Binds: Substitutes for Household Production, Returns to Working Long Hours, and the Skilled Gender Wage Gap. Journal of Labor Economics 37(2), 351–398.

Costa-Ramón, A., U. Schaede, M. Slotwinski, and A. A. Brenøe (2024). (Not) thinking about the future: inattention and maternal labor supply. Working paper series/department of economics, University of Zurich.

Davidoff-Gore, S. (2017). Assessing School Choice as a Mechanism for Ethnic, Racial, and Socioeconomic Integration in New Haven Public Schools. Technical report, Yale University, New Haven, CT.

De Chaisemartin, C. and L. Behaghel (2020). Estimating the Effect of Treatments Allocated by Randomized Waiting Lists. Econometrica 88(4), 1453–1477.

Deming, D. J. (2023). Why Do Wages Grow Faster for Educated Workers? Working paper, National Bureau of Economic Research.

District of Columbia Public Schools (2024). Aferschool programs.

Duncan, G. J. and R. J. Murnane (2014). Restoring opportunity: The crisis of inequality and the challenge for American education. Harvard Education Press.

Durkin, K., M. W. Lipsey, D. C. Farran, and S. E. Wiesen (2022). Effects of a Statewide Pre-Kindergarten Program on Children's Achievement and Behavior through Sixth Grade. Developmental Psychology 58(3), 470–484.

Elm City Montessori (2024). Frequently Asked Questions. Available at https://elmcitymontessori.org/about-us/frequently-asked-questions/. Accessed March 13th, 2024.

Fitzpatrick, M. D. (2010). Preschoolers Enrolled and Mothers at Work? The Effects of Universal Prekindergarten. Journal of Labor Economics 28(1), 51–85. Publisher: The University of Chicago Press.

Fitzpatrick, M. D. (2012). Revising Our Thinking About the Relationship Between Maternal Labor Supply and Preschool. Journal of Human Resources 47(3), 583–612.

Friedman-Krauss, A. H., W. S. Barnett, and J. K. Duer (2022). The State(s) of Head Start and Early Head Start: Looking at Equity. Technical report, National Institute for Early Education Research, New Brunswick, NJ.

Friedman-Krauss, A. H., W. S. Barnett, K. S. Hodges, K. A. Garver, G. Weisenfeld, B. A. Gardiner, and T. M. Jost (2023). The State of Preschool 2022: State Preschool Yearbook. Technical report, National Institute for Early Education Research, New Brunswick, NJ.

García, J. L., J. J. Heckman, D. E. Leaf, and M. J. Prados (2020). Quantifying the Life-Cycle Benefits of an Influential Early-Childhood Program. Journal of Political Economy 128(7), 2502–2541.

Gelbach, J. B. (2002). Public Schooling for Young Children and Maternal Labor Supply. American Economic Review 92(1), 307–322.

Gibbs, C. R., J. Wikle, and R. Wilson (2024, July 10). A Matter of Time? Measuring Effects of Public Schooling Expansions on Families' Constraints. EdWorkingPaper 24-987, Brown University.

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.

Havnes, T. and M. Mogstad (2011). Money for Nothing? Universal Child Care and Maternal Employment. Journal of Public Economics 95(11-12), 1455–1465.

Hays, E. (2021). Middle School Reopens Gradually. New Haven Independent. Available at https://www.newhavenindependent.org/index.php/article/middle_school_reopens_gradually/.

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), 2052–2086.

Hendren, N. and B. Sprung-Keyser (2020). A Unified Welfare Analysis of Government Policies. The Quarterly Journal of Economics 135(3), 1209–1318.

Jacobson, L. S., R. J. LaLonde, and D. G. Sullivan (1993). Earnings losses of displaced workers. The American Economic Review 83(4), 685–709.

Kleven, H., C. Landais, and J. E. Søgaard (2019). Children and Gender Inequality: Evidence from Denmark. American Economic Journal: Applied Economics 11(4), 181–209.

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.

Kuhn, P. and F. Lozano (2008). The Expanding Workweek? Understanding Trends in Long Work Hours among US Men, 1979–2006. Journal of Labor Economics 26(2), 311–343.

Lefebvre, P. and P. Merrigan (2008). Child-Care Policy and the Labor Supply of Mothers with Young Children: A Natural Experiment from Canada. Journal of Labor Economics 26(3), 519–548.

Lefebvre, P., P. Merrigan, and M. Verstraete (2009). Dynamic Labour Supply Effects of Childcare Subsidies: Evidence from a Canadian Natural Experiment on Low-Fee Universal Child Care. Labour Economics 16(5), 490–502.

Lin, W. and J. M. Wooldridge (2019). Chapter 2 - Testing and Correcting for Endogeneity in Nonlinear Unobserved Effects Models. In M. Tsionas (Ed.), Panel Data Econometrics, pp. 21–43. Academic Press.

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.

Malik, R. (2018). The effects of universal preschool in washington, dc: ChildrenâĂŹs learning and mothersâĂŹ earnings. Technical report, Washington, DC: Center for American Progress. Accessed June.

New York City Public Schools (2023). Pre-k enrollment.

New York City Public Schools (2024). Aferschool programs.

Peak, C. (2019). Profs Tapped to Fix School-Choice Lottery. New Haven Independent. Available at https://www.newhavenindependent.org/index.php/archives/entry/professors_school_choice_lottery/.

Povich, E. (2024). When 'Universal' Pre-K Really Isn't: Barriers to Participating Abound. Stateline. Available at https://stateline.org/2024/07/03/when-universal-pre-k-really-isnt-barriers-to-participating-abound/.

Rosenbaum, P. R. and D. B. Rubin (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. Biometrika 70(1), 41–55.

Sabol, T. J. and P. L. Chase-Lansdale (2015). The Influence of Low-Income Children's Participation in Head Start on Their Parents' Education and Employment. Journal of Policy Analysis and Management 34(1), 136–161.

Schiman, C. (2022). Experimental Evidence of the Effect of Head Start on Mothers' Labor Supply and Human Capital Investments. Review of Economics of the Household 20(1), 199–241.

Shapiro, E. (2023). In an Expensive City, Who Should Get Free Preschool? New York Times. Available at https://www.nytimes.com/2023/10/11/nyregion/nyc-3k-preschool-costs.html. Accessed February 29th, 2024.

State of Connecticut (1999). An Act Concerning School Choice and Interdistrict Programs. State of Connecticut Substitute House Bill No. 6950.

Swenson, K. and K. Burgess Simms (2021). Increases in Out-of-Pocket Child Care Costs: 1995 to 2016. Technical report, Office of The Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. Available at https://aspe.hhs.gov/sites/default/files/migrated_legacy_files//200606/increases-in-out-of-pocket-child-care-costs.pdf. Accessed: 02/28/2024.

US Department of the Treasury (2021). The Economics of Child Care Supply in the United States. Technical report, US Department of the Treasury.

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.

White House (2021). Fact Sheet: The American Families Plan. Available at https://bidenwhitehouse.archives.gov/wp-content/uploads/2021/04/American-Families-Plan-Fact-Sheet-FINAL.pdf. Accessed: 04/14/2025.

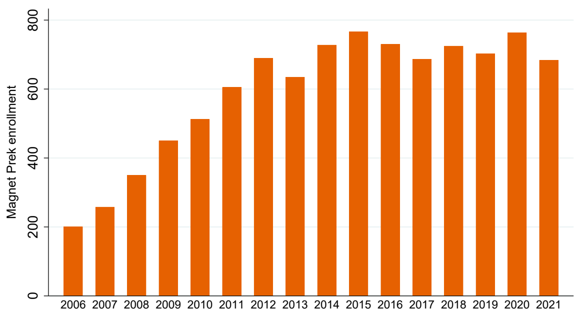
Wikle, J. and R. Wilson (2023). Access to Head Start and Maternal Labor Supply: Experimental and Quasi-Experimental Evidence. Journal of Labor Economics 41(4), 1081–1127.

Woodyard, H., T. R. Sass, and I. Fazlul (2023). Assessing the Benefits of Education in Early Childhood: Evidence from a Pre-K Lottery in Georgia. Technical Report No. 23-880, Annenberg Institute for School Reform at Brown University.

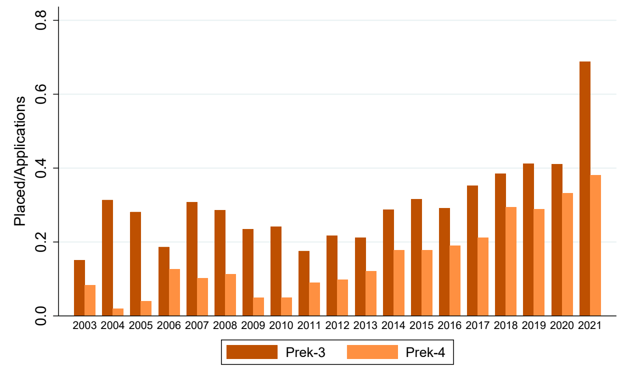
Zimmerman, S. D. (2014). The Returns to College Admission for Academically Marginal Students. Journal of Labor Economics 32(4), 711–754.

Figures

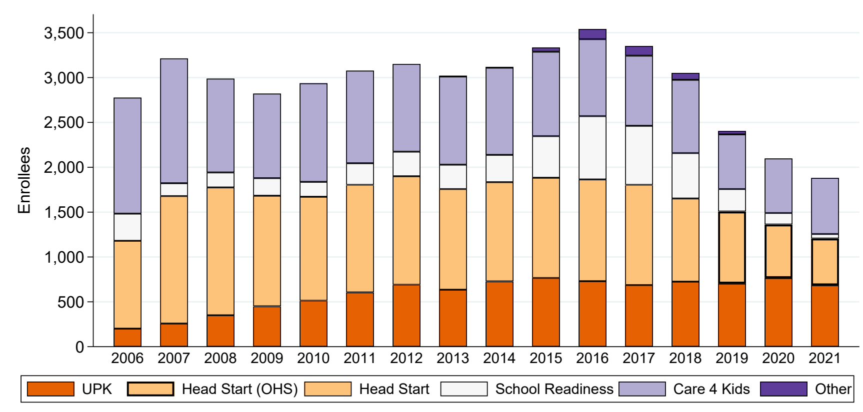
Figure 1: New Haven pre-kindergarten enrollment and applications



(a) UPK enrollment



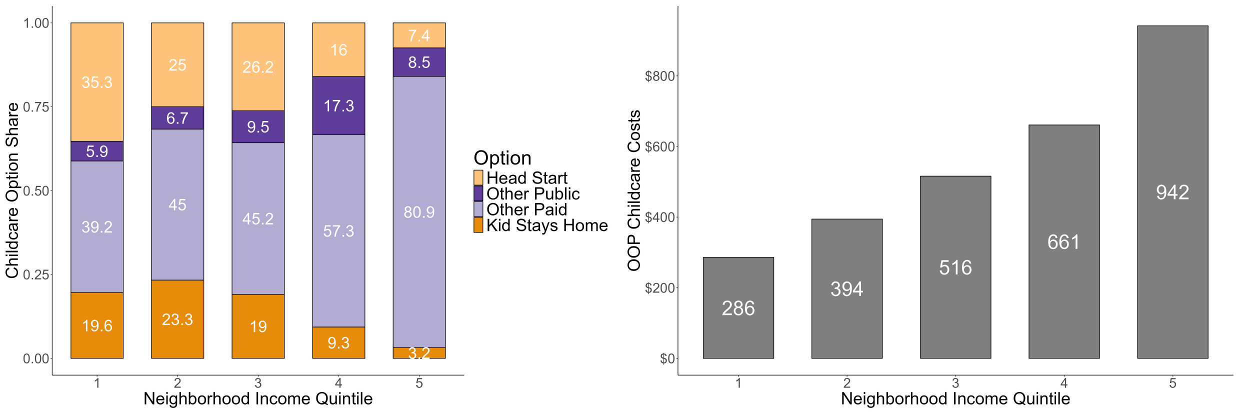
(b) Oversubscription by grade



(c) Enrollment in programs with subsidies

Notes: Panel (a): total enrollment in NHPS UPK programs by year. Panel (b): the ratio of placement counts to unique applicants by application grade and lottery year. Panel (c): enrollment in subsidized New Haven pre-kindergarten programs among three- and four-year olds by application year. Head Start enrollment counts are imputed from 2019 to 2021 using aggregate Office of Head Start (OHS) data. Care 4 Kids, School Readiness, and Other enrollment bars impute unique enrollment counts from data on enrollment spells, using the ratio of spells to unique students among applicants to the UPK program. Source: authors' calculations from NHPS data, CT Department of Education data, CT Office of Early Childhood data, and aggregate data from the Office of Head Start. See Section 2 for details.

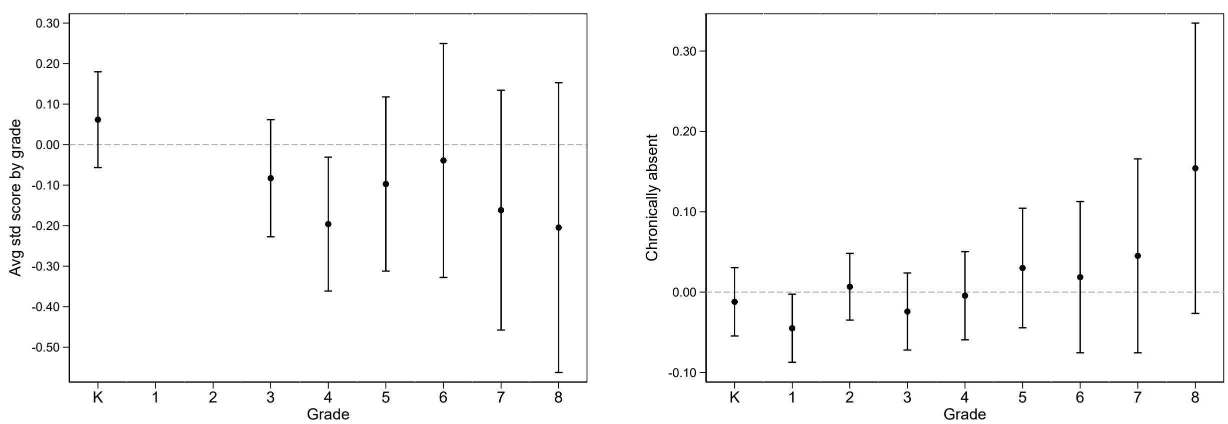
Figure 2: Outside options and out-of-pocket costs in survey data



(a) Childcare Options by Income **(b) OOP Costs by Income**

Notes: This figure shows the composition of outside options for childcare and the out-of-pocket (OOP) costs associated with them based on data from our survey of lottery applicants. Panel (a) shows the composition of non-UPK childcare options for children who applied but did not enroll in UPK. Each bar shows the composition for a specific neighborhood income quintile based on the median ACS household in the Census block group where the child lived at the time of the UPK application. Panels (b) show the average monthly OOP costs by ACS neighborhood median household income quintiles. See Section 3 for details.

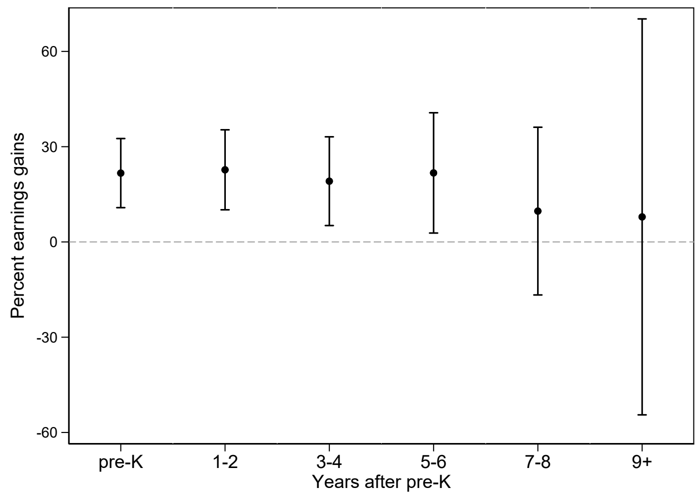
Figure 3: The effects of UPK on children's academic outcomes



(a) Test scores **(b) Chronic absenteeism**

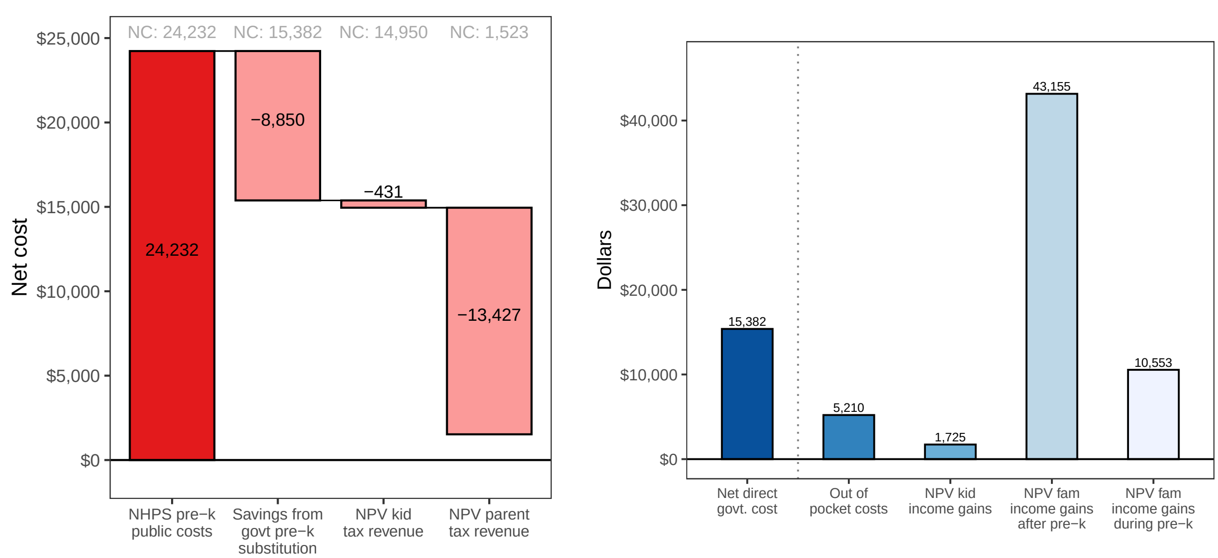
Notes: This figure shows IV estimates of the effect of UPK enrollment on average standardized test scores and chronic absenteeism. Test scores in Panel (a) are standardized to have a mean of 0 and variance of one by grade and cohort. Chronic absenteeism in Panel (b) is measured using Connecticut's definition: missing 10 percent or more of the total number of days enrolled during the school year. Each point estimate uses all available observations in a given grade for students subject to random assignment. Black dots correspond to point estimates with the surrounding error bars indicating the 90% confidence interval. Joint tests on the grades 3-8 coefficients in panels (a) and (b) fail to reject the null that the coefficients are all zero ($p = 0.140$ and $p = 0.631$, respectively).

Figure 4: The effects of UPK on parent earnings



Notes: This figure reports Poisson IV estimates where the outcome is parent income in dollars and the endogenous regressor is UPK enrollment. We report $exp(\hat{\beta}) - 1$, which is an estimate of the proportional change $(E[Y(1) - Y(0)]/E[Y(0)])$. Each point is an effect estimate (and associated 90% CI) for the time interval reported on the horizontal axis. All estimates include control for demographics and assignment propensities; see Section 4.1 for details. The sample consists of all lottery records for which parent earnings information is available over the specified time horizon. See Section 4.5 for discussion and Online Appendix H for details on the Poisson specification.

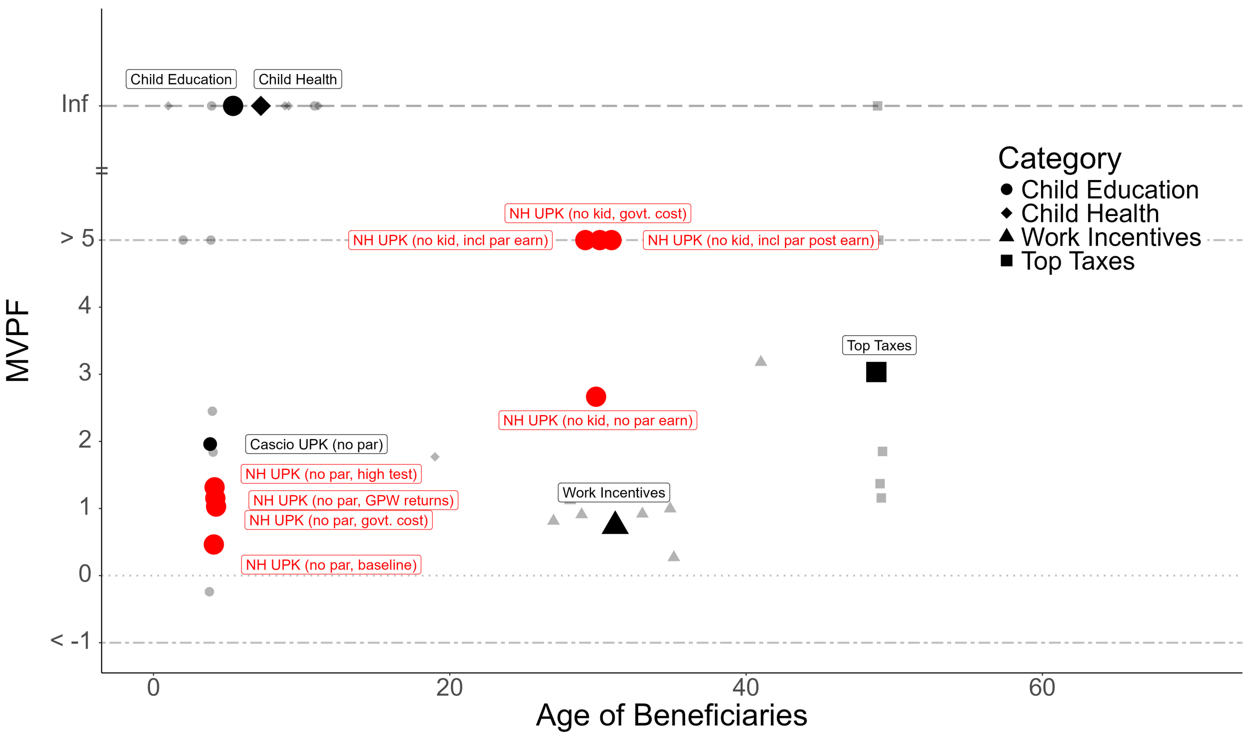
Figure 5: MVPF and inputs into willingness to pay and net costs



(a) Net costs **(b) Willingness to pay**

Notes: Panel (a): Net government costs of UPK provision are constructed by summing four different inputs. We add inputs sequentially, moving from left to right on the graph. (1) The direct public costs of providing the magnet pre-k slot, (2) the public savings from substitution away from other publicly funded pre-k and childcare programs, (3) changes in discounted tax revenue due to the estimated changes in kids' earnings, and (4) changes in discounted tax revenue due to changes in parents' earnings. Panel (b): we consider five possible inputs to willingness to pay (WTP) for UPK. (1) Net direct government cost, (2) reduction in out-of-pocket costs, (3) child income gains estimated using changes in kindergarten test scores, (4) parental earnings gains after pre-k, (5) parental earnings gains during pre-k. See Section 5 for details.

Figure 6: MVPF estimates across programs and program categories



Notes: This figure combines our own calculations of MVPFs for New Haven's UPK program with Cascio (2023)'s estimated MVPF for UPK and with estimates of other MVPF values for other program types from Hendren and Sprung-Keyser (2020). The horizontal axis is the age of program beneficiaries. The vertical axis is the MVPF value, pooling all values from $[5, \infty)$. The points labeled "NH UPK (no par, ...)" report MVPFs that exclude earnings gains for parents from the calculation, under different assumptions about what children's gains are. "Baseline" is the estimated score gains in NHPS data. "High test" assumes a pre-kindergarten score gain of 0.4 as in Lipsey et al. (2018), with earnings gains projected using Chetty et al. (2011). "GPW returns" assumes children's gains from the NHPS UPK program are as implied by the Gray-Lobe et al. (2023)'s estimates of four-year college attendance, projecting earnings gains using Zimmerman (2014). "NH UPK (no kid...)" report MVPFs that exclude earnings gains for children from the calculation. "incl par earn", "govt. cost", "no par earn", and "incl par post earn" represent the four different assumptions we make about inputs into willingness to pay, which we believe likely bracket the true value. See Section 5.2. Lighter gray markers are estimates for other individual policies or programs and the larger black markers are group-specific aggregates, as reported in Hendren and Sprung-Keyser (2020). The shape of the marker indicates the category of the policy or program as described in the legend. See Section 5 for details.

Tables

Table 1: Sample characteristics

	Lottery Sample	State Sample	Parent Earnings Sample	Survey Sample
Panel A: Lottery demographics				
Black	0.418	0.421	0.455	0.305
White	0.217	0.217	0.233	0.314
Hispanic	0.285	0.290	0.219	0.280
Female	0.506	0.505	0.507	0.519
Age at application	3.68	3.68	3.69	3.43
Pre-K 4 applications	0.479	0.485	0.495	0.258
New Haven applicant	0.617	0.617	0.586	0.576
ACS median HH income	59,713	59,735	61,600	65,960
Matched to state data	0.911	1.000	0.912	0.936
Matched to earnings	0.609	0.634	1.000	0.594
Survey respondent	0.043			1.000
Panel B: Lottery outcomes				
UPK assignment	0.251	0.246	0.244	0.419
UPK enrollment	0.265	0.282	0.276	0.546
Panel C: Test scores				
Avg std score K		0.123	0.160	
Avg std grade 3 test score		0.069	0.097	
Avg std grade 8 test score		0.038	0.041	
Panel D: Parent demographics				
Share with two parents	0.329	0.335	0.398	
Share of two parents post 2013	0.562	0.566	0.612	
Share of moms post 2013	0.690	0.691	0.676	
Any positive baseline income			0.785	
Baseline income (with zeros)			25,157	
Panel E: Survey data				
Employed full-time				0.696
Employed part-time				0.164
Hours worked per week				33.51
Respondent mom				0.897
N individuals	16037	13917	9162	840
N applications	18795	16485	10866	

Notes: This table shows the means of variables listed in the rows within samples defined by the columns. The "Lottery Sample" column describes the full set of UPK applicants who applied through the choice process. The "State Sample" column describes applicants matched to enrollment data. The "Parent Earnings Sample" column describes applicants whose parents we identify in the earnings data. The "Survey Sample" column describes applicants whose parents took the survey. Match rates reported in the lower part of Panel A are computed relative to attempted matches. Test scores in Panel C are standardized to have a mean zero and a standard deviation of one in the population of pre-kindergarten students in New Haven County. See Section 3 for details.

Table 2: Lottery design validation and first stage

	Comp Cont. Mean	Control Mean	NHPS sample	NHPS sample	State sample	Earnings sample	Survey sample
Panel A: Balance							
Black	0.348	0.430	-0.027	-0.006	-0.005	-0.011	-0.034
	(0.034)		(0.009)	(0.013)	(0.014)	(0.017)	(0.053)
White	0.246	0.199	0.034	0.009	0.008	0.014	0.060
	(0.028)		(0.007)	(0.011)	(0.012)	(0.016)	(0.050)
Female	0.566	0.510	-0.033	-0.028	-0.033	-0.019	-0.073
	(0.034)		(0.009)	(0.014)	(0.015)	(0.019)	(0.058)
Age at application	3.488	3.727	-0.010	0.002	0.008	-0.004	0.039
	(0.034)		(0.005)	(0.008)	(0.009)	(0.011)	(0.032)
ACS median HH income	63,589	58,650	1,518	-496	-267	-1,353	-3,495
	(2,176)		(575)	(777)	(828)	(1,066)	(3,982)
Pre-period income (dollars)	27,409	23,766				232	
	(2,280)					(949)	
Any pre-period income	0.836	0.776				-0.002	
	(0.034)					(0.013)	
Earnings-weighted index			824	-39	26	-207	-482
			(197)	(260)	(278)	(357)	(1,317)
Joint test			0.000	0.526	0.387	0.811	0.367
Panel B: Match							
Matched to state	0.895	0.862	0.026	0.011			
	(0.027)		(0.005)	(0.009)			
Matched to earnings	0.757	0.616	-0.005	-0.006			
	(0.033)		(0.008)	(0.013)			
Matched to survey	0.053	0.028	0.007	0.012			
	(0.013)		(0.004)	(0.006)			
Panel C: First Stage							
Enrolled NHPS UPK	0.000	0.000	0.581	0.390	0.402	0.416	0.409
			(0.008)	(0.013)	(0.014)	(0.018)	(0.053)
Years NHPS UPK	-0.043	0.076	0.833	0.576	0.634	0.625	
	(0.030)		(0.013)	(0.022)	(0.023)	(0.029)	
Year and Grade FEs			✓	✓	✓	✓	✓
Admit prob. indicators				✓	✓	✓	✓
First stage partial F-stat			5,496.1	844.1	838.2	568.7	63.0
N individuals			16037	15931	13847	9078	829
N applications			18795	18669	16389	10753	

Notes: Panels A and B of this table report results from reduced-form versions of Equation 1, taking either predetermined student and parent covariates (Panel A) or indicators for match to the listed data source (Panel B) as the dependent variable of interest. The joint test in Panel A evaluates the hypothesis that all coefficients shown in a given column (except for the coefficient on the earnings-weighted index) as well as the coefficients for an additional set of ACS tract-level controls, shown in Appendix Table A.3, are zero. Panel C reports first-stage estimates of Equation 1 where the outcome is either following-year enrollment in an NHPS UPK program or years of enrollment in an NHPS UPK program. Columns 1 and 2 report the control complier and control group means of the dependent variable listed in the row. Columns 3-7 report regression results from a specification where the dependent variable is as listed in the table row and the controls and samples vary across columns. Each cell reports results from a separate regression. The reported estimates are coefficients on an indicator for being offered a UPK spot, with standard errors in parentheses. Column 3 uses all available application data and includes only grade-by-year fixed effects. Column 4 uses all available application data and adds controls for the P_i, as described in Section 4.1. Column 5 has the same controls as column 4, but restricts to application data that is successfully matched to state records. Column 6 has the same controls as column 4 but restricts to application data linked to parent earnings records. Column 7 restricts to the survey sample and uses the recentered instrument. Standard errors are clustered at the application level (columns 3-5 and 7), or two ways at the application and parent level (column 6). See Section 4.2 for details.

Table 3: UPK effects on childcare outcomes

	Comp Cont. Mean	Control Mean	State sample	Survey sample
Panel A: Substitution				
Aggregate Measures				
Any pre-k or childcare (admin)	0.624	0.540	0.393	
	(0.034)		(0.025)	
Any pre-k or childcare (survey)	1.096	0.877		0.022
	(0.122)			(0.044)
Specific Program Types				
Enrolled Head Start (admin)	0.203	0.202	-0.139	
	(0.029)		(0.019)	
Enrolled School Readiness (admin)	0.066	0.086	-0.076	
	(0.021)		(0.013)	
Care4Kids (admin)	0.186	0.197	-0.050	
	(0.028)		(0.026)	
Any other SDE/OEC pre-k (admin)	0.456	0.351	-0.170	
	(0.036)		(0.032)	
Enrolled Head Start (survey)	0.162	0.217		-0.192
	(0.151)			(0.068)
Other paid option (survey)	0.720	0.539		-0.623
	(0.183)			(0.082)
Another public option (survey)	0.144	0.094		-0.117
	(0.111)			(0.051)
Panel B: Usage intensity				
Weekly childcare hours (survey)	51.235	28.857		11.319
	(13.022)			(3.200)
Monthly OOP costs (survey)	487	617		-375
	(216)			(90)
First stage partial F-stat			839.2	63.0
N individuals			13842	829
N applications			16331	

Notes: This table describes how compliers substitute away from other programs when they enroll in UPK. Results are from IV estimates of Equation 1 where the outcome is as listed in the rows. All specifications include controls for assignment propensity and demographics as described in Section 4.1. Specifications using survey outcomes use re-centered instruments. The first two columns report complier control means and control group means. The third and fourth columns report the IV estimates for our different datasets. "State sample" is the set of applications matched to state records. Results restricting the state sample to 2015-2017-the years with the most complete data-can be found in Appendix Table K.1. "Survey sample" is the set of survey respondents. Panel A reports results from specifications where outcomes are enrollment indicators for other pre-kindergarten programs. The top two rows indicate whether a child enrolls in any pre-k program (including UPK) based on administrative or survey data, respectively. The remaining rows in Panel A show enrollment in each specific pre-K program, with data sources in parentheses. Panel B reports IV regressions where outcomes are measures of childcare usage based on survey data. "Weekly hours" reports the change in weekly hours of childcare or pre-k provided. "Monthly OOP costs" is the self-reported monthly cost of childcare. Standard errors are in parentheses and are clustered at the applicant level. See Section 4.3 for details.

Table 4: Labor market effects

	Income Cntrl Mn	Work Indicator	Income (Incl. 0s)	Poisson IV	N individuals
Panel A: Earnings from admin data					
Pre-K years	34,318	0.057	5,464	0.217	10619
	(2,486)	(0.026)	(1,718)	(0.066)	
Years after pre-K 1-2	36,542	0.037	6,485	0.227	10282
	(2,787)	(0.028)	(2,217)	(0.077)	
Years after pre-K 3-4	34,961	0.050	6,417	0.191	9989
	(3,300)	(0.031)	(2,445)	(0.085)	
Years after pre-K 5-6	36,159	0.062	6,857	0.217	8221
	(3,619)	(0.040)	(3,200)	(0.115)	
Pooled post pre-k					
Years after pre-K 1-6	35,920	0.048	6,474	0.209	10282
	(2,825)	(0.028)	(2,260)	(0.079)	
Years after pre-K 7+	33,470	0.054	3,560	0.094	6128
	(6,744)	(0.077)	(5,851)	(0.213)	
Panel B: Earnings from admin data (Balanced)					
Pre-K years	36,528	0.044	4,380	0.153	9139
	(2,575)	(0.027)	(1,839)	(0.068)	
Years after pre-K 1-2	38,781	0.035	5,234	0.164	9139
	(2,761)	(0.029)	(2,280)	(0.077)	
Years after pre-K 3-4	36,509	0.042	6,253	0.181	9139
	(2,953)	(0.032)	(2,562)	(0.088)	
Pooled post pre-k					
Years after pre-K 1-4	37,645	0.039	5,743	0.172	9139
	(2,768)	(0.028)	(2,301)	(0.079)	
	Weekly Hrs Cntrl Mn	Employed FT	Employed PT	Hours/ week	
Panel C: Survey data					
During pre-k	27.87	0.221	-0.133	12.80	726; 726; 721
	(5.89)	(0.138)	(0.119)	(4.25)	
After pre-k	37.44	0.00	0.08	1.48	497; 497; 487
	(5.46)	(0.11)	(0.09)	(3.91)	

Notes: This table reports IV estimates of Equation 1 where outcomes are measures of parent earnings and labor supply. All specifications include controls for assignment propensity and demographics, as described in Section 4.1. For alternate specifications of Panel A, see Appendix Table A.5. Rows are samples defined by time relative to pre-kindergarten enrollment. Panels A and B report results obtained using administrative earnings records. Column 1 is the control complier mean of base-period income, in dollars. The remaining columns report regression results for different outcome variables or specifications. "Work indicator" takes a dummy for annual earnings (computed as the sum of earnings over 4 quarters) being greater than 0 as the outcome. "Income (incl 0s)" takes dollar income as the outcome, including zero income data points. "Poisson IV" takes dollar income as the outcome and estimates a Poisson regression using a control function approach (see Online Appendix H). For the Poisson regression, we report $exp(\hat{\beta})-1$, which is an estimate of the proportional change $(E[Y(1)-Y(0)]/E[Y(0)])$. Panel A uses all available data in each specification. Panel B restricts the sample to a balanced panel of individuals in cohorts that we can follow through four years after pre-kindergarten. Panel C uses survey records to estimate the labor supply effects of UPK enrollment. In Panel C, column 1 reports the control complier mean hours worked. Columns 2-4 report IV estimates for different labor supply outcomes from the survey. "Employed FT" is an indicator for if the respondent reported working full time. "Employed PT" is an indicator for if the respondent reported working part time. "Hours/week" is the hours per week the respondent reported working. Sample sizes for each specification are included in the last column. In Panels A and B, standard errors are two-way clustered at the applicant and parent level, except for the Poisson specification, where standard errors are estimated via bootstrap clustered at the at the application level with 500 bootstrap draws. In Panel C standard errors are clustered at the respondent level. See Section 4.5 for details.

Table 5: Career disruption

	Switch main industry	One job over \$ 4,000	Quarters earn. $\leq$ \$ 4,000 (Incl. 0s)	Total qts earn. $\leq$ \$ 4,000 since PK	N individuals
Disaggregated					
Pre-K years	-0.072	0.300	-0.206	-0.316	9205; 10621; 10727; 10727
	(0.027)	(0.108)	(0.098)	(0.143)	
Yrs after PK 1-2	-0.022	0.220	-0.152	-0.659	9010; 10285; 10391; 10391
	(0.024)	(0.117)	(0.096)	(0.267)	
Yrs after PK 3-4	-0.040	0.132	-0.094	-0.762	8385; 9990; 10096; 10096
	(0.024)	(0.117)	(0.094)	(0.392)	
Yrs after PK 5-6	-0.017	0.043	0.020	-0.319	6633; 8221; 8327; 8327
	(0.030)	(0.143)	(0.112)	(0.563)	
Pooled post pre-k					
Yrs after PK 1-6	-0.026	0.139	-0.087	-0.649	9476; 10285; 10391; 10391
	(0.017)	(0.104)	(0.079)	(0.360)	
Yrs after PK 7+	-0.030	0.016	0.065	-0.989	4964; 6128; 6234; 6234
	(0.043)	(0.269)	(0.168)	(1.122)	

Notes: This table reports IV estimates of Equation 1 where outcomes are measures of career disruption. All specifications include controls for assignment propensity and demographics, as described in Section 4.1. Rows are samples defined by time relative to pre-kindergarten enrollment. Columns are outcomes. "Switch main industry" is a binary variable equal to one if the applicant's top-earning industry in the current academic year differs from the prior year, conditional on being employed in the prior year, with unemployment in the current year also counted as a switch. "One job over \$4,000" counts the number of quarters in an academic year that an individual has exactly one job and that job pays at least \$4,000. "Quarters earn $\leq$ \$4,000" counts quarters with earnings of less than \$4,000 in an academic year. "Total qts earn $\leq$ \$4,000 since PK" counts the total number of quarters, including those in the present year, in which an individual has earned less than \$4,000 since the start of the pre-k years. Complier means for are reported in Table A.6. Standard errors are two-way clustered at the application and parent levels.

Table 6: UPK effects by family income tercile

	ACS 1st tercile		ACS 2nd tercile		ACS 3rd tercile	
Panel A: First Stage	FS	F-stat	FS	F-stat	FS	F-stat
Magnet Enrollment	0.419	229.7	0.440	371.1	0.355	261.0
	(0.028)		(0.023)		(0.022)	
N individuals	5223		5038		5089	
N applications	6027		5776		5842	
Panel B: Substitution	CCM	IV	CCM	IV	CCM	IV
Weekly childcare hours (survey)	57.0	6.1	43.3	10.9	40.3	12.5
	(40.3)	(10.0)	(32.8)	(5.9)	(12.0)	(5.3)
Any pre-k or childcare (survey)	1.356	-0.031	0.944	0.063	1.078	-0.025
	(0.722)	(0.180)	(0.295)	(0.103)	(0.098)	(0.034)
Enrolled Head Start (admin)	0.285	-0.273	0.113	-0.114	0.039	-0.037
	(0.048)	(0.048)	(0.042)	(0.026)	(0.028)	(0.025)
Enrolled Head Start (survey)	0.371	-0.372	0.187	-0.190	0.225	-0.161
	(0.560)	(0.157)	(0.375)	(0.160)	(0.159)	(0.079)
Another public option (survey)	0.137	-0.083	0.111	-0.110	0.130	-0.134
	(0.186)	(0.054)	(0.218)	(0.100)	(0.179)	(0.087)
Other paid option (survey)	0.791	-0.547	0.597	-0.589	0.632	-0.671
	(0.636)	(0.182)	(0.416)	(0.168)	(0.242)	(0.116)
Monthly OOP costs (survey)	-20	-61	596	-400	675	-565
	(787)	(127)	(468)	(188)	(308)	(149)
N individuals (survey)	174	174	250	250	274	274
N individuals (admin)	5221	5221	5033	5033	5087	5087
N applications (admin)	6024	6024	5771	5771	5840	5840
Panel C: Test scores	CCM	IV	CCM	IV	CCM	IV
Avg std score K	0.125	-0.176	0.170	0.293	0.222	0.013
	(0.128)	(0.183)	(0.153)	(0.108)	(0.090)	(0.114)
Avg std score grade 8	-0.230	-0.513	-0.063	0.538	0.416	-0.538
	(0.382)	(0.820)	(0.319)	(0.269)	(0.191)	(0.374)
N individuals	2996	2996	2910	2910	2968	2968
N applications	3548	3548	3398	3398	3478	3478
Panel D: Parent Earnings	CCM	IV	CCM	IV	CCM	IV
Pre-K years	18,159	0.106	34,324	0.247	43,153	0.268
	(2,201)	(0.135)	(3,566)	(0.120)	(5,191)	(0.108)
Years after pre-K 1-2	20,050	0.034	36,826	0.277	42,344	0.293
	(2,353)	(0.155)	(4,326)	(0.159)	(5,596)	(0.138)
Years after pre-K 3-4	19,324	-0.063	36,776	0.351	41,032	0.176
	(2,468)	(0.140)	(4,529)	(0.150)	(6,192)	(0.138)
Years after pre-K 5-6	15,853	-0.015	36,629	0.336	43,150	0.193
	(5,114)	(0.282)	(5,077)	(0.232)	(7,540)	(0.157)
N applicants	3336	3336	3296	3296	3781	3781

Notes: This table reports IV estimates of substitution patterns, test score impacts, and parental wage income impacts by tercile of ACS median block-group household income. Terciles are computed based on the distribution of neighborhood median household income among lottery applicants, with the 2nd tercile starting at \$39,323 and the 3rd tercile starting at \$69,619. All specifications are IV estimates of Equation 1 except for parents' earnings which uses the Poisson specification and reports $exp(\hat{\beta}) - 1$ which is an estimate of the proportional change ($E[Y(1) - Y(0)]/E[Y(0)]$). See Online Appendix H for details. All regressions include controls for demographics and assignment propensity, as described in Section 4.1. Panel A reports first stage (FS) estimates of Equation 1 where the outcome is following year enrollment in an NHPS UPK Program. Panel B reports substitution patterns from other pre-k programs, weekly hours of pre-k, and monthly out-of-pocket costs. Implausible values of the complier control mean for "Enrolled Head Start (admin)", "Monthly OOP costs", and "Another public option" for the first tercile are due to small samples. Panel C reports impacts on students' average Kindergarten Entrance Inventory (KEI) scores. See Online Appendix E for details. Panel D reports estimates on parental wage income gains. The first two columns report the control-group complier mean (CCM) and the Poisson estimate for the lowest ACS median household income tercile, while the remaining columns report the same estimates for the 2nd and 3rd terciles. Standard errors are clustered at the child level in Panels A, B, and C, and at the application level in Panel D. See Section 4.6 for details.

Table 7: UPK labor market effects by demographic group

	Student Race			Family Structure (Apps Post 2013)					
	Black	Hispanic	White	All Post 2013	Moms	Dads	1-Parent	2-Parents, Individual Income	HH Income, All Parents Matched
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pre-K years									
Estimate	0.141	0.174	0.385	0.221	0.200	0.278	0.179	0.244	0.195
	(0.101)	(0.143)	(0.162)	(0.068)	(0.080)	(0.119)	(0.091)	(0.113)	(0.069)
N applicants	4770	2236	2434	6147	4989	2622	3944	1884	6115
Years after pre-K 1-2									
Estimate	0.044	0.129	0.512	0.233	0.186	0.298	0.150	0.245	0.212
	(0.110)	(0.171)	(0.205)	(0.085)	(0.101)	(0.143)	(0.106)	(0.133)	(0.081)
N applicants	4629	2147	2354	5808	4711	2452	3768	1738	5790
Years after pre-K 3-4									
Estimate	0.043	0.237	0.423	0.183	0.222	0.123	0.223	0.114	0.173
	(0.117)	(0.207)	(0.214)	(0.087)	(0.122)	(0.137)	(0.124)	(0.147)	(0.092)
N applicants	4524	2075	2308	5515	4484	2341	3593	1639	5507
Years after pre-K 5-6									
Estimate	-0.014	0.210	0.396	0.231	0.190	0.351	0.414	0.086	0.195
	(0.166)	(0.386)	(0.238)	(0.123)	(0.148)	(0.242)	(0.203)	(0.192)	(0.116)
N applicants	3779	1629	1992	3747	3015	1533	2472	1048	3746

Notes: This table replicates the Poisson specifications from Panel A of Table 4 for different subsets of applicants. Columns 1-3 restrict the sample to parents listed on applications from Black, Hispanic, and White students. Columns 4-9 restrict the sample to applications in 2014 or later, as starting in 2014 information on guardians was collected systematically on applications. Column 4 restricts the sample to all applications in 2014 or later. Columns 5-6 restrict the sample to only mothers or only fathers listed on UPK applications. Columns 7-8 restrict the sample to parents listed on applications that had one or two parents listed. Column 9 restricts the sample to families in which all listed guardians matched to earnings and reports estimated effects on household income. Labor market outcomes are reported for the period the child was in pre-k as well as 1-2 years, 3-4 years, and 5-6 years after pre-k. Estimates are derived from Poisson regression run on annual income (including 0s) as described in Online Appendix H. The table reports $exp(\hat{\beta}) - 1$, which is an estimate of the proportional change $(E[Y(1) - Y(0)]/E[Y(0)])$. Standard errors, in parentheses, are estimated via bootstrap using 500 bootstrap samples and are clustered at the applicant level. Complier control means are shown in Appendix Table A.7. The number of observations is shown below standard errors. All regressions include controls for applicant race, gender, age, and neighborhood characteristics, as described in Section 4.1.

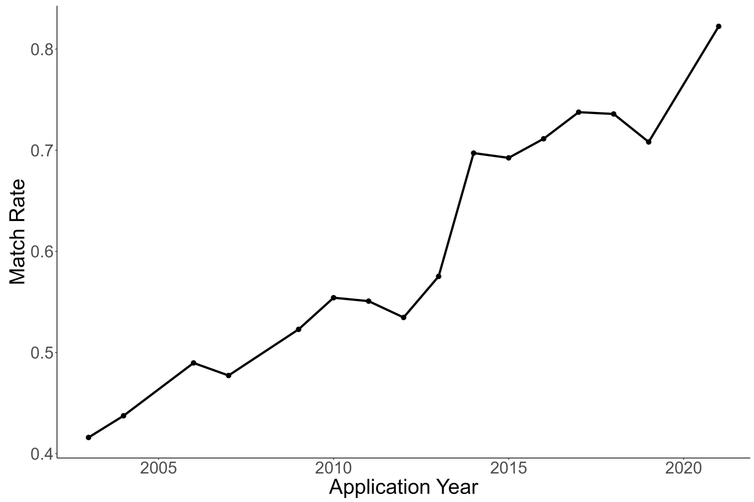
Table 8: MVPF estimates using varying willingness to pay constructions

Specification	WTP is net direct govt cost	WTP is OOP costs, kid earnings	WTP is OOP costs, kid earnings, parent post earnings	WTP is OOP costs, kid earnings, all parent earnings
Main	10.10	4.55	32.88	39.81
	[1.87, Inf]	[0.58, Inf]	[3.16, Inf]	[4.03, Inf]
Larger Effects on Kids				
Kids' effects of 0.4σ	Inf	Inf	Inf	Inf
	[2.60, Inf]	[2.68, Inf]	[6.18, Inf]	[7.26, Inf]
GPW kids' effects	Inf	Inf	Inf	Inf
	[2.41, Inf]	[2.27, Inf]	[5.50, Inf]	[6.49, Inf]
Decompositions				
No parent earnings gains	1.03	0.46	0.46	0.46
	[0.97, 1.09]	[0.18, 0.77]	[0.18, 0.77]	[0.18, 0.77]
No kids earnings gains	7.87	2.67	24.75	30.14
	[1.82, Inf]	[0.55, Inf]	[3.01, Inf]	[3.73, Inf]

Notes: This table reports estimates of the MVPF of UPK under different assumptions about earnings effects for children (upper panel) and inclusion/exclusion criteria for parent and child earnings gains (lower panel), using four different constructions of willingness to pay (one in each column). The first column assumes parents value the program at its net direct government cost. The next three columns report results from different hedonic approaches to WTP calculation. The second excludes parental earnings from WTP and considers only change in out of pocket expenditures and kids' future earnings. The third adds post-pre-k parental earnings, and the fourth considers all parental earnings. The first three rows show results under different assumptions about the effect of UPK on students. "Main" assumes the kindergarten score gains found in our analysis. "Kids' effects of 0.4σ" assumes test score gains of 0.40 standard deviations. "GPW kids' effects" assumes kids benefit through increased college enrollment as reported in Gray-Lobe et al. (2023). The bottom two rows report results when excluding parents' earnings gains ("No parent earnings gains") or kids' earnings gains ("No kids earnings gains") from the calculation.

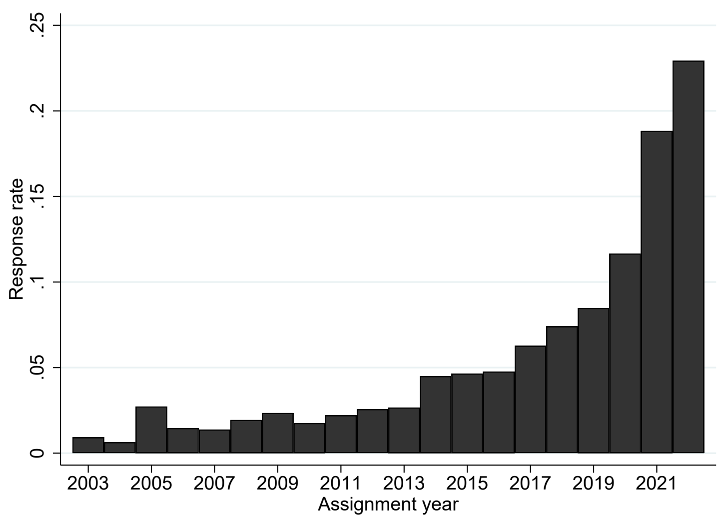
A Additional Tables and Figures

Figure A.1: Parental earnings match rate by application cohort



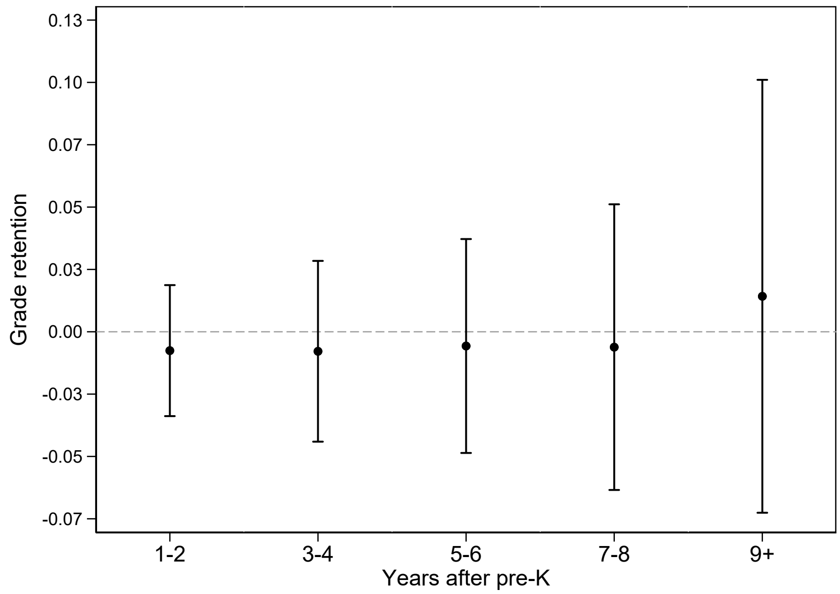
Notes: This figure shows the evolution of the match rate for parental earnings by application cohort. We can observe up to two contacts per applicant, usually one or both of their parents. We consider an applicant record matched if we can match at least one of their provided contacts to the administrative earnings records. We do not consider parental earnings for the 2005, 2008, and 2020 application cohorts. We did not recover historical records of the 2005 and 2008 application processes until after the merge with state records was conducted. As in our other analyses, we drop observations for the 2020 application cohort due to Covid. See Section 3 for details.

Figure A.2: Survey response rate by application cohort



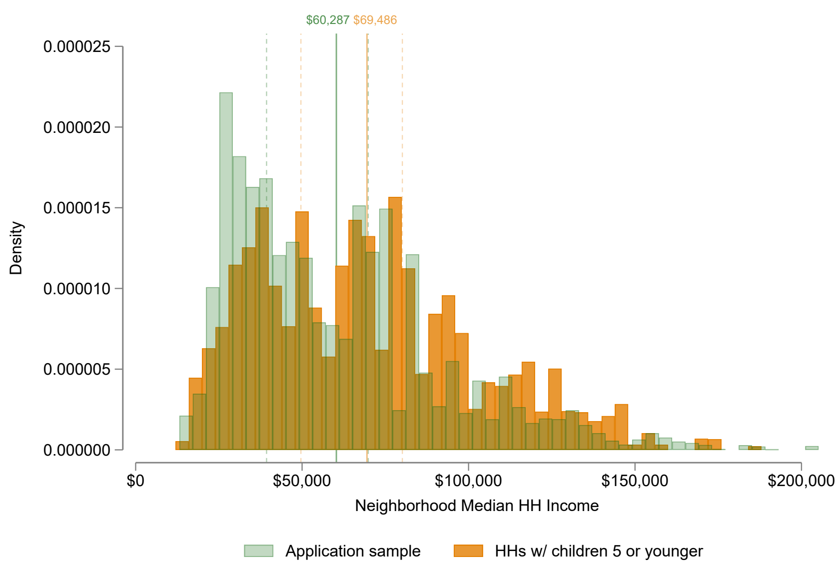
Notes: This figure reports survey response rates by application cohort. See Section 3 for details.

Figure A.3: The effects of UPK on grade retention



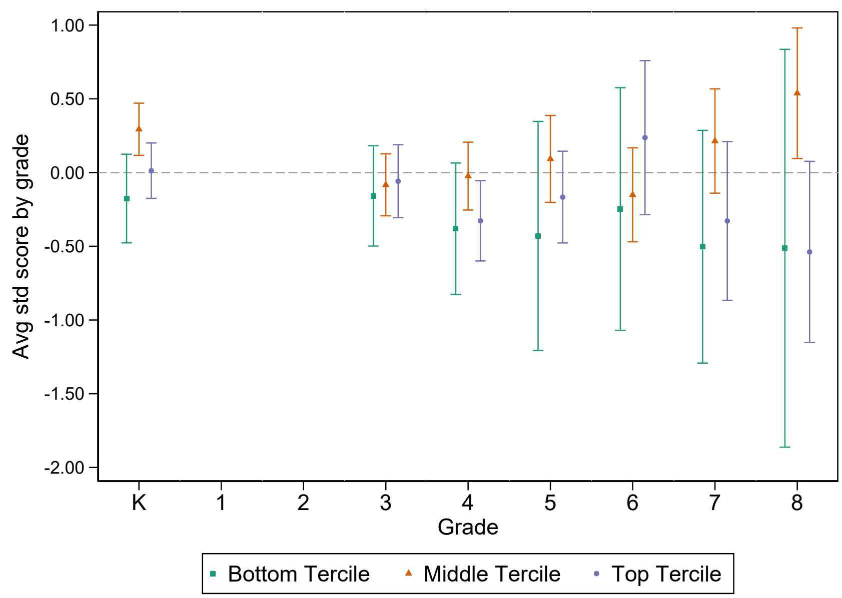
Notes: This figure shows IV estimates of the effect of UPK enrollment on grade retention. Grade retention is defined as a cumulative indicator for ever being retained up to that point. Each point estimate uses all available observations in a given range of years after pre-k for the students subject to random assignment. Black dots correspond to point estimates with the surrounding error bars indicating the 90% confidence interval.

Figure A.4: Neighborhood income distribution for applicants and households with children under age six

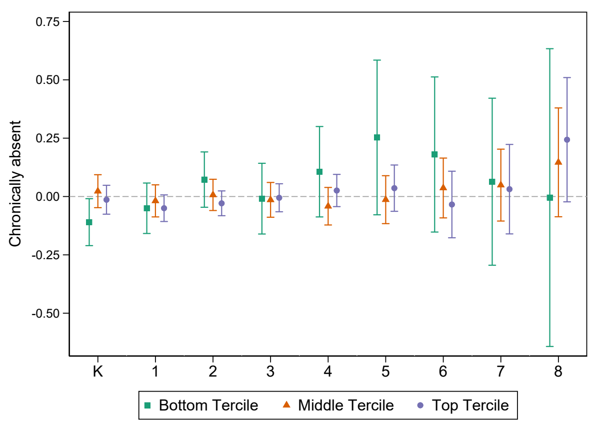


Notes: Distribution of neighborhood median household income in the sample of UPK applicants (light green bars) and all households with children under age six in New Haven County (orange bars). Source: block-group data from the 2019 ACS 5-year sample. Solid vertical lines represent each sample's mean, with the value shown on top of the lines. Dashed lines to the left and right represent each sample's 33.3rd and 66.7th percentiles, respectively. The 33.3rd and 66.7th percentiles in the applicant distribution are $39,323 and $69,868 respectively. For NH county, these values are $49,645 and $80,106. See Section 4.6 for details.

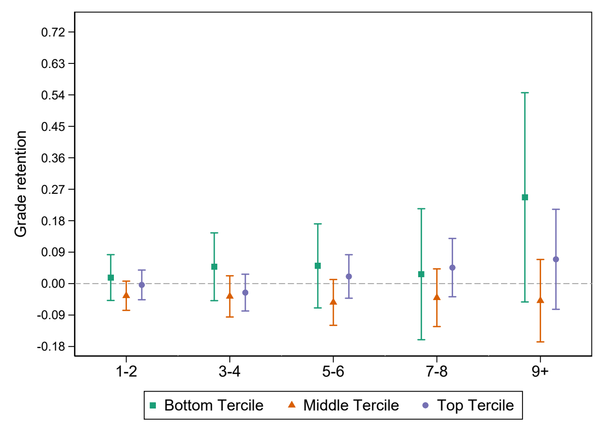
Figure A.5: Effects of UPK on children's academic outcomes by family income tercile



(a) Test scores



(b) Chronic absenteeism



(c) Grade retention

Notes: This figure shows IV estimates for the effect of UPK enrollment on standardized test scores (Panel (a)), chronic absenteeism (Panel (b)), and a cumulative indicator for ever being retained in a grade (Panel (c)) by tercile of ACS median block-group household income analogous to those presented in Figures 3 and A.3. Dots correspond to point estimates with the surrounding error bars indicating the 90% confidence interval.

Table A.1: Demographics in New Haven and Connecticut public schools

	NHPS	CT public schools
Asian	3.2%	5.2%
Black or African American	34.4%	12.5%
Hispanic	48.5%	30.0%
White	10.5%	47.5%
English Learners	20.5 %	9.7%
Free or reduced price meals	65.9%	42.4%
Students with disabilities	15.6%	17.1%

Notes: This table reports population characteristics (expressed as percentages) for New Haven Public Schools (column 1) and all public schools in Connecticut (column 2). Calculations from the State Department of Education for the 2022-23 school year (Connecticut State Department of Education, 2023).

Table A.2: Descriptive Statistics - Comparison Set

Variable	Comparison Set
Asian	0.04
Black	0.22
Hispanic	0.36
White	0.35
English learners	0.17
Free or reduced price meals	0.72
Students with disabilities	0.34
KEI Score	-0.03
N	54517

Notes: This table shows descriptive statistics for the comparison set of children enrolled in any public pre-kindergarten program in New Haven County during our sample period. See Section 3.2 for details.

Table A.3: Lottery design validation - All Variables

	Comp Cont. Mean	Control Mean	NHPS sample	NHPS sample	State sample	Earnings sample	Survey sample
Black	0.348	0.430	-0.027	-0.006	-0.005	-0.011	-0.034
	(0.034)		(0.009)	(0.013)	(0.014)	(0.017)	(0.053)
White	0.246	0.199	0.034	0.009	0.008	0.014	0.060
	(0.028)		(0.007)	(0.011)	(0.012)	(0.016)	(0.050)
Female	0.566	0.510	-0.033	-0.028	-0.033	-0.019	-0.073
	(0.034)		(0.009)	(0.014)	(0.015)	(0.019)	(0.058)
Age at application	3.488	3.727	-0.010	0.002	0.008	-0.004	0.039
	(0.034)		(0.005)	(0.008)	(0.009)	(0.011)	(0.032)
ACS median HH income	63,589	58,650	1,518	-496	-267	-1,353	-3,495
	(2,176)		(575)	(777)	(828)	(1,066)	(3,982)
Fraction renters	48.732	55.202	-2.394	0.071	-0.153	0.281	2.184
	(2.041)		(0.520)	(0.708)	(0.754)	(0.944)	(3.447)
Fraction HH below poverty	0.063	0.075	-0.002	-0.001	-0.001	0.001	0.004
	(0.004)		(0.001)	(0.001)	(0.002)	(0.002)	(0.006)
Fraction employed over 16	62.035	60.426	-0.250	0.097	0.227	-0.093	0.943
	(0.738)		(0.196)	(0.287)	(0.306)	(0.393)	(1.112)
Pre-period income (dollars)	27,409	23,766				232	
	(2,280)					(949)	
Any pre-period income	0.836	0.776				-0.002	
	(0.034)					(0.013)	
Earnings-weighted index			824	-39	26	-207	-482
			(197)	(260)	(278)	(357)	(1,317)
Joint test			0.000	0.526	0.387	0.811	0.367
Year and Grade FEs			✓	✓	✓	✓	✓
Admit prob. indicators				✓	✓	✓	✓
First stage partial F-stat			5,496.1	844.1	838.2	568.7	63.0
N individuals			16037	15931	13847	9078	829
N applications			18795	18669	16389	10753	

Notes: This table expands the results reported in Table 2 by reporting reduced-form results for all variables used in the joint balance test. The reduced-form regressions are again following Equation 1 and take predetermined student and parent covariates as the dependent variables of interest. The joint test considers the hypothesis that all coefficients in a given column (except for the coefficient on the earnings-weighted index) are zero. Columns 1 and 2 report the control complier and control group means of the dependent variable listed in the row. Columns 3-7 report regression results from a specification where the dependent variable is as listed in the table row and the controls and samples vary across columns. Each cell reports results from a separate regression. The reported estimates are coefficients on an indicator for being offered a UPK spot, with standard errors in parentheses. Column 3 uses all available application data and includes only grade-by-year fixed effects. Column 4 uses all available application data and adds controls for the P_i, as described in Section 4.1. Column 5 has the same controls as column 4, but restricts to application data that is successfully matched to state records. Column 6 has the same controls as column 4 but restricts to application data linked to parent earnings records. Column 7 restricts to the survey sample. Standard errors are clustered at the application level (columns 3-5 and 7), or two ways at the application and parent level (column 6). See Section 4.2 for details.

Table A.4: The effect of UPK enrollment on KEI subscores

	Comp Cont. Mean	Control Mean	Overall	ACS 1st tercile	ACS 2nd tercile	ACS 3rd tercile
Avg KEI Score	0.180	0.099	0.062	-0.176	0.293	0.013
	(0.067)		(0.072)	(0.183)	(0.108)	(0.114)
Language Skills	0.134	0.105	0.108	-0.040	0.345	-0.063
	(0.077)		(0.082)	(0.208)	(0.122)	(0.133)
Literacy Skills	0.258	0.124	0.040	-0.182	0.294	-0.010
	(0.079)		(0.084)	(0.207)	(0.124)	(0.138)
Numeracy Skills	0.240	0.072	-0.018	-0.411	0.205	0.061
	(0.081)		(0.086)	(0.221)	(0.127)	(0.136)
Physical/Motor Skills	0.168	0.080	0.011	-0.029	0.272	-0.045
	(0.077)		(0.085)	(0.218)	(0.128)	(0.134)
Creative/Aesthetic Skills	0.119	0.090	0.113	-0.205	0.281	0.096
	(0.077)		(0.084)	(0.210)	(0.125)	(0.134)
Personal/Social Skills	0.163	0.123	0.117	-0.191	0.365	0.038
	(0.076)		(0.081)	(0.203)	(0.120)	(0.133)
First-stage partial F-stat			578.1	105.3	314.5	203.4
N			8716	2996	2910	2968

Notes: This table reports IV estimates of Equation 1 for the effect of UPK enrollment on the Connecticut Kindergarten Entrance Inventory (KEI). The first row reports results on the overall KEI score and subsequent rows report results for each of the six subscores. Complier control mean and control mean columns report statistics for the full sample. The three rightmost columns report IV estimates within samples defined by tercile of neighborhood median household income, as measured in the 2019 ACS. All scores have been z-scored by year. Clustered standard errors are in parentheses. See Section 4 for details.

Table A.5: Alternate Specifications for Results in Table 4

	Income CCM	Work Indicator	Income (Incl. 0s)	Poisson IV	N individuals
Panel A: Recentered Instrument					
Pre-K years	34,760	0.060	4,519	0.153	10691; 10691; 10691
	(1,814)	(0.026)	(1,780)	(0.062)	
Years after pre-K 1-2	35,809	0.037	6,251	0.202	10346; 10346; 10346
	(2,216)	(0.028)	(2,185)	(0.074)	
Years after pre-K 3-4	34,821	0.049	5,755	0.170	10033; 10033; 10033
	(2,473)	(0.030)	(2,393)	(0.078)	
Years after pre-K 5-6	35,537	0.053	6,496	0.194	8250; 8250; 8250
	(2,996)	(0.039)	(3,165)	(0.103)	
Pooled post pre-k					
Years after pre-K 1-6	35,384	0.045	6,112	0.189	10346; 10346; 10346
	(2,204)	(0.027)	(2,224)	(0.074)	
Years after pre-K 7+	32,278	0.067	2,453	0.060	6152; 6152; 6152
	(5,069)	(0.071)	(5,243)	(0.176)	
Panel B: No Extra Controls					
Pre-K years	32,214	0.065	7,938	0.245	10731; 10731; 10731
	(2,759)	(0.029)	(2,968)	(0.101)	
Years after pre-K 1-2	34,670	0.038	8,502	0.255	10394; 10394; 10394
	(2,949)	(0.031)	(3,131)	(0.103)	
Years after pre-K 3-4	32,567	0.050	8,825	0.272	10101; 10101; 10101
	(3,222)	(0.032)	(3,241)	(0.111)	
Years after pre-K 5-6	31,033	0.063	11,695	0.382	8333; 8333; 8333
	(4,285)	(0.041)	(4,252)	(0.162)	
Pooled post pre-k					
Years after pre-K 1-6	32,993	0.048	9,409	0.291	10394; 10394; 10394
	(3,110)	(0.030)	(3,201)	(0.110)	
Years after pre-K 7+	28,643	0.042	6,949	0.239	6240; 6240; 6240
	(7,381)	(0.077)	(6,644)	(0.257)	

Notes: This table presents IV estimates of Equation 1, analogous to those in Table 4, under alternative specifications. Panel A applies the recentered instrument approach described in Section 4.1. Panel B omits all controls except for interacted propensity score groups. Standard errors are two-way clustered at the applicant and parent levels, except in the Poisson specification, where they are estimated via bootstrap clustering at the application level with 500 bootstrap draws.

Table A.6: Control complier means for specifications in Table 5

	Switch main industry	One job over $ 4,000	Quarters earn. $\leq$ $ 4,000 (Incl. 0s)	Total qts earn. $\leq$ $ 4,000 since PK	N individuals
Disaggregated					
Pre-K years	0.21	2.00	1.49	2.11	9205; 10621; 10727; 10727
	(0.02)	(0.12)	(0.10)	(0.14)	
Yrs after PK 1-2	0.17	2.08	1.46	4.71	9010; 10285; 10391; 10391
	(0.02)	(0.11)	(0.10)	(0.29)	
Yrs after PK 3-4	0.19	2.01	1.31	7.27	8385; 9990; 10096; 10096
	(0.02)	(0.11)	(0.10)	(0.46)	
Yrs after PK 5-6	0.15	2.05	0.99	9.14	6633; 8221; 8327; 8327
	(0.03)	(0.14)	(0.13)	(0.79)	
Pooled post pre-k					
Yrs after PK 1-6	0.17	2.05	1.29	6.73	9476; 10285; 10391; 10391
	(0.02)	(0.10)	(0.09)	(0.45)	
Yrs after PK 7+	0.18	1.97	1.01	13.57	4964; 6128; 6234; 6234
	(0.04)	(0.27)	(0.24)	(2.52)	

Notes: This table reports the complier control means for estimates in Table 5. See the note for that table and Section 4.5.3 for details.

Table A.7: Control complier means for specifications in Table 7

	Student Race			Family Structure (Apps Post 2013)					
	Black	Hispanic	White	All Post 2013	Moms	Dads	1-Parent	2-Parents, Individual Inc.	HH Income, All Parents Matched
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pre-K years									
CCM	25,886	28,018	43,025	31,344	28,417	39,710	27,301	36,783	42,336
	(3,770)	(3,220)	(7,507)	(2,465)	(2,671)	(4,530)	(1,784)	(7,042)	(3,425)
Years after pre-K 1-2									
CCM	28,177	31,768	44,019	33,871	30,735	44,139	29,307	41,670	45,687
	(4,380)	(4,377)	(6,374)	(2,826)	(3,035)	(5,102)	(2,727)	(6,537)	(4,144)
Years after pre-K 3-4									
CCM	25,344	28,101	38,790	31,945	28,563	41,582	27,390	39,395	43,597
	(4,533)	(4,329)	(7,884)	(3,156)	(3,137)	(5,108)	(2,759)	(5,961)	(4,487)
Years after pre-K 5-6									
CCM	25,255	29,874	36,410	27,334	27,498	31,647	22,351	37,348	37,858
	(5,989)	(8,464)	(8,967)	(4,191)	(3,792)	(7,428)	(5,158)	(7,935)	(5,805)

Notes: This table reports the complier control means for estimates in Table 7. See the note for that table and Section 4.6 for details.

B Assignment mechanisms

NHPS used a centralized process to assign students to UPK programs over our entire study period, from 2003 to 2022. Though run simultaneously, the assignment processes for students from New Haven and those from other towns are fully separate, with separate capacities and no competition between students. The city's assignment mechanism and other elements of the application process changed several times over the study period. This appendix describes changes in the NHPS UPK assignment system over time and how we use the data available under different assignment regimes to construct our school assignment instrument.

B.1 Assignment mechanisms and procedures

2003-2013

NHPS assigned students using an Immediate Acceptance (IA) mechanism with sibling and neighborhood priority, with ties broken using random draws. Students could apply to at most three schools. The mechanism was implemented by an IT consultant hired by the district, who wrote an NHPS-specific software package. Applications were primarily on paper, typically with one contact listed.

2014-2015

NHPS assigned students to schools using the "New Haven" mechanism, in which school preferences are lexicographic over 1) priority group (neighborhood, sibling, or both), and then 2) listed rank. Students apply to three schools. The district hired Smartchoice, a provider of school choice software to many districts, to implement the mechanisms. Application is primarily online, and applicants have the opportunity to list multiple contacts and define their relationship to the student.

2016-2017

NHPS returned to IA, and raised the maximum number of listed schools to four. The broader system continued to include neighborhood and sibling priority.

2018

NHPS continued as in 2016-2017, but added a priority system based on student zip code that comes after neighborhood and sibling preference in the lexicographic ordering.

2019

NHPS adopted a deferred acceptance (DA) procedure, keeping other process elements as in the previous year.

2020-2022

NHPS increased the length of the rank list from four to six. Other elements were unchanged.

B.2 Data on assignment processes and instrumental variable construction

The data we have about the assignment process changes when the district switches from the "old" procedure, a bespoke process run by a district-hired IT consultant, to the "new" procedure, run by a large purveyor of school choice services. This shift takes place after the 2013 assignment process.

2003-2013

Between 2003 and 2013, we observe applications, capacities, priorities, and realizations of the random lottery draws. We do not directly observe initial placement outcomes. Instead, we observe something closer to realized placements, inclusive of declined placements and the aftermarket processes through which NHPS filled declined spots.

To construct our UPK placement instrument, we simulate the school assignment process based on the applications, capacities, priorities, and realized values of the random draws. Using this simulation, we create an indicator equal to one if the student is assigned an offer in the main process. The feature of the data that lets us do this is that we observe the values of the randomized draws used to make assignments.

2014-2022

In 2014 and later, we observe applications, capacities, priorities, and initial placement outcomes. We use the presence of a main-round offer as our UPK placement instrument.

B.2.1 Simulated probabilities

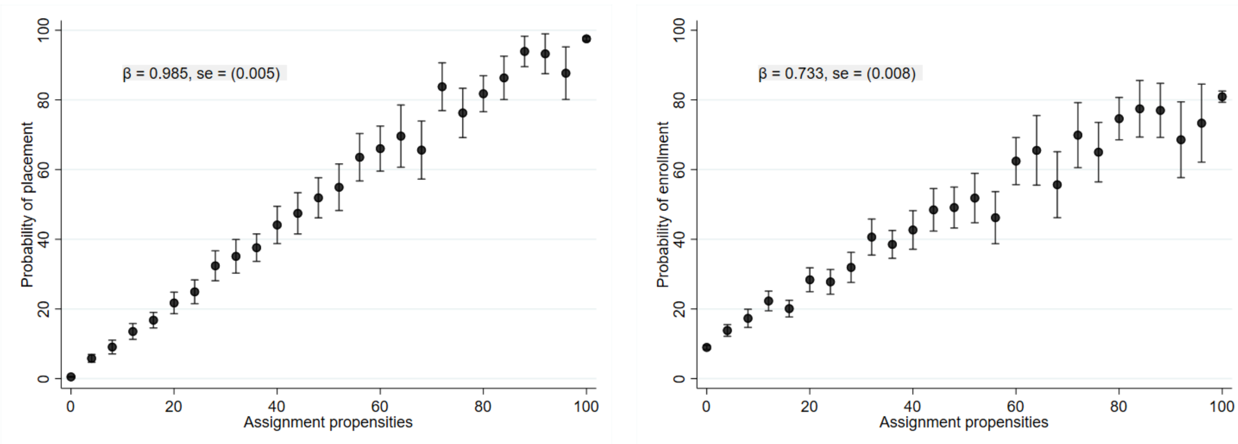
Our approach follows Azevedo and Leshno (2016) and Agarwal and Somaini (2018). We simulate random tiebreaker draws 500 times and run the assignment mechanism under each set of draws. We compute the admissions cutoff score in each simulation (based on the RSP+C representation of the mechanism), then take the average cutoff score across all simulations and use it to compute the assignment probabilities for each application.

B.2.2 Tests of simulation validity

Our process for constructing the placement instruments Z_i and assignment propensities P_i relies on reconstructing the assignment process. To test the accuracy of our reconstruction, we regress an indicator for observed assignment on simulated assignment probability P_i. Panel (a) of Figure B.1 reports the results of this exercise. We see that, as predicted, observed assignment probabilities rise one-to-one with the simulated values.

We also see observed *enrollment* probabilities rising steeply with the P_i. We report these results in Panel (b) of Figure B.1. The slope here is not one-for-one, and we do not expect it to be due to noncompliance. However, the fact that enrollment rises steeply with simulated assignment probability helps confirm that our approach accurately describes the assignment and enrollment procedures that students encounter.

Figure B.1: Validating assignment propensities



(a) Placement **(b) Enrollment**

Notes: Panel (a) plots the share of observed UPK assignments (vertical axis) against simulated assignment probabilities P_i (horizontal axis). Each dot represents a 5pp bin of values of P_i. The coefficient from a linear regression of observed placement on P_i is reported in the figure. Panel (b) has the same structure as Panel (a), but the vertical axis variable reflects the probability of enrolling in UPK.

C Details on other subsidized programs

This appendix describes how we calculate enrollment in various subsidized pre-kindergarten programs in New Haven and provides additional information on those programs. We aim to estimate the number of three- and four-year-old children enrolled in each subsidized pre-kindergarten program in New Haven from 2006 through 2022. We use data from the Connecticut State Department of Education (SDE), NHPS' Public Schools Information System (PSIS), Connecticut's Office of Early Childhood (OEC), and aggregate data from the federal Office of Head Start (OHS) to categorize five main program types and determine enrollment figures for each.

The five main program types are as follows. First, several local centers offer Head Start programming. These programs are means-tested and provided mostly part-day slots in 2013 and earlier before switching to mostly full-time slots of at least six hours per day thereafter.[18] NHPS runs some of these centers at school facilities, while others are run by third-party providers. Second, NHPS offers subsidized childcare through a state-funded "School Readiness" grant. School Readiness programs are run by NHPS but are either not extended-day, not free, or both (NHPS Office of Early Childhoood, 2024). The School Readiness program also provides subsidized spots at private providers. Third, the state runs a program called "Care 4 Kids" that provides means-tested vouchers for families to use at private pre-kindergartens.[19] Fourth, the state offers a variety of smaller subsidized programs, often payable to daycare centers enrolling lower-income students. Fifth, families may enroll their children in private pre-kindergarten programs without any subsidy. Online Appendix C details the different program types and the data available for each.

See Section 3 for more details on each data source, and Section 2 for discussion of the counts we obtain from this exercise, which are reported in Figure 1, Panel (c). The main point to take away from this figure is that it is surprisingly hard to count subsidized pre-kindergarten enrollment using administrative records. While we have complete data on enrollment in UPK programs, School Readiness programs, and Care 4 Kids programs from 2006-2021, state- and district-level datasets do not include student records from the major New Haven Head Start programs after 2018. We estimate counts for 2019 and later using aggregate data from federal Office of Head Start records. Similarly, we only observe records of enrollment in non-Care 4 Kids subsidized programs from 2013 to 2019, which make up the "other" category in the graph. The limited availability of microdata contrasts with the public K-12 setting and highlights the importance of augmenting administrative data with survey reports.

C.1 UPK enrollment

See Section 2 for details on New Haven's UPK program. To calculate UPK enrollment, we use data from the SDE and the PSIS to track enrollment in New Haven UPK programs in each school year.

C.2 School Readiness

School readiness programs are funded through grants to high-need communities. To be eligible, a school must serve one of the 50 lowest-wealth towns in Connecticut, or be designated as a priority school. School Readiness preschool programs vary in the hours of care they offer (full-time, school-time, part-time, wrap-around) and family

[18]Source: authors' calculations from Office of Head Start data.

[19]In addition to income requirements, Care 4 Kids requires the parent to be working or in an approved educational program.

fees depend on the type of program, its duration, and family size. For example, for a full-time program, those earning less than 12% of State Median Income (SMI) have a weekly family fee that is equivalent to 4% of family income (if attended for the full year) and slowly increases up to 10% for those making more than 150% of SMI (Connecticut Office of Early Childhood, 2024).

For School Readiness program enrollment, we rely on data from SDE, PSIS and OEC. School Readiness enrollment is recorded in PSIS data (for children enrolled in New Haven), in SDE data (for children enrolled in publicly-run non-NHPS programs), and OEC data (for children enrolled in subsidized programs run by other providers). SDE and PSIS data are available across the full analysis period, but OEC data are available only from 2013-2019 and are most complete from 2016-2018. We keep data from facilities based in New Haven. We erase duplicates in cases where we can identify the same child and school year combination in multiple datasets. Using the deduplicated data, we count the number of unique children across years.

Data on enrollment in OEC files do not include identifiers for children enrolled in subsidized programs who did not match to our record of applicants to the New Haven UPK lottery. The data are identified at the level of the enrollment spell. It is possible one child may account for multiple enrollment spell observations. We estimate the number of unique children from the spell data using the ratio of spells to unique applicants among children that did match our records, then compute total School Readiness enrollment in each year by adding the scaled number of spells for non-applicant children to the number of unique individuals from our applicant records.

C.3 Care 4 Kids

Care 4 Kids offers families vouchers that can be used to pay for childcare. The program is primarily funded by the Federal Child Care Development Fund (CCDF) Plan. In addition to income requirements, Care 4 Kids requires the parent to be working or in an approved educational program. To be eligible, new applicants must earn less than 60% of the State Median Income (SMI), must remain below 85% of SMI while enrolled, and must be below 65% of SMI for redetermination (Care 4 Kids, 2024). Families must additionally pay a fee that is tied to annual gross income. This is 2% annually for those below 20% SMI, 4% for those above 20 but below 30% SMI, 6% for those above 30 but below 40% SMI, 8% for those above 40 but below 50% SMI, and 10% for those above 50 and below 85% SMI.

To estimate Care 4 Kids enrollment, we use data from the OEC to identify children enrolled in Care 4 Kids programs. Information on enrollment in Care 4 Kids is split between two datasets from the OEC, which jointly span the full analysis period. Unlike information used to calculate enrollment in the School Readiness program, here we cannot identify the town where a facility is located for most facilities. We can observe the town of residence for the child. Our approach is to include enrolled children in the program that matched to our record of applicants to the New Haven lottery, and from this set, we keep only those whose facility is in New Haven, if we can identify its location, or those that are New Haven residents in cases where we don't know where their facility is located.

Information on enrollment in Care 4 Kids programs does not include a variable that identifies individuals for those who did not match our record of applicants but are New Haven residents. The data are identified at the level of the enrollment spell. It is possible one child may account for multiple enrollment spell observations. We estimate the number of unique children from the spell data using the ratio of spells to unique applicants among children that did match our records, then compute total Care 4 Kids enrollment in each year by adding the scaled number of spells for non-applicant children to the number of unique individuals from our applicant records.

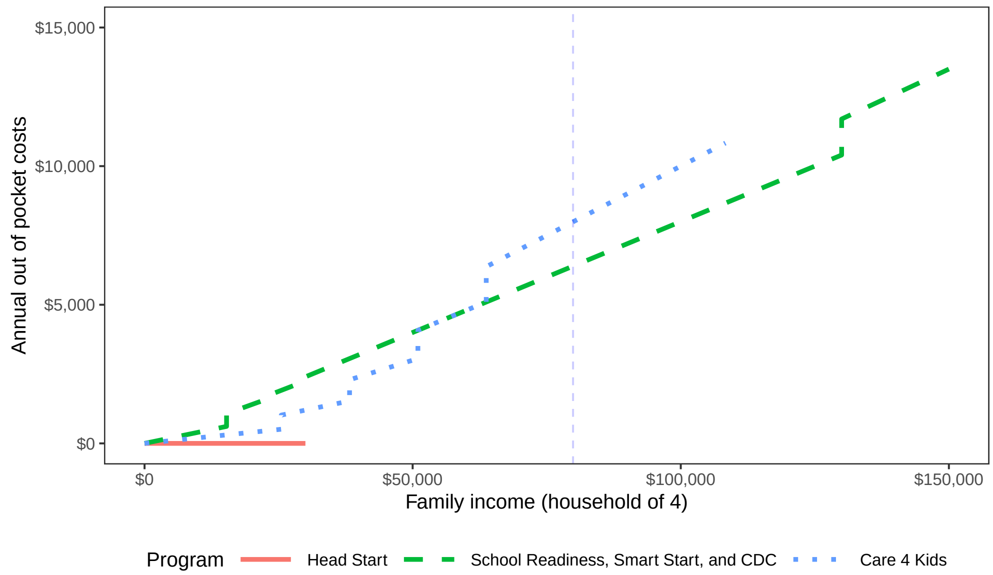
C.4 Head Start

Head Start is a nationally funded program for families with household incomes below the federal poverty line. In Connecticut, a child is also eligible if their family receives TANF or if they are homeless or in foster care. Head Start is free to eligible families. Head Start programs in New Haven provided mostly part-day slots in 2013 and earlier before switching to mostly full-time slots of at least six hours per day thereafter (Source: authors' calculations from Office of Head Start data.). See Friedman-Krauss et al. (2022) for additional details on Head Start and Head Start in Connecticut.

To build estimates of Head Start enrollment in New Haven, we use four data sources. First, we use data from the SDE and PSIS to identify children enrolled in Head Start programs administered by public school systems. We then supplement with OEC data to capture additional enrollees from the New Haven choice process that we might not cover using the SDE source.

The public and state sources of Head Start data have some limitations. In 2019, one of New Haven's largest Head Start providers, LULAC, leaves both district and state datasets despite continuing to operate. This appears to be related to the source of funding for the programs provided by LULAC. To address this issue, we use aggregate records of Head Start enrollment in New Haven from OHS. These data report enrollment counts by year and center and generate counts close to those we observe in administrative data for years 2018 and earlier. We construct predicted enrollment using the OHS data by regressing our observed enrollment counts on OHS counts using pre-2018 data. We then use predicted values from this regression to estimate Head Start enrollment counts for 2019 through 2022.

Figure C.1: Payment schedules and eligibility



Notes: This figure plots the annual out-of-pocket cost for families in Head Start, Care 4 Kids, and other OEC programs (School Readiness, Smart Start, and CDC) as a function of household income. Numbers are for a family of four enrolling in a full-time preschool program. All amounts and thresholds are based on 2023 values. The vertical line is the eligibility cutoff for new applicants for Care 4 Kids. Sources: Care 4 Kids (2024); Connecticut Office of Early Childhood (2024)

C.5 Other programs

The OEC data includes information on other smaller programs in the state. In the category of "Other programs", we include children enrolled in the rest of subsidized options from the OEC data. This includes children enrolled in Smart Start programs, Child Day Care Contracts (CDCC), and the Preschool Development Grant (PDG) across years that matched to our records of applicants to the New Haven UPK lottery.

We also consider deflected spells to account for children enrolled in these programs that did not match to our records of applicants to the New Haven UPK lottery. We calculate total enrollment in these programs by adding the number of deflected spells and unique individuals that matched to our record of applicants whose facility is based in New Haven.

C.6 Payment schedules

To compare out-of-pocket costs for School Readiness, Care 4 Kids, and Head Start, Figure C.1 plots the annual out-of-pocket cost to families for enrolling in each of these programs as a function of family income. The figure is for a household of four enrolling their child in a full-time preschool (ages 3 or 4). The dashed vertical line is the eligibility threshold for Care 4 Kids for new enrollees.

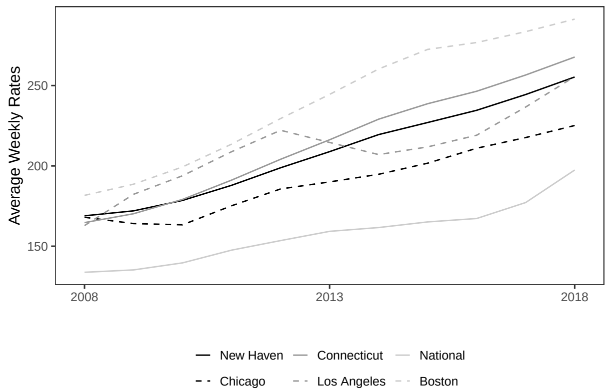
D Childcare costs over time

This appendix provides estimates of childcare costs over time from three sources, all of which provide evidence of increasing childcare costs. First, we use the National Database on Childcare Prices to estimate changes in the market rate for center-based preschool care. Panel (a) of Figure D.1 plots the market rate for center-based preschool in New Haven County (solid black line), Connecticut (dark gray line), and in the US overall (light gray line) from 2008 to 2018 in 2015 dollars. We additionally include dotted lines for Chicago (Cook County), Los Angeles (Los Angeles County) and Boston (Suffolk County). New Haven County is above the national average, but somewhat below the average for Connecticut. New Haven is somewhat above Chicago, similar to Los Angeles, and somewhat below Boston.

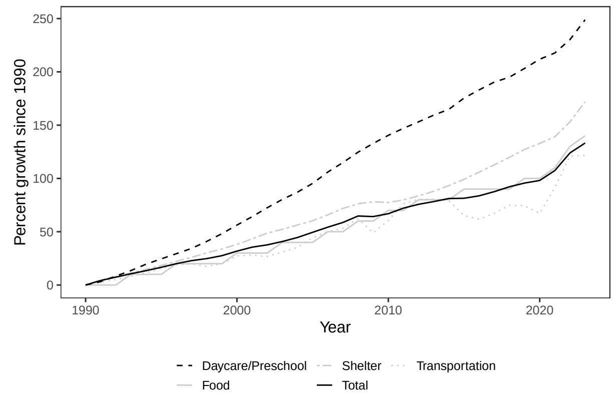
Second, we plot price changes using the seasonally adjusted CPI-U for US cities to plot the Full-basket CPI and the CPI for Daycare/Preschool, Shelter, Transportation, and Food. Specifically, we plot the percent change since 1990 for each basket from 1990 in Panel (b) of Figure D.1. We see that the CPI for Daycare/Preschool increased much more rapidly than the total or any of the other categories we consider, rising by 250% nominally since 1990. The total CPI-U increased by 130% in the same time period.

Third, we use the Current Population Survey Annual Social and Economic Supplements (CPS ASEC) to calculate households reported spending on childcare, restricted to households with children under the age of five. Total expenditure on childcare is divided by the number of children to approximate average cost per child. Panel (c) of Figure D.1 shows that per-child household expenditure on childcare increased notably since 2005, with a drop in 2020, likely due to the Covid-19 pandemic.

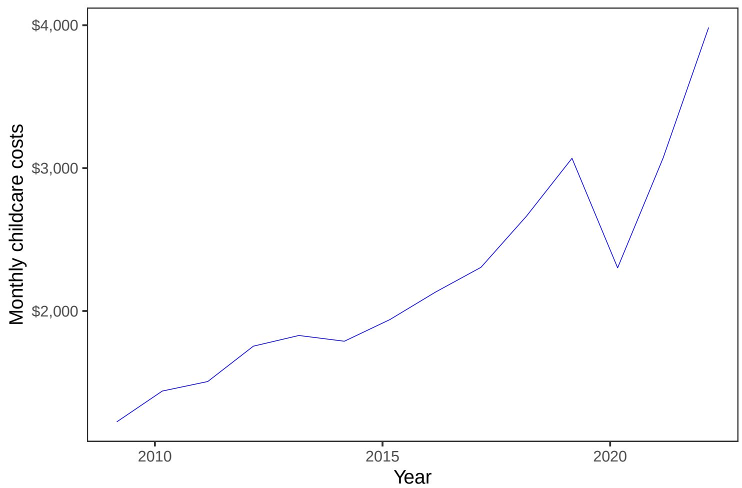
Figure D.1: Growing childcare costs over time



(a) Market rates for formal preschool care in from 2008 to 2018



(b) Percent growth in CPI-U for daycare/preschool vs other goods



(c) Average annual household expenditure on childcare per child in households with children under the age of five

Notes: Panel (a) plots weekly market prices for center-based preschool-aged childcare. The plot contains the following lines: New Haven County (solid black line), Connecticut (dark gray line), and Nationally (light gray line) from 2008 to 2018 in 2015 dollars. We additionally include dashed lines for Chicago (Cook County, black), Los Angeles (Los Angeles County, dark gray) and Boston (Suffolk County, light gray). Source: National Database on Childcare Prices (Landivar et al., 2023). Panel (b) plots the percent change since 1990 for the seasonally adjusted CPI-U for all goods (Total SA0), Daycare/Preschool (DAYCARE / PRESCHOOL SEEB03), Shelter (SHELTER AH1), Transportation (TRANSPORTATION AT), and food (FOOD AF1). Source: Bureau of Labor Statistics Consumer Price Index (CPI). Panel (c) plots the average monthly household spending on childcare per child for households with children under the age of five using data from the Current Population Survey Annual Social and Economic Supplements (CPS ASEC). Amounts are in 2015 dollars.

E Student achievement data

E.1 Test score data

This section describes our data on student achievement. We pay particular attention to achievement measures for kindergarteners, because these are the earliest measures we observe and likely also the least familiar to readers.

The earliest assessment we observe is the Connecticut Kindergarten Entrance Inventory (KEI). Kindergarten teachers conduct this assessment in the fall of each academic year. Students receive scores ranging from one to three on six readiness measures: creativity, language proficiency, literacy, numeracy, personal/social readiness, and physical readiness.

The Fall Kindergarten Entrance Inventory (KEI) is an evaluation performed at the beginning of the school year to document the skills students demonstrate upon arriving in kindergarten. The inventory was introduced in 2007 to comply with new state regulations and is administered statewide (Connecticut State Department of Education, 2021, 2024). Scores are based on teachers' evaluations across six domains: Creative/Aesthetics, Language, Literacy, Numeracy, Personal/Social skills, and Physical/Motor skills. Students are rated on a 3-point scale reflecting the degree to which a child demonstrates the skills in a given domain and the amount of instructional support required. To give an example, for Language the teacher is asked the following question: At what level does the student:

- Participate in conversations
- Retell information from a story read to him/her
- Follow simple two-step verbal directions
- Speak using sentences of at least 5 words
- Communicate feelings and needs
- Listen attentively to a speaker.

The teacher then provides a single score of one, two, or three, where a higher score represents a more consistent demonstration of the skills and that the student requires less instructional support for the skills (Connecticut State Department of Education, 2008). Our data covers results for the KEI for the school years 2008/09 through 2018/19, 2020/21, and 2021/22. We standardize test scores by school year and subtest using data from all New Haven pre-kindergarten students. We do not use the 2020/21 years in our analysis because Covid school closures interfered with the administration of the evaluation.

Figure E.1 plots the average of the six raw scores for students in our sample, while Table E.1 reports the min, max, and quartiles of the normalized KEI scores for our sample.

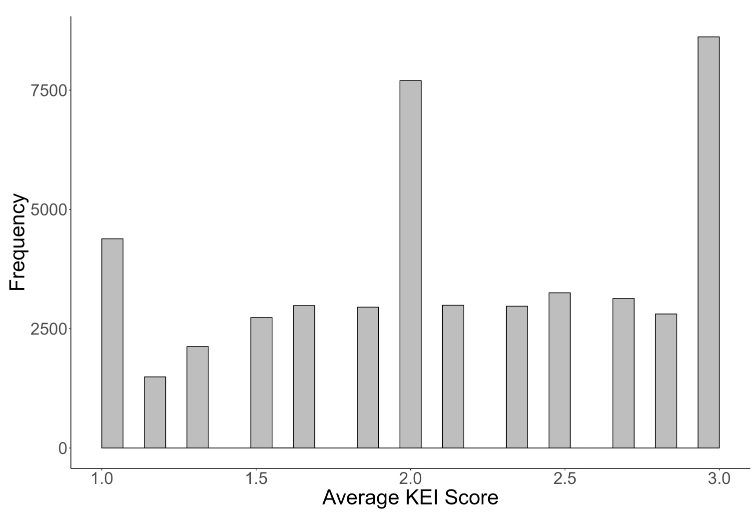
We also observe results from the Connecticut Mastery Test (CMT) and Smarter Balanced Assessment (SBA). These are high-stakes exams administered in grades three through eight that form the basis for Connecticut's school accountability system. We observe these scores from 2007-08 through 2021-22.

All test score data is missing for the 2019-20 school year due to the state Covid response; many Connecticut schools shut down in March 2020 just prior to the scheduled administration of most standardized exams. In addition, school closures in Fall 2020 interfered with the in-person administration of the KEI for the 2020-21 school year.

E.2 Testing KEI scores

Kindergarten readiness scores are strongly correlated with later achievement measures. Figure E.2 plots the relationship between the average KEI score in kindergarten against average test scores in grade 3, grade 8, and grades 3-8. Each subpanel plots the standardized average KEI score on the x-axis and the standardized other test score on the y-axis. The dots are binned means of the test listed in the title within quantiles of KEI score, while the line from a linear regression of the later test score on the KEI score. On each plot we report the slope of the linear regression (beta). Overall, there is a strong, linear relationship between the KEI score and later test scores, though the slope of line of best fit decreases as more time has passed, with a slope of 0.40 in 3rd grade and 0.27 in eighth grade.

Figure E.1: Distribution of KEI Scores



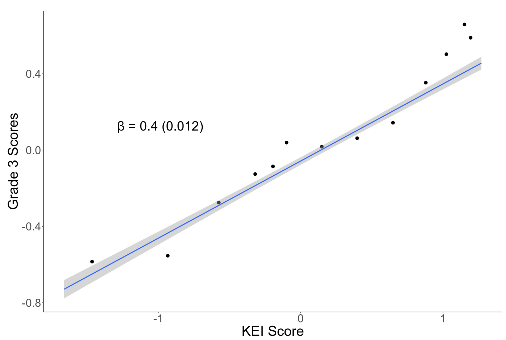
Notes: This figure shows the distribution of raw KEI scores, averaged across all six subtests, among all students in our SDE sample. The maximum attainable score in each subtest is 3, the minimum attainable score is 1. The time period covered is 2008/09-2018/19 and 2020/21-2021/22.

Table E.1: KEI Subtest Score Distribution

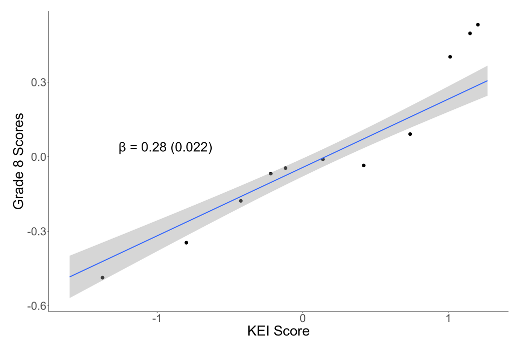
Subtest	Min	25th	Median	75th	Max
Average	-1.66	-0.40	0.09	0.94	1.27
Creative	-1.89	-0.38	-0.32	1.04	1.17
Language	-1.45	-0.18	-0.02	1.23	1.35
Literacy	-1.52	-0.25	-0.05	1.20	1.38
Numeracy	-1.56	-0.19	-0.07	1.19	1.33
Personal	-1.62	-0.23	-0.13	1.16	1.28
Physical	-1.96	-0.42	-0.29	1.02	1.13

Notes: This table shows the minimum, 25th percentile, median, 75th percentile, and maximum of the standardized KEI score for the overall average across all six subtests and for each subtest. The underlying sample consists of all students contained in our SDE dataset. The time period covered is 2008/09-2018/19 and 2020/21-2021/22. Scores are standardized at the subtest-school year level.

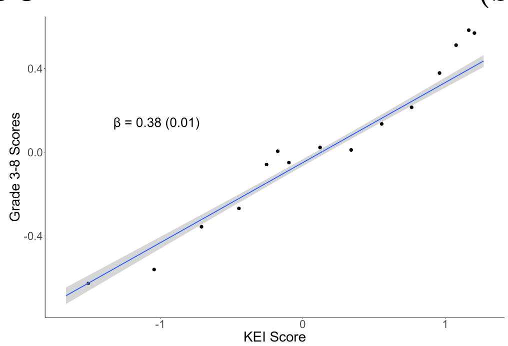
Figure E.2: KEI Score and Average Upper Grade Score



(a) KEI and Grade 3



(b) KEI and Grade 8



(c) KEI and Grades 3-8

Notes: This figure shows binscatter plots visualizing the correlation between the overall standardized KEI score, i.e. the average over all six standardized subscores, and a selection of standardized test scores from later grades. The x-axis in each panel reflects the standardized KEI score. For the y-axis, Panel (a) presents the grade 3 test scores, Panel (b) grade 8 test scores, Panel (c) shows the average test score across grades 3 through 8. Each subfigure additionally plots the regression line and reports the slope of the line.

F Survey procedures, variable construction, and match quality

F.1 Earnings match

Two descriptive exercises suggest our merge procedure of administrative, earnings, and survey data generates accurate matches. The first explores how earnings evolve for male and female parents around the time of childbirth. Previous research shows that earnings for mothers drop relative to earnings for fathers following the birth of a child and do not fully recover (Kleven et al., 2019). As reported in Panels (a) and (b) of Figure F.1, we observe this pattern in our data. Earnings for female parents fall by $6,000 relative to male parents in the academic year the applicant is born, 19% of the mean two years prior to birth. This gap persists for at least five years.

The second exercise compares individual earnings reports measured in the two years prior to application to median household income in the Census block group where the applicant lives. If our merge procedure generates accurate matches, we would expect to see a strong positive relationship since people with higher incomes tend to live in neighborhoods with higher household incomes. This is what we see. As reported in Panel (c) of Figure F.1, a $1,000 increase in median neighborhood household income is associated with a $360 (SE=$15) increase in individual earnings. Panel (d) reports similar results but uses household income as reported in the survey instead of from administrative records. We again see a positive relationship, a $1,000 increase in median neighborhood household is associated with a $560 (SE = $54) in individual earnings.

F.2 Survey implementation

We worked with NHPS to survey the parents of past UPK applicants. The survey ran from May to November 2023. Using contact information provided on application forms, survey enumerators at NORC emailed parents of past applicants and followed up by phone with applicants who did not complete the survey or lacked email addresses. We prioritized phone follow-ups for parents whose children had interior placement probabilities to maximize statistical power within budget constraints. Survey respondents were rewarded with a chance to win a prize of $100. We awarded 25 such prizes.

The survey launched in early May with emails to all parents in the survey sample. Emails highlighted the collaboration between NHPS, Yale, and NORC. The email also emphasized that responses would help improve pre-kindergarten programs in New Haven and that responses were confidential. See Figure F.2 for the email invitation. Upon clicking the email invitation link, respondents were directed to the landing page that reiterated the purpose of the survey, provided contact information for the research team, and asked for consent to continue. Each legal guardian with an available email received up to five reminders. Survey efforts continued until the second week of November.

Survey logic allowed for different questions based on the application status of the child, distinguishing between those children who applied but were not offered a seat, those who were offered a seat but chose not to enroll, and those who enrolled after receiving an offer. The survey questionnaire is available at https://github.com/economics/prek_survey.

F.3 Survey variable construction

In most cases our analysis relies on direct reports from survey responses. In some cases we combine responses from multiple questions into summary variables. We describe pertinent data construction choices below.

F.3.1 Pre-kindergarten outside options

If a survey respondent indicated that their child did not enroll in a UPK program, they were asked about the kind of childcare program their child was enrolled in. The answer options for this question consisted of

1. Head Start/Early Head Start
2. Another childcare center or pre-k (not Head Start)
3. A paid childcare provider operating out of their home (not Head Start)
4. Another town's public pre-k or childcare program
5. Babysitter, nanny, or another private option
6. Other, please specify:

For our analysis, we code these responses into three categories as follows:

- *Head Start*: response option 1
- *Other Public PreK*: response option 4
- *Other Paid PreK*: response options 2, 3, 5

We additionally hand-code all free text responses to answer option 6.

Finally, if a survey respondent indicates that their child did not enroll in any of these childcare options we assign that child to the category *Kid Stays Home*.

F.3.2 Out-of-pocket costs

Out-of-pocket (OOP) costs for non-UPK pre-kindergarten programs are taken directly from survey responses. Survey respondents could choose between seven bins to describe their monthly OOP costs, with a low value of zero dollars, a high value of \$2,000 or more, and five intermediate choices corresponding to ranges of dollar values. In our analysis we assign dollar values based on the midpoint of the selected bin for the middle bins and assign a dollar value of \$2,000 for the top bin. Figure F.3 reports the average monthly out-of-pocket costs reported by survey participants who did not enroll in UPK, categorized by the type of childcare they used that year. These costs include expenditures on care beyond the primary care reported in the survey, which explains why parents of Head Start students-despite Head Start being free-still incur some out-of-pocket costs. Due to an error in survey logic, we did not collect information on OOP costs for children enrolled in UPK programs. While enrollment in magnet programs is free of charge and includes wraparound before- and after-care prior to 2021 (i.e., for almost all of our sample period), families may nevertheless incur other childcare costs. To address this issue, we impute OOP costs for UPK enrollees as the average OOP costs reported for children enrolling for Head Start programs. We choose Head Start because the programs are also free to participants. The key difference in structure is that Head Start programs do not provide as much extended-hours care as UPKs through most of our sample period. We therefore regard this approach as conservative in the sense that it is likely to overstate the OOP costs for UPK enrollees.

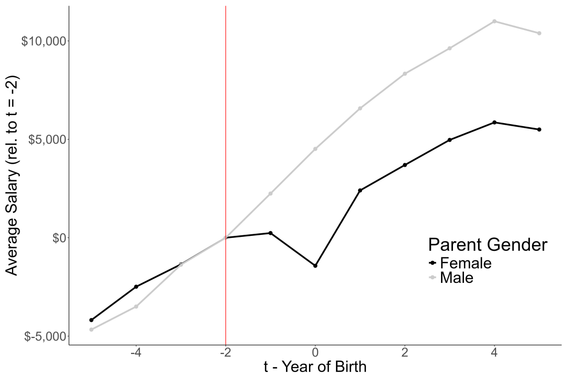
F.4 Survey response reliability

Several tests support the idea that survey participants provided accurate responses related to their schooling choices and earnings. First, reported enrollment from the survey is consistent with the enrollment choices that we observe in administrative records. Overall, survey reports of UPK enrollment match administrative data in 89.6% of cases in the pre-2021 application years, for which we can match the survey to administrative records. Of respondents who report that their child did not enroll in a UPK program, we observe UPK enrollment in only 2.4% of cases, for a 97.6% accuracy rate. Of respondents who report their child enrolled in a UPK program in the year following application, we confirm UPK enrollment for 82.4%. We regard this number as high given the many potential sources of confusion for parents about which program their children were enrolled in. For example, of the 17.6% of "false positive" UPK enrollees, 28.3% enrolled in a non-UPK school readiness program physically located in one of the magnet schools or enrolled in UPK in a different year than the one asked about in the survey. Another 34.8% of these students enrolled at some other form of non-universal pre-kindergarten offered by NHPS. Confusion on the part of parents about the administrative classification of the pre-kindergarten program their children attended seems reasonable. If one re-classifies cases of reasonable confusion as correct, the accuracy rate rises from 89.6% to 94.6%. Among those reporting UPK enrollment, 82.4% are confirmed in administrative records. Many "false positives" are reasonable misclassifications: 28.3% of the false positives enrolled in non-UPK programs located in magnet schools or enrolled in UPK in a different year, while 34.8% enrolled in other NHPS pre-kindergarten programs. If one re-classifies cases of reasonable confusion as correct, the accuracy rate rises from 82.4% to 93.5%.

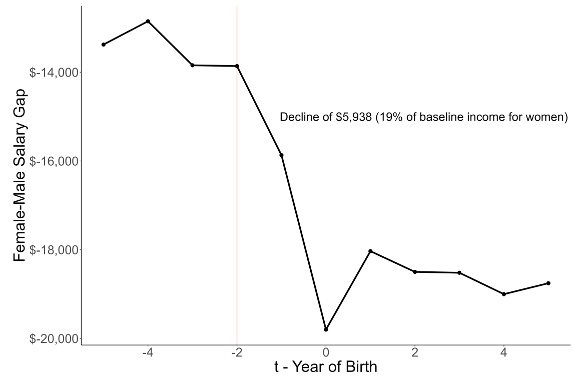
F.5 Survey Balance

As shown in Table 2, survey respondents differ from the full sample of lottery applicants in several ways. Respondents were more likely to be White, female, reside in higher-income neighborhoods, and have been placed in a magnet school. To ensure the validity of estimates using survey data, we compare our main results to results using a reweighted survey sample that reweights the survey responses to match the full-sample characteristics on: child race, child gender, New Haven residency, and neighborhood income tercile. Table F.1 presents IV estimates for both the unweighted and reweighted survey samples. After reweighting, substitution patterns shift slightly, and estimated effects on out-of-pocket costs decrease. These changes are consistent with greater weight being placed on lower-income individuals, who are more likely to transition out of Head Start or other subsidized programs rather than private pre-K. Importantly, the main results-such as increased hours worked and greater childcare coverage-remain largely unchanged, reinforcing the robustness of our findings.

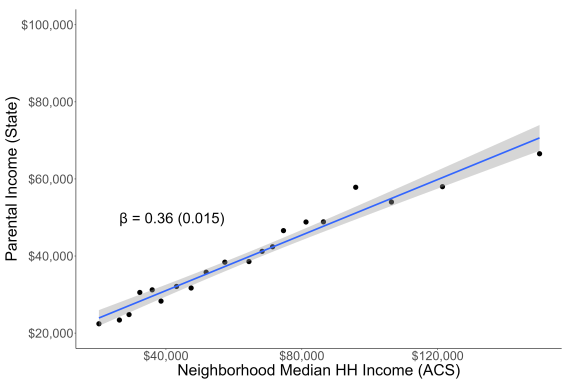
Figure F.1: Validating administrative and survey data



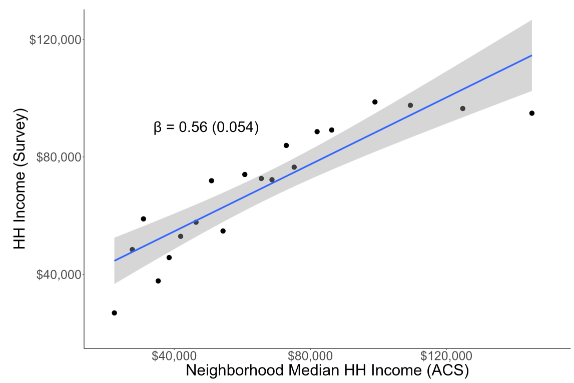
(a) Mothers' and fathers' income around childbirth



(b) Parental income gap around childbirth



(c) Parental income (state) vs. HH income (ACS)



(d) HH income (survey) vs. HH income (ACS)

Notes: Panels (a) and (b) show the evolution of mothers' and fathers' incomes around the year of birth for future UPK applicants. Panel (a) normalizes incomes to 0 in period t = -2 and plots mothers' income in black and fathers' income in gray. Panel (b) plots the gap between moms' and dads' incomes over time. The red vertical line indicates two years before a child's birth. Panels (c) and (d) show binscatter plots of median household income at the Census block level from the ACS against parental income taken from administrative records (Panel c) and household income reported in the survey (Panel d). Panels (c) and (d) also plot the regression line and report the slope. See Section 3 for details.

Figure F.2: UPK survey invitation

NEW HAVEN PUBLIC SCHOOLS

Yale University

New Haven Public Schools Pre-K Survey

Dear Parents,

Records indicate that your household previously applied to a pre-kindergarten lottery within New Haven Public Schools. We are inviting you to take a survey to help us better understand the impacts of pre-kindergarten programs being offered by New Haven Public Schools. Those who complete the survey will be entered into a raffle to win one of 25 $100 prizes.

Begin your Survey

Your feedback is important and will be used to improve pre-k programs in New Haven Public Schools and help better understand how the NHPS pre-kindergarten programs help families.

Your responses are CONFIDENTIAL. This means that your individual responses will not be shared. Neither your name nor your child's name will be shared with anyone outside of the research team.

The survey is around 10 minutes long, and those who complete the survey will be entered into a raffle to win one of 25 $100 prizes.

Begin your Survey

To learn more about this survey and research project, please visit tobin.yale.edu/preksurvey.

If you have questions or concerns, please contact the New Haven School Choice Research Team from Yale University at (203) 436-4477 or nhps_prek_survey@yale.edu.

Sincerely,

New Haven School Choice Research Team
Yale University

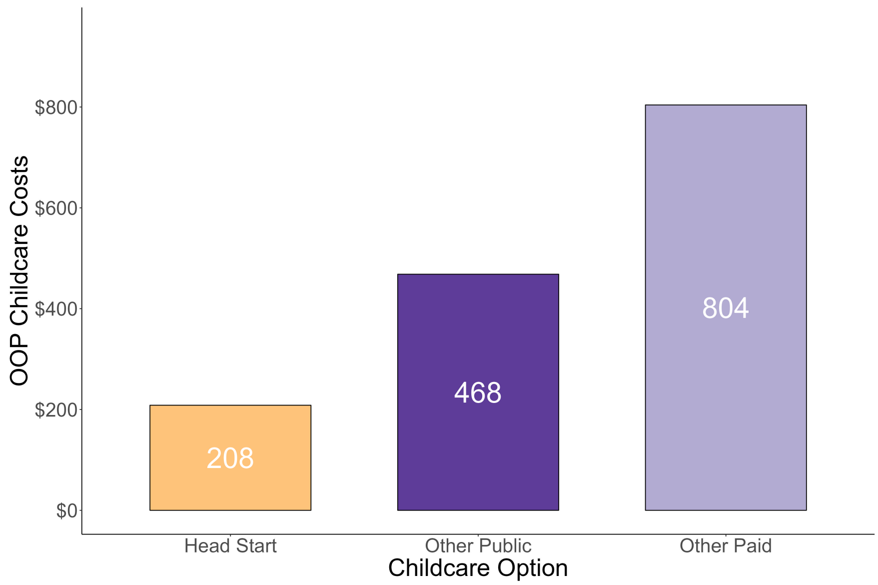
Yale University NORC at the University of Chicago

This email is intended for the parent of **[CHILDNAME] ([MEMBER_PIN])**.

Having trouble with the link? You can copy and paste this into your browser: https://survey.amerispeak.org/SE/?[$PURLPARAMS] Need more help? Email support@AmeriSpeak.org or call toll-free (800) 993-0495. **NORC Support**: 55 E Monroe St, 19th Floor, Chicago, IL 60603.

Notes: This is the email sent to parents or guardians of past UPK applicants inviting them to participate in the survey. Respondents accessed the survey by clicking on any of the buttons displayed, and subsequently were directed to a landing page where they received additional information on the purpose of the survey.

Figure F.3: Out-of-pocket costs by childcare type



Notes: This figure shows the composition of outside options for childcare and the out-of-pocket (OOP) costs associated with them based on data from our survey of lottery applicants. See Section 3 for details.

Table F.1: Reweighted Survey Sample

	Unweighted Sample		Reweighted Sample	
	Control Mean	IV	Control Mean	IV
Panel A: Labor Supply				
During Pre-k				
Employed full-time	0.657	0.221	0.637	0.222
		(0.138)		(0.155)
Employed part-time	0.170	-0.133	0.178	-0.102
		(0.119)		(0.132)
Hours worked per week	31.62	12.80	31.14	14.29
		(4.25)		(4.47)
After Pre-k				
Employed full-time	0.740	0.005	0.732	-0.048
		(0.106)		(0.124)
Employed part-time	0.140	0.081	0.146	0.118
		(0.090)		(0.102)
Hours worked per week	33.74	1.48	33.48	0.26
		(3.91)		(4.32)
Panel B: Childcare Usage				
Substitution Patterns				
Any pre-k or childcare (survey)	0.877	0.022	0.867	0.016
		(0.044)		(0.050)
Enrolled Head Start (survey)	0.217	-0.192	0.270	-0.267
		(0.068)		(0.074)
Another public option (survey)	0.094	-0.117	0.091	-0.104
		(0.051)		(0.047)
Other paid option (survey)	0.539	-0.623	0.480	-0.572
		(0.082)		(0.082)
Usage Intensity				
Weekly childcare hours (survey)	28.9	11.3	28.6	10.8
		(3.2)		(3.2)
Monthly OOP costs (survey)	617	-375	537	-293
		(90)		(79)

Notes: This table reports IV estimates of Equation 1 for UPK enrollment substitution patterns and employment, analogous to the results in Tables 3 and 4, among survey respondents. Columns 1 and 2 present control means and IV estimates using the main survey dataset. Columns 3 and 4 show control means and IV estimates using a reweighted dataset that more closely matches the full lottery applicant sample across key characteristics. All specifications control for student sex and race/ethnicity. The probability of winning the lottery is accounted for using the recentered instrument approach discussed in Section 4.1. Standard errors are clustered at the student level.

G Pre-Kindergarten Hours

G.1 Childcare Hours Construction

Main Method

We assign each student in our survey dataset a value reflecting the daily hours of care the student has access to at their childcare facility in the year following application to the New Haven choice process (i.e., the first year they might be enrolled in a UPK program). We draw on and/or construct program-year level data from a variety of sources, including NHPS, the Connecticut Office of Early Childhood, and the Office of Head Start. To construct our main measure, we follow the steps outlined below.

1. Students with no reported pre-kindergarten are assigned a value of zero hours.
2. Students reported as enrolled in a UPK program are assigned a value of 10 hours for the school years 2003-04 through 2020-21 and a value of 6.5 hours for 2021-22 and 2022-23. These numbers reflect the official minimum hours offered by UPK programs during those periods. The 6.5-hour figure during 2021-22 and later is a lower bound: many programs in practice did offer before- and after-care during this period.
3. For students reporting enrollment in other programs that appear in OEC records, we assign the modal value of hours among all students enrolled in a given program in a given year. If the OEC records don't indicate the number of childcare hours directly, we assign the midpoint on the range of hours that regulations permit.[20]
4. Remaining students without any information on pre-kindergarten hours who report in the survey that they have enrolled in a Head Start program are assigned the modal hours recorded for their respective program taken from the Office of Head Start (OHS).[21]
5. For all remaining students without any information on pre-kindergarten hours, we look up the childcare programs reported in the survey and use the hours currently offered by the program.

Alternative Methods

We prepare a variety of alternative childcare hours measures that are used for robustness checks of our main measure:

[20] See https://www.ctoec.org/wp-content/uploads/2020/02/GP-B-04-Definition-of-Space-Types-Categories-of-Care-and-Eligibility-for-Enrollment_FY25.pdf and https://www.ctcare4kids.com/wp-content/uploads/2021/12/sample-certificate.pdf (retrieved 25.06.2024).

[21] Office of Head Start, Program Information Reports 2008-2022 (https://headstart.gov/program-data/article/program-information-report-pir).

- *Lower/Upper Bound Hours*: Instead of using the average number of hours observed for a given childcare program, we take the minimum/maximum value.
- *Childcare 211 hours*: We directly collect offered childcare hours, as reported by Connecticut 211 Childcare, for the childcare providers survey respondents reported using.[22] For the majority of centers, we use their reported daily hours in this new data source. However, we manually adjust some of the hours for a small number of children based on several criteria.
  - For providers missing hours information from the 211 website, we use data collected from the centers' websites.[23]
  - For cases where the childcare name indicated no formal childcare (e.g., home-schooling or care by a family member), we assume zero hours.
  - We assign 10 hours to children who are flagged as UPK students in the survey if the application year is 2020 or earlier. For application years after 2020, we assign 6.5 hours.
  - Children who are flagged as Head Start students in the survey with missing hour are assigned hours from the main method unless the childcare name is "Reggie Mayo", in which case we assign 6 hours. If both hours measures are missing, we assign 7 hours as a default.

  Overall, this approach is conservative as CT 211 Childcare typically reports the maximum possible hours offered by the center, likely overstating the hours accessed by those not in UPK.

Finally, we similarly impute missing childcare hours using several different approaches:

- *Zero Hours*: Impute missing childcare hours as 0.
- *Average Non-UPK Hours*: Impute missing childcare hours as the average number of hours observed among all non-UPK childcare enrollees.
- *Median Non-UPK Hours*: Impute missing childcare hours as the median number of hours observed among all non-UPK childcare enrollees.
- *Maximum Non-UPK Hours*: Impute missing childcare hours as the maximum number of hours observed among all non-UPK childcare enrollees.
- *UPK Hours*: Impute missing childcare hours as the number of hours a UPK enrollee would have received in the same year.

[22] Connecticut 211 Childcare is a website administered by The United Way of Connecticut and supported by the Connecticut Office of Early Childhood to facilitate access to childcare (`https://www.211childcare.org/`). Data retrieved on 07.18.2024.

[23] The centers are Alice Peck Early Learning Center (9 hours - see website), and Overbrook Early Learning Center (10.5 hours - see website).

G.2 IV Results

We estimate alternate versions of the hours of care IV specification described in Section 4.3 and reported in Table 3 using the approaches to hours imputation described above. We report our findings for daily hours in Tables G.1 and G.2.

Table G.1: Daily childcare hours IV

	Childcare Hours			
	Main (1)	Lower (2)	Upper (3)	Childcare 211 (4)
UPK Enrollment	2.27	4.61	1.77	1.52
	(0.640)	(0.537)	(0.637)	(0.596)
Grade and Year	✓	✓	✓	✓
Race	✓	✓	✓	✓
Gender	✓	✓	✓	✓
Dependent variable mean	7.44	6.76	7.69	7.85
Observations	724	724	724	756

Notes: This table reports IV estimates of Equation 1 for UPK enrollment on childcare hours received among the sample of survey respondents. Column 1 presents results using our main measure of childcare hours. Columns 2-4 show results using a series of alternative methods to construct the measure of childcare hours. For details on the different methods, see Appendix G. All specifications control for student sex and race/ethnicity. We account for the probability of winning the lottery using the recentered instrument approach discussed in Section 4.1. Standard errors are clustered at the student level.

Table G.2: Daily childcare hours IV - impute missing values

	Childcare Hours, Missing Imputed as				
	Zero (1)	Average (2)	Median (3)	Max (4)	UPK (5)
UPK Enrollment	4.46	2.49	2.09	1.08	1.20
	(0.648)	(0.504)	(0.519)	(0.618)	(0.557)
Grade and Year	✓	✓	✓	✓	✓
Race	✓	✓	✓	✓	✓
Gender	✓	✓	✓	✓	✓
Dependent variable mean	6.49	7.23	7.38	7.77	7.69
Observations	830	830	830	830	830

Notes: This table replicates the IV estimation in Table G.1, column 1, but imputes missing values in the measure of childcare hours. For details on the different methods used to impute missing values, see Appendix G. All specifications control for student sex and race/ethnicity. We account for the probability of winning the lottery using the recentered instrument approach discussed in Section 4.1. Standard errors are clustered at the student level.

H Poisson specifications

We follow Lin and Wooldridge (2019) and estimate a Poisson specification using a control function approach. We estimate a linear first-stage equation given by

$$D_i = \delta Z_i + X_i'\pi + \sum_p \rho_p 1[P_i = p] + \eta_i. \tag{2}$$

We obtain the estimated residual $\hat{\eta}_i$, and include this residual as a control in the second-stage Poisson specification, given by

$$E[Y_i|D_i, X_i, P_i, \hat{\eta}_i] = \exp\left(\beta D_i + X_i'\Gamma + \sum_p \alpha_p 1[P_i = p] + \phi\hat{\eta}_i\right) \tag{3}$$

We estimate this specification using Poisson pseudo maximum likelihood, as implemented in Correia et al. (2020). We report $e^{\hat{\beta}} - 1$, which is an estimate of the proportional change $(E[Y(1) - Y(0)]/E[Y(0)])$. We compute bootstrapped standard errors clustered at the application level using 500 bootstrap samples. In contrast to our approach to linear specifications, we do not two-way cluster the Poisson specifications. Estimates of one-way clustered linear specifications suggest that clustering on the second dimension (parent identifier potentially spanning multiple applications) has little effect on estimated standard errors.

I Subjective effects and qualitative reports

This section describes the qualitative reports given by survey respondents when asked about their experience with UPK. The findings from our empirical analysis in Section 4.5 are consistent with survey respondents' stated priorities when forming preferences over preschool programs and with their beliefs about how access to UPK programming affected their lives. Panel (a) of Figure I.1 describes how respondents valued different program attributes when choosing where to enroll their child in pre-kindergarten. Our survey asked applicants to rate the importance of six different attributes on a 1-5 scale, with one being "not important" and five being "very important." Location and schedule, the two attributes we asked about that are most closely tied to work opportunities, are among parents' top priorities, with mean importance scores ranging from 4.3 to 4.6 across groups. These values are slightly below the average rating parents assign to teachers, where values range from 4.6 to 4.7. However, they are well above mean ratings for class size or school peers, two other potentially important academic inputs, for which reported values range from 3.5 to 3.9. Parents appear to be thinking about their own work lives when they make pre-kindergarten choices for their children.

Panel (b) of Figure I.1 describes parents' beliefs about how access to the NHPS UPK programs affected their lives (if they enrolled in a UPK program) or would have affected their lives (if they did not). 76% of parents whose children enrolled in a UPK program report that enrolling in the program allowed them to work more. Though somewhat below the 91% share who report that enrolling in a UPK program led to a better pre-kindergarten experience for their child and the 85% who report less stress about money, this is still a large share. Findings are similar for survey respondents whose children did not attend UPK: 67% think that attending would have helped them work more, 79% think it would have improved their child's pre-kindergarten experience, and 87% think it would have reduced financial stress. These reports support the findings from our main analysis that UPK enrollment raises parent earnings while reducing out-of-pocket costs. They also suggest that UPK programs may benefit children in ways that are not captured by test scores.

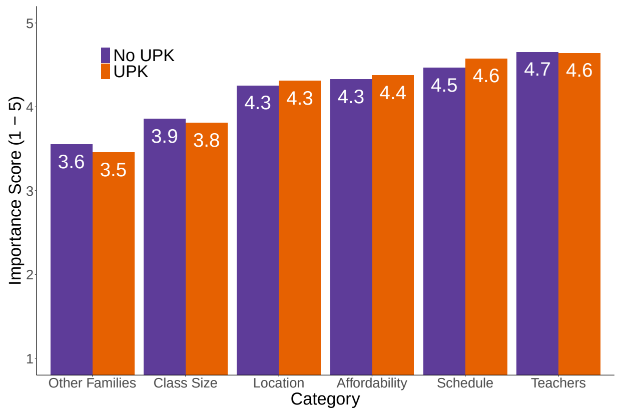
We asked parents who said that UPK enrollment did or would have helped them work more what kind of job changes UPK enabled. Panel (c) of Appendix Figure I.1 tabulates responses to this question. The modal response was that UPK allowed (or would have allowed) the respondent to increase their hours (54% in the untreated group, 33% in the treated group); many respondents also reported that UPK allowed another household member to increase their hours (16% in the untreated group, 15% in the treated group). These responses are consistent with our findings of increased labor supply during pre-kindergarten in both the administrative and survey data.

Qualitative responses provide some additional insight into the channels through which UPK access affected parents' labor market outcomes. Parents who did not fit into one of the survey-provided types of labor market gains (e.g., "increased hours," "got a full-time job," "switched part-time to full-time"; see Panel (c) of Appendix Figure I.1) were prompted to enter a text description of the job change, and many did.

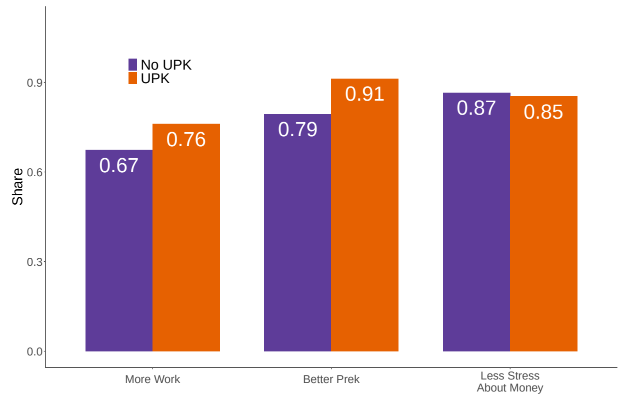
These responses illuminate the diverse channels through which UPK can improve current and future labor market outcomes and the challenges facing parents trying to purchase similar care on the private market. A major theme was the ability to maintain existing jobs. One respondent wrote, "I was able to continue working full time. Without the program I would've had to quit my job." Some respondents specifically credited wraparound care for increased hours. One wrote that "[because my child] was in before and after care programs, I was able to stay at work and not have to leave early" while another described how "after school hours were available which enabled us to pick up later." These reports included expressions of regret from those who did not gain enrollment: "I could have worked a different full time schedule—more normal hours. My husband and I ended up doing what felt like shift work [so that] I could pick up [my child] at 3:00 when the private preschool program ended." Others credited the program for improved productivity while they were working, writing, for example, that "[I] was better able to focus on work during the time worked" or that "I was able to focus more because I work from home." Finally, some respondents reported that UPK enabled them to invest in human capital likely to pay off in the future. One reported that she started an accelerated nursing program. Another described how they were "a full time college student and both parents graduated with bachelor's degrees."

These responses paint a picture in which UPK helps people work in their preferred jobs, work more, work more effectively, and invest more in their careers. This is consistent with our finding of sustained labor market gains after children age out of UPK.

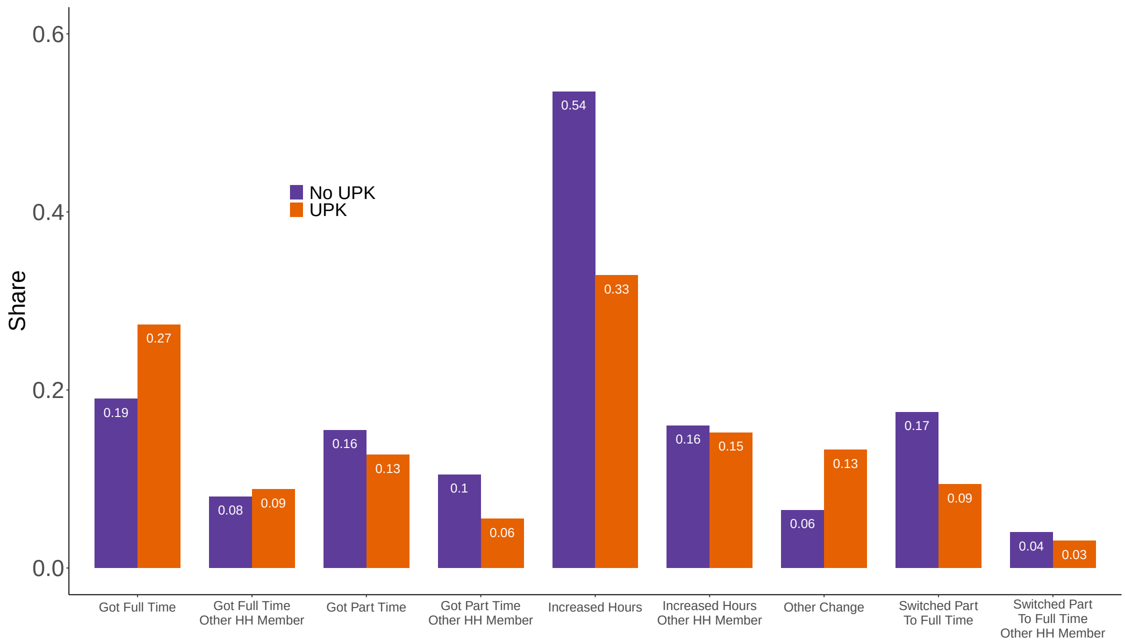
Figure I.1: Reported priorities and subjective treatment effects



(a) Survey reports of attribute importance



(b) Survey reports of subjective treatment effects of UPK



(c) Type of increase in work for those reporting an increase

Notes: Panel (a) describes survey respondents' stated priorities over different program attributes when choosing a preschool. Ratings are on a 1-5 scale, with one being "not important" and five being "very important." Panel (b) describes survey reports of subjective treatment effects. Applicants were asked whether they believed receiving a UPK slot did (if enrolled) or would have (if not enrolled) resulted in "you or other adults in your household being able to work more," "less stress about money," and "better pre-k education for [the child]." Panel (c) reports the type of increase in work for survey respondents reporting that enrolling in a UPK program either did (if enrolled) or would have (if not enrolled) allowed them to work more. Responses are reported both for the survey respondent and for other household members (labeled with "Other HH Member"). Purple "No UPK" bars in each panel report data for the group that did not receive UPK and orange "UPK" bars for the group that did.

J Characteristics of UPK applicants

This section describes how families applying for New Haven's UPK program compare to the population of families with children in the New Haven area and to the samples used in previous studies of the labor market effects of childcare. Our focus is on levels and gender differences in rates of labor force participation.

We restrict our sample to applications submitted after 2013. As discussed in Section 3, post-2013 applications typically include records for multiple adults in each household and also list the relationship between each adult and the child. We further restrict our sample to applications for which we match all listed adults to the earnings sample.[24] We report labor market outcomes for parents in the two years preceding their child's pre-K entry, when children are typically 1 to 2 years old.

Figure J.1 reports earnings distributions for mothers and fathers in two-parent households, comparing UPK applicants (Panel A) to all parents of one- and two-year-olds in New Haven County (Panel B). UPK applicants tend to be lower earners but exhibit a relatively small gender wage gap. In contrast, parents of young children in New Haven County overall have higher earnings, with a much more pronounced gender gap.

UPK applicants have much higher rates of female labor force participation than the samples examined in prior studies. Table J.1 reports these comparisons. As reported in Panel A, 81% of single mothers and 79% of partnered mothers in our sample worked in the years leading up to pre-kindergarten application, for an overall participation rate of 80%. This is well above the labor force participation rates for women in previous studies. For example, Gelbach (2002) uses data from the 1980 Census in which labor force participation rates for single and married mothers are 62% and 51%, respectively, as we report in the columns labeled "1980 Census." Similarly, Gibbs et al. (2024) uses data from the 1999-2022 CPS ASEC in which the female labor force participation rate is 67%. Other studies of childcare programs, such as Lefebvre and Merrigan (2008), also focused on mothers with significantly lower labor force participation.

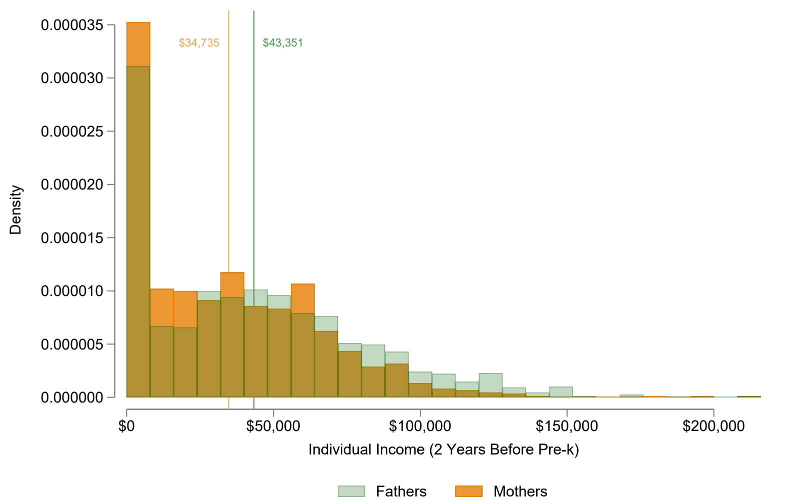
High rates of female labor force participation translate to relatively small gender gaps for UPK applicants. In the UPK applicant data, the labor force participation rate for fathers is 81%, essentially the same as the rate for mothers. This contrasts with the 19 percentage point gap reported in Gibbs et al. (2024). Similarly, mean earnings for fathers are about \$42,800 among UPK applicants, 59% above the \$26,900 value for mothers. This is far smaller than the 157% gap reported in Gibbs et al. (2024).

These facts provide insight into the effects of UPK enrollment on fathers discussed in Section 4.6.2. A stylized story is that extended-day UPK programs are particularly

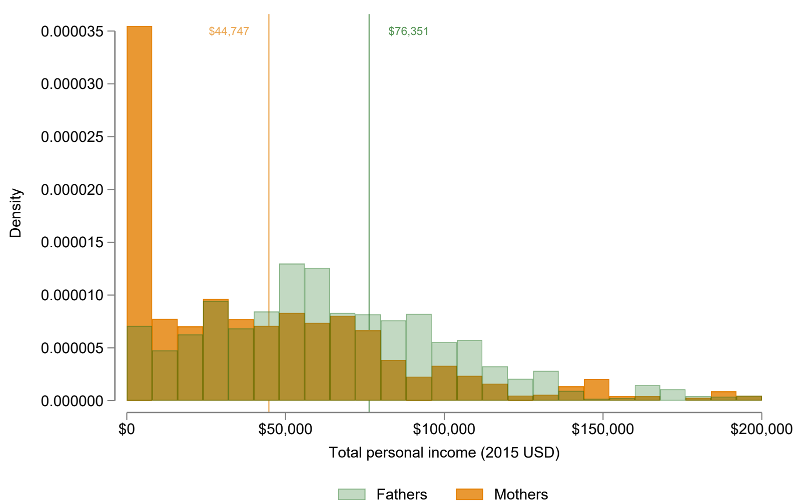
[24] Given we have two decades of earnings records, we believe individuals who do not match to earnings records are likely failed matches rather than participants who never work.

appealing to families where both parents have strong labor market attachment and are also both involved in childcare. Access to extended day care therefore relaxes time constraints and improves labor market outcomes for both mothers and fathers. While likely similar to other contemporary urban UPK programs, which share key features, our study population contrasts with the populations studied in past evaluations of the effects of childcare on parent labor supply, where mothers were much less likely to work than fathers and presumably bore a much greater share of the childcare responsibility.

Figure J.1: Earnings Distribution (2-parent families with 1-2 year olds)



(a) NHPS Application sample (2-parent families)



(b) New Haven County (2-parent families, ACS)

Notes: This figure shows the distribution of earnings for parents of young children (aged 1-2). Panel A shows earnings for NHPS applicants who applied after 2013, listed two parents on the application, and who both matched to earnings records. Earnings are shown for 2 years before pre-k application. Panel B shows earnings for parents of 1-2 year olds from two-parent families in New Haven county from the 2019 ACS, topcoded at $200,000. All income is in 2015 dollars.

Table J.1: Parent Labor Force Behavior Compared to Other Studies

	NHPS Applicants			1980 Census		CPS ASEC
Panel A: Mothers	Single	Partnered	All	Single	Married	All
Work Indicator	0.81	0.79	0.80	0.62	0.51	0.67
Quarters Worked	2.92	2.94	2.93			
Weeks Worked				23.56	18.02	31.65
Income (Incl. 0s)	19,516.33	34,735.35	26,852.75	13,592.38	9,321.27	22,546.19
Observations	2098	1863	4048	17817	105297	326267
Panel B: Fathers	Single	Partnered	All			All
Work Indicator	0.75	0.82	0.81			0.88
Quarters Worked	2.79	3.02	3.01			
Weeks Worked						46.62
Income (Incl. 0s)	32,875.90	43,350.92	42,838.97			57,837.31
Observations	107	1863	2003			266083

Notes: This table presents summary statistics for parents who applied to the NHPS lottery alongside parents analyzed in Gibbs et al. (2024) and Gelbach (2002). The NHPS applicant sample is restricted to parents who applied in 2013 or later and for whom all listed guardians matched to the earnings sample. Statistics are shown for two years prior to pre-K. The 1980 Census columns present statistics for single and married mothers from Gelbach (2002), which reports data separately for mothers with a youngest child aged 5 and under 5. We report a weighted average based on sample counts from Gelbach (2002). The CPS ASEC column provides statistics from the 1999-2022 CPS ASEC as reported in Gibbs et al. (2024). "Work indicator" denotes having any positive income. "Weeks worked" refers to weeks worked in the previous year, while "Quarters worked" represents quarters worked. All income values are in 2015 dollars.

K Cost-benefit calculations

This appendix provides details on the construction of the cost-benefit calculations described in Section 5. We focus on the marginal value of public funds (MVPF), though we also consider benefit-cost ratios (BCR) and net social benefit calculations (NSB).

K.1 Framework for MVPF calculations

This section shows how different models of the market for pre-kindergarten services affect what one should include in the numerator of the MVPF, i.e., the willingness to pay. A conceptual challenge here, and common in MVPF analyses in general, is that the MVPF framework focuses on marginal changes in policy and relies heavily on the logic of the envelope theorem, while the policy effects we measure (access to UPK) reflect large interventions in the lives of specific families. We abstract from this issue by considering marginal shifts in two types of UPK-like policies: price subsidies for childcare and expansions in childcare hours.

K.1.1 UPK as a price subsidy in an unconstrained market

We first show that if one models UPK as a price subsidy in an otherwise unconstrained market for childcare, WTP a) does not include parents own earnings and b) includes children's earnings if parents do not take children's outcomes fully into consideration when making labor supply choices.

Parents live for periods $t = 0$ through $t = T^p$. Period zero is pre-kindergarten, which differs from other periods because parents must pay for childcare while they work. Assume that in period zero there are perfect markets for childcare at price p and the government can offer some subsidy s. Parents' individual utility is given by $U_p = \sum_{t=0}^{T^p} \beta^t u(c_t, l_t)$. Parents may also take into account child utility U_k when they make labor supply choices, so we think of parents as maximizing $\tilde{U}_p = U_p + \theta U_k$, where $\theta \in [0, 1]$ is the weight parents place on child utility, subject to the budget constraint

$$\sum_{t=0}^{T^p} \beta^t c_t + l_0(p - s) \leq (1 - \tau) \left(w_0 l_0 + \sum_{t=1}^{T^p} \beta^t w_t(l_0) l_t \right) = (1 - \tau) Y^p. \tag{4}$$

w_t are wages in period t, τ is the tax rate on income, and Y^p denotes the PDV of pre-tax parent earnings. Wages in periods $t > 0$ can depend on labor supply l_0 in period 0. This captures the idea that career investments during pre-kindergarten may have long-run effects.

Children reach adulthood in period $\underline{T}^k$ and live through period $\overline{T}^k$. They have income $Y_t^k(l_0)$ in each period $t \in [\underline{T}^k, \overline{T}^k]$. We allow children's earnings to depend on parents' labor supply/childcare choices in period 0 to reflect the possible impacts of childcare use on children's human capital. Children maximize utility $U_k = \sum_{t=\underline{T}^k}^{\overline{T}^k} \beta^t u(c_t^k)$ subject to the budget constraint

$$\sum_{t=\underline{T}^k}^{\overline{T}^k} \beta^t c_t^k \leq (1 - \tau) \sum_{t=\underline{T}^k}^{\overline{T}^k} \beta^t Y_t^k(l_0) = (1 - \tau) Y^k. \tag{5}$$

Y^k denotes the PDV of pre-tax child earnings. An assumption built into this model is that neither parents nor children can borrow against children's future earnings to fund childcare expenses.

Net government expenditures are given by

$$G = l_0 s - \tau \left(Y^p + Y^k \right) \tag{6}$$

Let λ_p denote the multiplier on the budget constraint in the parent's optimization problem and λ_k denote the multiplier on the constraint in the children's problem. We can then write the MVPF of the policy as the sum of parent and child WTP divided by total government costs:

$$MVPF_1 = \frac{\frac{1}{\lambda_p} \frac{dU_p}{ds} + \frac{1}{\lambda_k} \frac{dU_k}{ds}}{\frac{dG}{ds}} = \frac{\frac{1}{\lambda_p} \frac{d\tilde{U}_p}{ds} + \left(\frac{1}{\lambda_k} - \frac{\theta}{\lambda_p} \right) \frac{dU_k}{ds}}{\frac{dG}{ds}}. \tag{7}$$

By the envelope condition, we know that $\frac{d\tilde{U}_p}{ds} = \lambda_p l_0$. The envelope condition does not apply to $\frac{dU_k}{ds}$ since children do not choose l_0; we therefore have $\frac{dU_k}{ds} = \lambda_k(1 - \tau)\frac{dY^k}{dl_0}\frac{dl_0}{ds}$. Finally, we have $\frac{dG}{ds} = l_0 + s\frac{dl_0}{ds} - \tau\left(\frac{dY^p}{ds} + \frac{dY^k}{ds}\right)$. We can then rewrite the MVPF using these expressions as

$$MVPF_1 = \frac{l_0 + (1-\tau)\frac{dY^k}{dl_0}\frac{dl_0}{ds}\left(1 - \theta\frac{\lambda_k}{\lambda_p}\right)}{l_0 + s\frac{dl_0}{ds} - \tau\left(\frac{dY^p}{ds} + \frac{dY^k}{ds}\right)}. \tag{8}$$

This expression provides guidance about how to compute the MVPF in this model. The denominator reflects to the total cost of the subsidy to government, net of the fiscal externality from additional earnings for children and parents. The numerator consists of two terms. The first is l_0, which corresponds to the cost of the subsidy absent labor supply responses. The second term, $(1-\tau)\frac{dY^k}{dl_0}\frac{dl_0}{ds}\left(1 - \theta\frac{\lambda_k}{\lambda_p}\right)$, describes how much of childrens' earnings should be included in the WTP. If $\theta = 0$ and parents do not weigh children's outcomes at all when making labor supply choices, we should include all of children's after-tax income. If $\theta > 0$, the picture is more complicated. What share of children's after-tax earnings we should include falls with θ (how much parents take children into account) and with λ_k/λ_p (the utility value of income for children relative to parents).

K.1.2 Hours constraints

A different way to think about the UPK subsidy is as a policy that relaxes a binding constraint on childcare hours. Evidence presented in the main text suggests that it may be difficult for people to find alternate programs that offer equal coverage to the UPK program. In this case the MVPF calculation is different.

Suppose individuals face the same problem as above but also face an additional constraint on labor supply based on the hours of childcare availability h^*, so that

$$l_0 \leq h^* \tag{9}$$

Clearly if the optimal value of l_0 in the baseline problem, l_0^*, is less than h^*, everything is the same. But if $l_0^* > h^*$, results change, because people would like to work more if they could.

Say the government is considering raising h^*, holding s fixed. Note that labor supply l_0 will rise 1-1 with h^* under the assumption that $l_0^* > h^*$. Then

$$\frac{dG}{dh^*} = s - \tau\left(\frac{dY^p}{dl_0} + \frac{dY^k}{dl_0}\right)$$

which is again simply the change in total costs less the fiscal externality from additional earnings for parents and children.

Turning to willingness to pay, we are now in a corner solution with respect to labor supply in period zero. The welfare gains for parents from increased work in period 0 are now positive even as the envelope condition continues to apply to the other optimized variables. In particular,

$$\frac{d\tilde{U_p}}{dl_0} = \lambda_p \left((1-\tau) \left(w_0 + \sum_{t=1}^{T^p} w_t'(l_0) l_t \right) - (p-s) \right) + \theta \frac{dU_k}{dl_0} + u_l(c_0, l_0) > 0 \tag{10}$$

i.e. it is the net amount of dollars the parent gets from labor supply this period and wage gains in the future scaled by the utility value of a dollar in income, plus whatever parent-valued benefits kids get from additional work, less the disutility of additional work.

Let $\frac{dE}{dl_0} = \left((1-\tau) \left(w_0 + \sum_{t=1}^{T^p} w_t'(l_0) l_t \right) - (p-s) \right)$ denote the utility-relevant change in parent earnings from an increase in l_0. Also note that $\frac{dU_k}{dl_0} = \lambda_k (1-\tau) \frac{dY^k}{dl_0}$. Then we can write an alternative MVPF formulation given by

$$MVPF_2 = \frac{\frac{dE}{dl_0} + \frac{u_l(c_0,l_0)}{\lambda_p} + (1-\tau)\frac{dY^k}{dl_0}\left(1 - \theta\frac{\lambda_k}{\lambda_p}\right)}{s - \tau\left(\frac{dY^p}{dl_0} + \frac{dY^k}{dl_0}\right)} \tag{11}$$

This MVPF formula includes the part of parents' earnings gains attributable to period-0 earnings gains and later wage gains in the numerator, less the utility cost of period zero labor supply. It also includes a term for children's earnings gains that parallels the one included in $MVPF_1$.

What we take away from this analysis is that if access to full-day childcare is constrained, it may make sense to include all or part of parents' earnings gains in the WTP term. An upper bound on the WTP for parents' future earnings gains is given by the effect of UPK on after tax parent earnings. This bound would be tight if 1) changes in earnings after pre-kindergarten are due to wages, not labor supply, and 2) there is no disutility from work, or equivalently, it's just as hard to take care of a kid as it is to work.

K.1.3 Misunderstanding the returns to career continuity

Another reason it might be reasonable to include a component of parent earnings in the numerator of the MVPF is if parents do not understand the dynamic returns to career continuity.

To explore this, return to the baseline model from Section K.1.1. Assume that when making choices about period 0 labor supply, parents believe that $w_t'(l_0) = 0$ for all $0 < t \leq T^p$. To simplify the analysis, assume that parents learn about the returns to experience after choosing l_0 but before choosing c_0, so that consumption is still perfectly smoothed given lifetime income.

The analysis then proceeds as in Section K.1.1, except that

$$\frac{d\tilde{U}_p}{ds} = \lambda_p \left(l_0 + (1-\tau) \sum_{t=1}^{T^p} \beta^t w_t'(l_0) \frac{dl_0}{ds} l_t \right).$$

The sum reflects the welfare gains from unanticipated wage effects. Denote this term $\frac{dW}{ds}$. Plugging into the MVPF formula, we have

$$MVPF_3 = \frac{l_0 + (1-\tau)\frac{dW}{ds} + (1-\tau)\frac{dY^k}{dl_0}\frac{dl_0}{ds}\left(1 - \theta\frac{\lambda_k}{\lambda_p}\right)}{l_0 + s\frac{dl_0}{ds} - \tau\left(\frac{dY^p}{ds} + \frac{dY^k}{ds}\right)}. \tag{12}$$

In this setup, we would want to include earnings gains from future wage changes in the WTP term.

K.1.4 Credit constraints

In the absence of constraints on the availability of full time childcare in the private market, credit constraints do not by themselves motivate the inclusion of parents' earnings in the numerator of the MVPF. They do, however, a) suggest that we may want to include a larger share of children's earnings in the MVPF, and b) motivate an alternate argument in favor of childcare subsidies not captured by the MVPF: namely, that the value of consumption is high for families of pre-kindergarten age children.

To see this, start with the baseline model from Section K.1.1. Add an additional constraint that rules out borrowing in period 0:

$$c_0 + l_0(p-s) \leq (1-\tau)w_0 l_0. \tag{13}$$

Let λ_{cc} denote the multiplier on this constraint. Now consider the MVPF of raising the subsidy s. Assuming that the credit constraint in period zero binds, the utility value of a \$1 cash transfer to parents in period 0 is $\lambda_p + \lambda_{cc}$. We may therefore write

$$MVPF_4 = \frac{\frac{1}{\lambda_p+\lambda_{cc}}\frac{dU_p}{ds} + \frac{1}{\lambda_k}\frac{dU_k}{ds}}{\frac{dG}{ds}} = \frac{\frac{1}{\lambda_p+\lambda_{cc}}\frac{d\tilde{U}_p}{ds} + \left(\frac{1}{\lambda_k} - \frac{\theta}{\lambda_p+\lambda_{cc}}\right)\frac{dU_k}{ds}}{\frac{dG}{ds}}. \tag{14}$$

Applying the envelope theorem, we obtain

$$MVPF_4 = \frac{l_0 + (1-\tau)\frac{dY^k}{dl_0}\frac{dl_0}{ds}\left(1 - \theta\frac{\lambda_k}{\lambda_p+\lambda_{cc}}\right)}{l_0 + s\frac{dl_0}{ds} - \tau\left(\frac{dY^p}{ds} + \frac{dY^k}{ds}\right)}. \tag{15}$$

As in our baseline case, parent earnings do not appear in the numerator. The one difference relative to $MVPF_1$ is that the discount applied to child earnings, $\left(1 - \theta \frac{\lambda_k}{\lambda_p + \lambda_{cc}}\right)$, will tend to be smaller for a given value of θ.

The second difference, of course, is that credit constraints increase the marginal utility of consumption for parents of young children. If one believes transfers should target groups with higher marginal utility, this strengthens the case for transfers aimed at UPK beneficiaries. This does not show up in the numerator of the MVPF formula because WTP scales marginal utility of the subsidy by the value of a distortion-free transfer, and these rise in proportion as credit constraints rise.

K.2 Inputs to MVPF calculation

Our cost-benefit calculations focus on four components: (1) the net change in per-pupil expenditure (PPE) before kindergarten, (2) the change in out-of-pocket childcare costs for families, (3) the discounted present value of the child's wage gains estimated from changes in kindergarten test scores, and (4) the discounted present value of increased parental wage income. We first discuss how we calculate each of these four components and then discuss how they are used in our various benefit calculations.

Change in public per-pupil expenditure

As an initial step in calculating the net change in PPE prior to kindergarten, we first calculate the causal impact of UPK enrollment on the number of years in the UPK program, the number of years in Head Start, the number of years in School Readiness programs, the number of years with Care 4 Kids subsidies, and the number of years in other public or subsidized pre-k programs as recorded in administrative data from the State Department of Education and Office of Early Childhood. For enrollment in Head Start, we focus on 2015-2017 as those are the years for which we have the most complete data. As show in Table K.1, the substitution patterns in those years are very similar to those for the full sample. We estimate these values using our standard 2SLS specification as in Table 3 but taking years of enrollment rather than enrollment indicators as dependent variables. We consider only enrollment at ages 3 and 4; we do not include enrollment at older or younger ages. We also calculate these values by tercile of median neighborhood income, using the neighborhood when applying to the UPK program.

Next, we use PPE estimates for the various childcare and pre-kindergarten options from the National Institute for Early Education Research (NIEER) (Friedman-Krauss et al., 2022, 2023) and New Haven Public Schools (Connecticut Office of Elementary and Secondary Education, 2020). For childcare and pre-kindergarten options, we use

PPE estimates for School Readiness, Care 4 Kids, and Head Start in Connecticut in 2021-2022 (2018-2019 for Head Start), all in real 2015 dollars. For all public programs in the State Department of Education and Office of Early Childhood data that we are not able to classify, we use the average PPE for state programs in Connecticut.

Combining our 2SLS estimates and our PPE estimates, we calculate the gross and net public program costs of UPK enrollment. Gross program costs are the per-year PPE of the UPK program (PPE_m) multiplied by the 2SLS estimate of the increase in years of UPK enrollment (Δ_{UPK}). Net program costs then use the PPE for other programs and the change in years of enrollment in those other programs giving us:

$$\begin{aligned}\text{Net Program Costs} = \Delta_{\text{UPK}} \cdot PPE_m + \Delta_{\text{headstart}} \cdot PPE_{hs} + \Delta_{\text{schoolreadiness}} \cdot PPE_{sr} \\ + \Delta_{\text{care4kids}} \cdot PPE_{c4k} + \Delta_{\text{otherpublic}} \cdot PPE_{op}.\end{aligned}$$

We estimate similar regressions within terciles of neighborhood median household income. For these, we use the same PPE, but estimate the 2SLS estimates of years enrolled in the various programs conditional on each tercile at the time of application.

Reduction in out-of-pocket costs

We estimate the reduction in parents' out-of-pocket (OOP) costs of childcare using the 2SLS estimates reported in Tables 3 and 6. We assume that the estimated monthly reduction applies to the nine-month school year in each year the child enrolls in UPK. We estimate the effect of enrolling in UPK on total years of UPK enrollment using IV specifications with the count of years enrolled as the outcome. For years of enrollment beyond one, we discount the value using an interest rate of 0.03. We estimate these values in the full sample and within terciles of neighborhood median household income.

Projected earnings gains for children

We estimate the impacts of UPK enrollment on children's kindergarten test scores, then use these impacts to project future earnings. To do this, we first estimate the impact of UPK on kids' kindergarten test scores (measured in standard deviations) using our 2SLS approach. Next, we use our estimates to predict future earnings, closely following Cascio (2023), which also estimates MVPFs of UPK. First, Cascio (2023) assumes that the average present discounted value of earnings at age 4 is \$291,287 in 2005 dollars, which is the age 10 estimate from Chetty et al. (2011) of \$522,000 in 2010 dollars discounted back to age 4 using a 3% discount rate, which is also the discount rate we use in our estimates. Adjusting to 2015 dollars, we have an estimate of the average present discounted value of earnings at age 4 of \$353,507. Next, following Cascio (2023) and Kline and Walters (2016), we make the assumption that a one standard deviation increase in test scores increases earnings by 10%. We can then calculate the discounted present value of earnings gains for children as the product of the IV estimate, 0.06, and $353,507. We produce similar estimates by tercile of neighborhood household income, using 2SLS estimates calculated by tercile. Finally, our $353,507 estimate is the average present discounted value of earnings. Kline and Walters (2016), who study children eligible for Head Start, multiply this number by 0.8 to account for the fact that these children are in lower-income families. When producing results by income tercile, we assume 0.8 for the bottom tercile, 1 for the middle tercile, and 1.2 for the top tercile.

While the approach above follows the prior literature closely, it involves many assumptions. We therefore consider two alternative approaches. First, we consider larger test score gains of 0.4, as found in Lipsey et al. (2018). Second, we calculate earnings gains for kids based on the increase in 4-year college enrollment after high school reported by Gray-Lobe et al. (2023), who study a similar UPK program in Boston (a 0.086 increase in the probability of enrollment). Then we use estimates from Zimmerman (2014) to estimate the returns to enrolling in college ($142,757 in 2012 dollars, based on calculations from Hendren and Sprung-Keyser (2020)). We additionally adjust net costs by 0.086 * 2,617, where the latter number is the estimated net cost to the government of an additional enrollee in 4-year college.

Discounted present value of increased parental wage income

Tables 4 and 6 report the 2SLS estimates of enrolling in the UPK program on earnings in (1) the years when the student is enrolled in the program, (2) one to two years after, (3) three to four years after, and (4) five to six years after. Using these estimates, we construct the discounted present value of wage income gains using an interest rate of 0.03. We assume the effect is constant in each of the four time periods defined above, and then zero afterwards. The assumption of zero gains beyond six years out is based on Figure 4, which shows large and persistent gains through six years after prekindergarten, after which confidence intervals grow large and we cannot rule out null effects. We scale individual earnings effects by 1.56, the average number of adults listed on applications during years when we systematically observe multiple family members. While parents may not have listed all adults when completing the application, we believe 1.56 captures nearly all parents. The figure closely aligns with the average number of adults in households with children under age 6 in New Haven County, which is 1.63 [25].

The gains to parents will differ if the child enrolls in one or two years of UPK. To address this we estimate the discounted present value of parents' wage income both under the assumption of 1 and of 2 years of UPK. We then weight these estimates

[25] Source: Authors' calculation from the ACS 5-year 2019 estimates

using the estimated increase in years of UPK enrollment from the 2SLS estimates. For example, in the full sample we estimate that enrolling in UPK results in 1.56 additional years of UPK enrollment. This reflects a mixture of individuals enrolling for one year and individuals enrolling for two years, so 56 percent of applicants enroll for two years.

K.3 MVPF Calculations

Using the four inputs described above, we calculate the MVPF for the program as a whole and by tercile of neighborhood median household income. Below, we describe the construction of willingness to pay and net costs.

Because UPK is an in-kind transfer, what enters willingness to pay depends on our underlying economic assumptions, as described in Online Appendix K.1. We consider four potential constructions of willingness to pay. One uses a cost-based approach, and three use hedonic approaches, making different assumptions about the welfare value of program benefits.

1. Our first approach is based on program costs and assumes that families value UPK at the additional childcare subsidy they receive (i.e., the government's cost of providing UPK net of savings from substituting away from other subsidized programs). Estimates using this approach are shown in column one of Tables 8 and K.4 and row one of Table K.5.

2. Our first hedonic approach constructs willingness to pay from estimates of the reduction in out of pocket costs paid by families and the future income gains for children. This specification assumes that the entire effect on parental earnings is driven by behavioral distortions due to changes in childcare pricing, and as such, these earnings effects are excluded from willingness to pay. Estimates are shown in column two of Tables 8 and K.4 and row two of Table K.5.

3. Our second hedonic approach additionally includes parent earnings after pre-k. This would make sense if, for example, parents didn't internalize the later earnings gains associated with increased hours of work and labor force participation during pre-k. These estimates are shown in column three of Tables 8 and K.4 and row three of Table K.5.

4. Our third hedonic approach additionally includes earnings effects during pre-k. As shown in Online Appendix K.1, this is consistent with a model where the UPK program relaxes constraints on the hours of childcare families can access. Estimates are shown in column four of Tables 8 and K.4 and row four of Table K.5. Our baseline version of this specification does not discount earnings for the potential disutility of work, particularly during pre-k, where our survey data shows a notable increase in hours. For robustness, we also consider a specification where

we assume the utility cost of work during pre-k is equal to 60% of pay, based on estimates from Mas and Pallais (2019), which is estimated for people on the margin of employment vs unemployment. Results are shown in Table K.4.

For children's wage gains, we use the discounted present value described above net of taxes. We assume an effective tax rate of 0.2. Out of pocket costs come directly from the reduction in out-of-pocket costs described above.

For parents' wage gains, we calculate total wage income from the quarterly earnings records from the CT Department of Labor. Following Cascio (2023), we assume an effective marginal tax rate of 20%. This follows closely from Hendren and Sprung-Keyser (2020) who use effective marginal tax rates from Congressional Budget Office (2016), which are approximately 20% for incomes from 100 to 400 percent of the federal poverty line.

To calculate net costs we use the change in per-pupil public expenditure discussed above, which accounts for substitution from other subsidized programs. We then additionally subtract the discounted present value of tax revenue increases associated with changes in wage income for the parents and children using the tax rates discussed above. The estimates above assume that the other publicly-funded pre-k programs students substitute away from are not rationed. As discussed in Kline and Walters (2016), if the programs children substitute from are also over-subscribed, the MVPF calculation would use the benefits to the child (and their family) who takes up the now vacated slot in the other program (and the costs of what they substitute away from). As we cannot estimate the returns to the other programs, this is beyond the scope of our paper. As a robustness check, we include estimates that assume there are no cost savings from substituting from other programs, and also no benefits for children who gain access to the vacated slots. We view the resulting MVPF estimates as a loose lower bound, since they incorporate the cost side of children's substitution into vacated slots, but not the benefits.

K.4 Benefit Cost Ratio and Net Social Benefit Calculations

An alternative approach to constructing the benefits of a program is the benefit cost ratio (BCR, García et al., 2020; García and Heckman, 2022). This largely uses the same inputs as the MVPF. The numerator includes the numerator from (1) the MVPF (willingness to pay) plus (2) the indirect cost saving (savings from substitution from other publicly-funded pre-k and childcare programs along with tax revenue from increased earnings of parents and children), where (2) is then multiplied by one plus the dead-weight loss associated with raising taxes to fund the program, which we assume to be 0.3. The denominator is the direct cost of the program, calculated as the yearly PPE of the magnet program multiplied by the increase in the number of years of magnet enrollment, multiplied by one plus the dead-weight loss term.

Lastly, we estimate the net social benefit (NSB, García and Heckman, 2022) of the program. This is the willingness to pay from the MVPF calculation minus the net cost term from the MVPF calculation multiplied by one plus the dead-weight loss term. For both of these calculations, we also consider the four potential constructions of willingness to pay described in the MVPF calculations section.

K.5 Additional cost-benefit analysis results

Table K.2 reports benefit-cost ratios as described in Section K.4. The first column reports the BCR for the full population, while the second through fourth columns report the BCR by tercile of neighborhood median household income based on the neighborhood of residence when applying for the UPK program. Each row represents a different construction of the willingness to pay. The overall BCR ranges from 1.16 (in our most conservative hedonic approach to constructing willigness to pay) to 2.86 (in our hedonic approach that includes all parent earnings gains. Similar to our MVPF calculations, the BCR is larger for the second and third terciles of neighborhood income (at the time of the application).

Table K.3 reports the net social benefit as described in Online Appendix K.4. Similar to the prior table, the first column reports the NSB for the full population, while the second through fourth columns report the NSB by tercile of neighborhood median household income based on the neighborhood of residence when applying for the UPK program. Each row represents a different construction of the willingness to pay. The overall NSB ranges from $4,900 to $58,600, depending on what we include in willingness to pay. The NSB is larger for the second and third terciles of neighborhood income.

Table K.4 reports estimates of the MVPF under four different constructions of willingness to pay:

- No program substitution: Excludes the public cost-saving of individuals substituting away from other publicly funded pre-kindergarten and childcare programs.
- Survey-based program substitution: Uses substitution patterns estimated from survey data rather than state administrative data.
- No out of pocket costs: Excludes the savings in out-of-pocket costs to families.
- 60% opportunity cost of work: Assumes that the WTP for parents are only 40% of their gains in wage income during the pre-k years since during pre-k years much of the wage gains come from increased hours.
- 10% higher tax rates: Scales up all taxes by 10% as a robustness test.

- 25% higher tax rates: Scales up all taxes by 25% as a robustness test.
- 10% lower tax rates: Scales down all taxes by 10% as a robustness test.
- 25% lower tax rates: Scales down all taxes by 25% as a robustness test.
- 25% higher UPK cost: Scales up UPK cost by 25% as a robustness test.
- 25% lower UPK cost: Scales down UPK cost by 25% as a robustness test.
- Alternative PPE Estimate: Uses New Haven School District budget data to estimate pre-k specific per-pupil expenditures of $12,591 based on the method used by Kabay et al. (2020).[26]
- Smaller family size: Uses a smaller family-scaling number of 1.31 for the full sample based on the average number of parents per household in the sample that matched to the earnings data.
- Downstream sibling enrollment: We estimate that winning the lottery leads to an additional 0.18 years in UPK enrollment for a child's siblings and an additional 0.1 siblings enrolled.[27] We adjust the net government costs and changes in out-of-pocket expenses by the additional years of enrollment, and scale the children's earnings effects by the increase in the number of siblings enrolled.
- Partial substitution: Some children are enrolled in multiple programs within a given year. For magnet students, we assume this additional enrollment occurs during the summer months. For non-magnet students, we assume an equal split, with half-time enrollment in each program. We then adjust government costs accordingly: for magnet students, the government funds a full year of UPK plus a quarter year of other subsidized programs, while for non-magnet students enrolled in multiple programs, it covers a half year of each subsidized program.
- Kid effects of 0.4σ, no parent earnings gains: We assume test score gains of 0.4 standard deviations for kids following Lipsey et al. (2018), and do not include parents' earnings gains in the calculations.
- GPW kids' effects, no parent earnings gains: We assume kids benefit through increased college enrollment as reported in Gray-Lobe et al. (2023), and do not include parents' earnings gains in the calculations.

For each estimate we report 90% confidence intervals in brackets, which are based on 500 bootstrap samples. Following Hendren and Sprung-Keyser (2020), when net costs are negative and willingness to pay is positive, we report the MVPF to be infinity.

[26] Estimates used the NHPS 2021-2022 budget (New Haven Board of Education, 2022), average CT teacher salaries (Teach Connecticut, 2021) and average CT paraprofessional salaries.

[27] As discussed in Online Appendix B.1, siblings of children already enrolled in magnet schools receive priority in UPK admissions.

K.6 Cost-benefit analysis by income

A natural question to ask about untargeted programs like UPK is whether it might be possible to obtain similar benefits with more desirable distributional properties through a means test. We repeat our MVPF analysis from Section 5, re-estimating the parameters used in the MVPF calculations within neighborhood income tercile. For our main estimates, we project children's earnings based on tercile-specific gains in kindergarten scores from Table 6.

Figure K.1 reports inputs into the MVPF by tercile of neighborhood income while Table K.5 reports estimates of MVPF for the full population and by tercile of neighborhood income. The rows represent the four different constructions of willingness to pay. 90% confidence intervals are reported in brackets, which are based on 500 bootstrap samples.

The central result here is that MVPFs are infinite (i.e., net costs of UPK are negative) in the top two terciles because that is where earnings gains are largest. Panel (b) of Figure K.1 repeats the net government cost analysis from Figure 5 but split by tercile of neighborhood income. Upfront costs are similar across the three groups. The small differences we do see are driven by differences in the number of years enrolling students in each group attend a UPK program. Savings from government costs are largest for the lowest tercile (roughly \$13,000, compared to \$8,000-\$9,000 in the upper terciles) because this group is more likely to substitute away from other subsidized programs. Additional tax revenue from children is fairly small relative to costs across all three groups. The additional tax revenue from parents is much larger in terciles two and three (\$18,000 and \$20,000, respectively) than in the bottom tercile (\$3,000). Together with the program substitution effects, the additional tax revenue from parents is large enough to push net costs negative in the top two terciles.

In short, the bulk of the return to UPK expenditure comes from middle- and higher-income families within the applicant pool. Recall that given the income distribution within our sample, even the upper tercile of the distribution consists mainly of middle-class people who are often the target of work support programs. For example, for a single household head with one child, the maximum AGI value for EITC eligibility in 2022 was just under \$40,000 (in 2015 dollars), similar to the mean individual income in our top tercile.

Table K.1: Head Start Substitution 2015-2017

	Full sample	ACS 1st tercile	ACS 2nd tercile	ACS 3rd tercile
Enrolled Head Start	-0.176	-0.332	-0.190	-0.060
	(0.061)	(0.061)	(0.037)	(0.038)
N individuals	3902	1402	1281	1259
N applications	4385	1558	1397	1400

Notes: This table shows the effect of UPK enrollment on substitution out of a Head Start program, analogous to the results presented in Tables 3 and 6. Here the sample is restriction to individuals who applied between 2015 and 2017 because these are the application years for which we have the most comprehensive data on Head Start Enrollment. Standard errors are clustered at the applicant level.

Table K.2: BCR estimates by income tercile

Specification	Full	Bottom	Middle	Top
WTP is net direct govt cost	1.43	0.96	1.70	1.68
	[1.14, 1.74]	[0.58, 1.31]	[1.27, 2.20]	[1.05, 2.30]
WTP is OOP costs, kid earnings	1.16	0.47	1.64	1.45
	[0.83, 1.50]	[-0.06, 0.86]	[1.06, 2.24]	[0.78, 2.12]
WTP is OOP cost, kid earnings, parent post earnings	2.53	0.63	3.65	3.43
	[1.50, 3.68]	[-0.77, 1.88]	[2.06, 5.51]	[1.07, 5.73]
WTP OOP cost, kid earnings, all parent earnings	2.86	0.85	4.01	3.99
	[1.68, 4.11]	[-0.74, 2.27]	[2.29, 6.00]	[1.54, 6.56]

Notes: This table reports estimates of the benefit-cost ratio (BCR) by tercile of ACS median block-group household income using four different constructions of willingness to pay (WTP). The first row assumes parents value the program at its net direct government cost. The second excludes parental earnings from WTP and considers only change in out of pocket expenditures and kids' future earnings. The third adds post-pre-k parental earnings, and the fourth considers all parental earnings. The "Full" column reports estimates for the whole sample while the remaining three columns report estimates by tercile of neighborhood median household income, which are based on the block group the family lived in at the time they applied. 90% confidence intervals are reported in brackets based on 500 bootstraps.

Table K.3: NSB estimates by income tercile

Specification	Full	Bottom	Middle	Top
WTP is net direct govt cost	13,401	-1,256	21,337	21,744
	[4,668; 22,857]	[-14,731; 10,744]	[8,411; 36,819]	[1,624; 42,231]
WTP is OOP costs, kid earnings	4,955	-17,959	19,474	14,290
	[-5,309; 15,681]	[-36,895; -4,502]	[1,712; 37,440]	[-6,876; 36,177]
WTP is OOP cost, kid earnings, parent post earnings	48,110	-12,795	80,166	77,636
	[16,050; 83,314]	[-61,996; 30,288]	[32,242; 139,537]	[2,232; 152,450]
WTP OOP cost, kid earnings, all parent earnings	58,663	-5,245	91,140	95,379
	[21,784; 95,955]	[-60,644; 42,994]	[39,054; 155,083]	[16,904; 179,019]

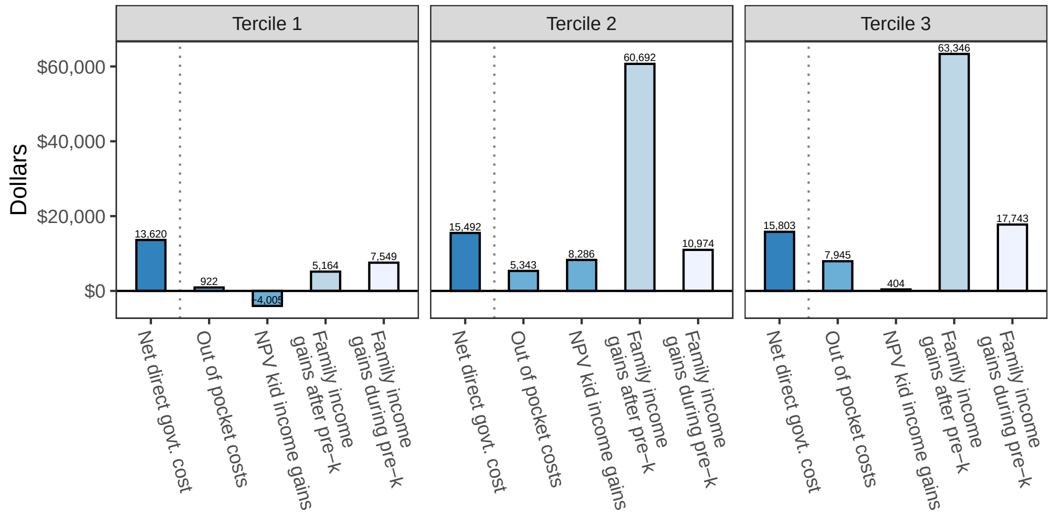
Notes: This table reports estimates of the Net Social Benefit (NSB) by tercile of ACS median block-group household income using four different constructions of willingness to pay (WTP). The first row assumes parents value the program at its net direct government cost. The second excludes parental earnings from WTP and considers only change in out of pocket expenditures and kids' future earnings. The third adds post-pre-k parental earnings, and the fourth considers all parental earnings. The "Full" column reports estimates for the whole sample while the remaining three columns report estimates by tercile of neighborhood median household income, which are based on the block group the family lived in at the time they applied. 90% confidence intervals are reported in brackets based on 500 bootstraps.

Table K.4: Sensitivity of MVPF calculations to underlying assumptions

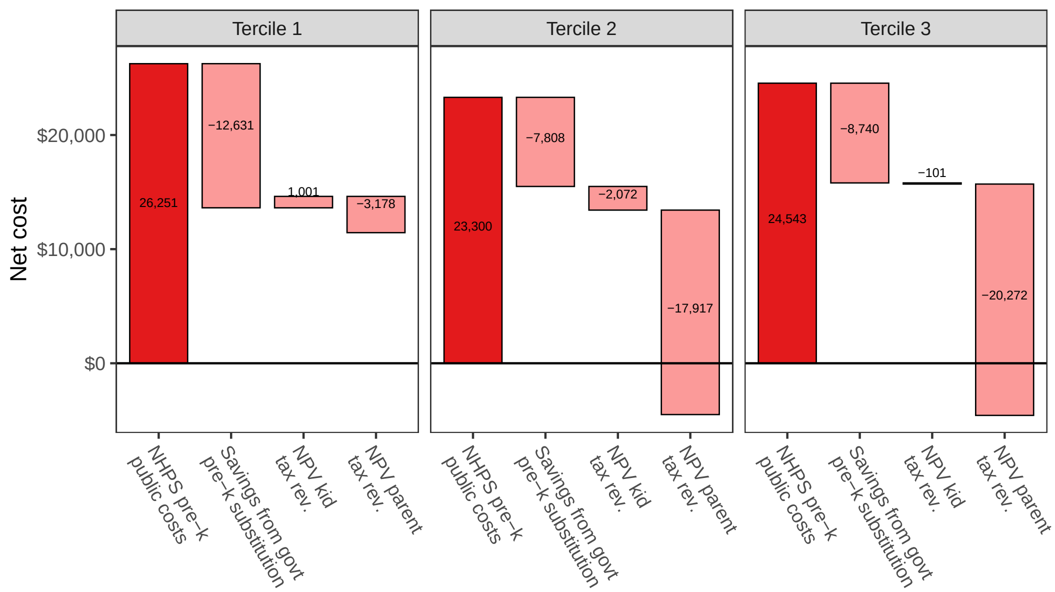
Specification	WTP is net direct govt cost	WTP is OOP costs, kid earnings	WTP is OOP costs, kid earnings, parent post earnings	WTP is OOP costs, kid earnings, all parent earnings
Robustness				
No program substitution	2.34	0.67	4.83	5.85
	[1.41, 8.46]	[0.23, 2.52]	[1.50, 24.63]	[1.89, 30.31]
Survey-based program subs.	3.47	1.24	8.94	10.83
	[1.58, Inf]	[0.34, Inf]	[2.09, Inf]	[2.72, Inf]
No out of pocket costs	10.10	1.13	29.46	36.39
	[1.87, Inf]	[-0.42, Inf]	[2.53, Inf]	[3.47, Inf]
60% opportunity cost of work	10.10	4.55	32.88	35.66
	[1.87, Inf]	[0.58, Inf]	[3.16, Inf]	[3.52, Inf]
10% higher tax rate	111.95	50.16	356.40	431.29
	[2.04, Inf]	[0.72, Inf]	[3.39, Inf]	[4.33, Inf]
25% higher tax rate	Inf	Inf	Inf	Inf
	[2.38, Inf]	[0.85, Inf]	[3.84, Inf]	[4.86, Inf]
10% lower tax rate	5.29	2.40	17.60	21.32
	[1.72, Inf]	[0.52, Inf]	[2.96, Inf]	[3.78, Inf]
25% lower tax rate	3.08	1.41	10.60	12.85
	[1.53, Inf]	[0.41, Inf]	[2.72, Inf]	[3.48, Inf]
25% higher UPK cost	2.83	0.91	6.61	8.00
	[1.49, Inf]	[0.28, Inf]	[1.87, Inf]	[2.25, Inf]
25% lower UPK cost	Inf	Inf	Inf	Inf
	[3.70, Inf]	[2.51, Inf]	[9.87, Inf]	[12.63, Inf]
Alternative PPE calculation	Inf	Inf	Inf	Inf
	[6.07, Inf]	[4.77, Inf]	[18.16, Inf]	[23.53, Inf]
Smaller family size	4.19	1.89	11.77	14.19
	[1.64, Inf]	[0.49, Inf]	[2.44, Inf]	[3.02, Inf]
Downstream sibling enrollment	4.24	1.88	11.93	14.39
	[1.65, Inf]	[0.50, Inf]	[2.46, Inf]	[3.07, Inf]
Partial substitution	9.01	4.01	28.94	35.03
	[1.83, Inf]	[0.55, Inf]	[3.20, Inf]	[3.84, Inf]
Additional Decompositions				
Kid effects of 0.4σ,	1.23	1.32	1.32	1.32
no parent earnings gains	[1.20, 1.27]	[1.09, 1.62]	[1.09, 1.62]	[1.09, 1.62]
GPW kids' effects,	1.18	1.15	1.15	1.15
no parent earnings gains	[1.15, 1.21]	[0.95, 1.42]	[0.95, 1.42]	[0.95, 1.42]

Notes: This table reports estimates of the MVPF under several alternative assumptions. The columns report our four assumptions regarding what enters willingness to pay. The first column assumes parents value the program at its net direct government cost. The second excludes parental earnings from WTP and considers only change in out of pocket expenditures and kids' future earnings. The third adds post-pre-k parental earnings, and the fourth considers all parental earnings. "No program substitution" excludes savings from reduced use of other public programs from the net cost calculation. "Survey-based program subs." uses the substitution patterns away from other public programs based on survey responses rather than the administrative data. "No out of pocket costs" excludes reductions in out of pocket costs from the willingness to pay. "60% opportunity cost of work" aims to account for the disutility of work. We assume this applies only while the child is in pre-k as we find the strongest evidence of labor supply effects during pre-k. We discount after-tax earnings in the pre-k period by 60% based on estimates from Mas and Pallais (2019). The next four rows scale the assumed taxes 10 and 25 percent higher or lower respectively. 25% higher/lower UPK cost scales up/down the per pupil expenditure. "Alternate PPE calculation" is based on New Haven School District budget data. "Smaller family size" uses a smaller family-scaling number based on the sample that matched to earnings data. "Downstream sibling enrollment" incorporates the effect of winning the lottery on sibling enrollment in UPK. "Partial substitution" assumes part time enrollment for children enrolled in multiple programs in a given year. "Kid effects of 0.4σ, no parent earnings gains" assumes test score gains of 0.4 standard deviations for kids following Lipsey et al. (2018), without including parents' earnings gains in the calculations. "GPW kids' effects, no parent earnings gains" assumes kids benefit through increased college enrollment as reported in Gray-Lobe et al. (2023), without including parents' earnings gains in the calculations. 90% confidence intervals are reported in brackets based on 500 bootstraps.

Figure K.1: Inputs into MVPF by neighborhood income tercile



(a) Inputs into WTP



(b) Inputs into net cost

Notes: This figure reports the inputs into the numerator (Panel (a)) and denominator (Panel (b)) of MVPF by tercile of ACS median block-group household income, based on the residential address of those who applied for the UPK program. Panel A reports five potential inputs into willingness to pay (WTP). (1) Net direct government cost, (2) reduction in out-of-pocket costs, (3) child income gains estimated based on changes in kindergarten test scores, (4) parental income gains after pre-kindergarten, (5) parental income gains during pre-kindergarten. Panel B reports inputs into net cost which are added sequentially from left to right. The dark red bar on the left of each panel shows the direct public costs of providing the UPK slot. The second bar from the left shows the public savings from substitution away from other publicly funded pre-k and childcare programs. The third bar from the left shows the changes in discounted tax revenue due to the estimated changes in kids' wage income. The fourth bar shows the changes in discounted tax revenue due to changes in parents' wage income. See Section 4.6 for details.

Table K.5: MVPF estimates by income tercile

Specification	Full	Bottom	Middle	Top
WTP is net direct govt cost	10.10	1.19	Inf	Inf
	[1.87, Inf]	[0.61, 5.82]	[2.51, Inf]	[1.45, Inf]
WTP is OOP costs, kid earnings	4.55	-0.27	Inf	Inf
	[0.58, Inf]	[-1.80, 0.46]	[1.81, Inf]	[0.39, Inf]
WTP is OOP cost, kid earnings, parent post earnings	32.88	0.18	Inf	Inf
	[3.16, Inf]	[-1.72, 15.12]	[5.54, Inf]	[1.52, Inf]
WTP OOP cost, kid earnings, all parent earnings	39.81	0.84	Inf	Inf
	[4.03, Inf]	[-1.50, 20.07]	[7.16, Inf]	[2.76, Inf]

Notes: This table reports estimates of the MVPF by tercile of ACS median block-group household income using four different constructions of willingness to pay (WTP). The first row assumes parents value the program at its net direct government cost. The second excludes parental earnings from WTP and considers only change in out of pocket expenditures and kids' future earnings. The third adds post-pre-k parental earnings, and the fourth considers all parental earnings. The "Full" column reports estimates for the whole sample while the remaining three columns report estimates by tercile of neighborhood median household income, which are based on the block group the family lived in at the time they applied. 90% confidence intervals are reported in brackets based on 500 bootstraps.

References

Agarwal, N. and P. Somaini (2018). Demand Analysis using Strategic Reports: An Application to a School Choice Mechanism. Econometrica 86(2), 391–444.

Azevedo, E. and J. Leshno (2016). A Supply and Demand Framework for Two-Sided Matching Markets. Journal of Political Economy 124(5), 1235–1268.

Care 4 Kids (2024). Care 4 Kids Program Income Guidelines. Available at https://www.ctcare4kids.com/care-4-kids-program/income-guidelines/. Accessed: 2024-03-13.

Cascio, E. U. (2023). Does Universal Preschool Hit the Target?: Program Access and Preschool Impacts. Journal of Human Resources 58(1), 1–42.

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.

Congressional Budget Office (2016). Effective Marginal Tax Rates for Low- and Moderate-Income Workers in 2016. https://www.cbo.gov/publication/50923. Accessed: 1/23/2024.

Connecticut Office of Early Childhood (2024). School Readiness Preschool Grant Program. https://www.ctoec.org/school-readiness-preschool-grant-program/. Accessed: 2024-03-13.

Connecticut Office of Elementary and Secondary Education (2020). Per Pupil Expenditures: Connecticut. https://oese.ed.gov/ppe/connecticut/. Accessed: January 12, 2024.

Connecticut State Department of Education (2008). Fall Kindergarten Entrance Inventory. Available at https://portal.ct.gov/-/media/SDE/Student-Assessment/K-Assessment/Fall-Kindergarten-Entrance-Inventory--FINAL2008.pdf. Accessed: 2024-03-01.

Connecticut State Department of Education (2021). Kindergarten Entry Inventory Handbook. Technical report, Connecticut State Department of Education. Available at https://portal.ct.gov/-/media/SDE/Student-Assessment/K-Assessment/KEI-handbook.pdf. Accessed: 2024-03-01.

Connecticut State Department of Education (2023). District Profile and Performance Report for School year 2022-23. https://edsight.ct.gov/Output/District/HighSchool/0930011_202223.pdf. Accessed: 2024-03-01.

Connecticut State Department of Education (2024). Kindergarten Inventory. https://portal.ct.gov/SDE/Student-Assessment/K-Assessment/Kindergarten-Inventory. Accessed: 2024-03-01.

Correia, S., P. Guimarães, and T. Zylkin (2020). Fast Poisson Estimation with High-Dimensional Fixed Effects. The Stata Journal 20(1), 95–115.

Friedman-Krauss, A. H., W. S. Barnett, and J. K. Duer (2022). The State(s) of Head Start and Early Head Start: Looking at Equity. Technical report, National Institute for Early Education Research, New Brunswick, NJ.

Friedman-Krauss, A. H., W. S. Barnett, K. S. Hodges, K. A. Garver, G. Weisenfeld, B. A. Gardiner, and T. M. Jost (2023). The State of Preschool 2022: State Preschool Yearbook. Technical report, National Institute for Early Education Research, New Brunswick, NJ.

García, J. L. and J. J. Heckman (2022). On Criteria for Evaluating Social Programs. Working Paper 30005, National Bureau of Economic Research.

García, J. L., J. J. Heckman, D. E. Leaf, and M. J. Prados (2020). Quantifying the Life-Cycle Benefits of an Influential Early-Childhood Program. Journal of Political Economy 128(7), 2502–2541.

Gelbach, J. B. (2002). Public Schooling for Young Children and Maternal Labor Supply. American Economic Review 92(1), 307–322.

Gibbs, C. R., J. Wikle, and R. Wilson (2024, July 10). A Matter of Time? Measuring Effects of Public Schooling Expansions on Families' Constraints. EdWorkingPaper 24-987, Brown University.

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.

Hendren, N. and B. Sprung-Keyser (2020). A Unified Welfare Analysis of Government Policies. The Quarterly Journal of Economics 135(3), 1209–1318.

Kabay, S., C. Weiland, and H. Yoshikawa (2020). Costs of the Boston Public Prekindergarten Program. Journal of Research on Educational Effectiveness 13(4), 574–600.

Kleven, H., C. Landais, and J. E. Søgaard (2019). Children and Gender Inequality: Evidence from Denmark. American Economic Journal: Applied Economics 11(4), 181–209.

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.

Landivar, L. C., N. L. Graf, and G. A. Rayo (2023). Issue Brief: Childcare Prices in Local Areas, Initial Findings from the National Database of Childcare Prices. Technical report, U.S. Department of Labor, Women's Bureau.

Lefebvre, P. and P. Merrigan (2008). Child-Care Policy and the Labor Supply of Mothers with Young Children: A Natural Experiment from Canada. Journal of Labor Economics 26(3), 519–548.

Lin, W. and J. M. Wooldridge (2019). Chapter 2 - Testing and Correcting for Endogeneity in Nonlinear Unobserved Effects Models. In M. Tsionas (Ed.), Panel Data Econometrics, pp. 21–43. Academic Press.

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.

Mas, A. and A. Pallais (2019). Labor Supply and the Value of Non-Work Time: Experimental Estimates from the Field. American Economic Review: Insights 1(1), 111–126. NBER Working Paper 23906.

New Haven Board of Education (2022). Superintendent's Proposed Budget Expenditure 2022-2023 School Year.

NHPS Office of Early Childhoood (2024). Office of Early Childhood. Available at https://www.nhps.net/domain/68. Accessed February 29th, 2024.

Teach Connecticut (2021). CT Teacher Salary: What You Need to Know. https://connecticut.teach.org/salary-benefits#references. [Accessed 22-08-2024].

Zimmerman, S. D. (2014). The Returns to College Admission for Academically Marginal Students. Journal of Labor Economics 32(4), 711–754.