The Long-Run Impacts of Universal Pre-K with Equilibrium Considerations

Jordan S. Berne
University of Michigan

Since 1995, publicly funded pre-K with universal eligibility has proliferated across the U.S. Universal pre-K (UPK) operates at great scale and serves children with a wide range of alternative childcare options. Because these programs are relatively young, very little is known about their long-run impacts on children. In this paper, I use a difference-in-differences (DiD) design to estimate the long-run impacts of Georgia UPK, the first statewide program. Children exposed to UPK were 1.7% more likely to graduate high school, 11.1% less likely to receive SNAP benefits as adults, and girls were 10.6% less likely to have children as teenagers. To help interpret those results, I develop a simple conceptual framework that considers how public pre-K expansions can affect the entire childcare market. For instance, greater competition could force private centers to adjust prices and quality, or to close entirely—creating spillover impacts on children not enrolled in public pre-K. Empirically, I find evidence consistent with large spillovers in Georgia, suggesting that a focus on UPK enrollees would miss a key part of the program's overall impact. Further, I show that conventional DiD estimates of treatment effects on the treated may be substantially biased in the presence of spillovers—in the Georgia context and in others.

VERSION: October 2024

Suggested citation: Berne, Jordan S.. (2024). The Long-Run Impacts of Universal Pre-K with Equilibrium Considerations. (EdWorkingPaper: 22-626). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/k4bh-0114

The Long-Run Impacts of Universal Pre-K with Equilibrium Considerations*

Jordan S. Berne
University of Michigan

October 22, 2024

Click here for latest draft

Abstract

Since 1995, publicly funded pre-K with universal eligibility has proliferated across the U.S. Universal pre-K (UPK) operates at great scale and serves children with a wide range of alternative childcare options. Because these programs are relatively young, very little is known about their long-run impacts on children. In this paper, I use a difference-in-differences (DiD) design to estimate the long-run impacts of Georgia UPK, the first statewide program. Children exposed to UPK were 1.7% more likely to graduate high school, 11.1% less likely to receive SNAP benefits as adults, and girls were 10.6% less likely to have children as teenagers. To help interpret those results, I develop a simple conceptual framework that considers how public pre-K expansions can affect the entire childcare market. For instance, greater competition could force private centers to adjust prices and quality, or to close entirely—creating spillover impacts on children not enrolled in public pre-K. Empirically, I find evidence consistent with large spillovers in Georgia, suggesting that a focus on UPK enrollees would miss a key part of the program's overall impact. Further, I show that conventional DiD estimates of treatment effects on the treated may be substantially biased in the presence of spillovers—in the Georgia context and in others.

Keywords: Universal preschool, state-funded pre-K, long-run effects, educational attainment, equilibrium effects, Georgia

*Thank you to Brian Jacob, Christina Weiland, Michael Mueller-Smith, and Ana Reynoso for excellent advising, and to many others who improved this project tremendously, such as Bruno Ferman, Ariza Gusti, Andrew Joung, Sarah Miller, Michael Ricks, Nathan Sotherland, and Kevin Stange. Thanks also to participants of the University of Michigan's Causal Inference in Education Research Seminar, Labor Economics Seminar, and Equity in Early Learning Lab. Thank you to J. Clint Carter for help using the Census data. I gratefully acknowledge funding support from the U.S. Department of Education's Institute of Education Sciences (R305B20011) and the NAEd/Spencer Dissertation Fellowship program. I am a member of the 1998 pre-K cohort exposed to Georgia UPK but do not judge there to be a conflict of interest estimating coefficients for event time 3. Any views expressed are those of the author and not those of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 3054. (CBDRB-FY24-P3054-R11701 and CBDRB-FY25-0013).

1 Introduction

In the United States, large achievement disparities emerge early in childhood based on socioeconomic status. At the start of kindergarten, for example, children at the bottom of the family income distribution score 1.3 standard deviations lower than children at the top in mathematics (Duncan and Magnuson, 2011). These early disparities persist through childhood and contribute to long-run disparities in educational attainment, health, and income (Currie and Almond, 2011; Nielsen, 2023).

Publicly funded prekindergarten (pre-K) has been a popular policy tool for combatting those disparities. Since 1980, the number of states with state-funded pre-K has risen from 4 to 45.[1] Most public pre-K programs are "targeted," meaning they have eligibility requirements related to a family's socioeconomic status. Since the 1990s, though, there has been consistent policy interest in universal pre-K (UPK), or programs with no eligibility criteria other than age. Ten states now have UPK programs and four more are creating new programs or converting targeted programs (Friedman-Krauss et al., 2023).[2] At the federal level, President Biden has pushed for a national UPK program for three- and four-year-olds.

In this paper, I investigate universal pre-K in Georgia, the first statewide program in the U.S. I document novel reduced form evidence on UPK's long-run impacts and illustrate how equilibrium responses in the childcare market add nuance to evaluations of large-scale pre-K programs. Specifically, I answer the following research questions: What are UPK's long-run impacts on teen fertility, educational attainment, government benefit receipt, employment, and earnings? How do those impacts vary across localities? How do equilibrium responses in the childcare market affect the interpretation of pre-K program evaluations?

Building the evidence base on universal programs is crucial as more states implement and expand UPK. Because most universal programs are too young, nearly all of the existing long-run research is on targeted programs (Cascio, 2021). It is unclear whether universal programs will deliver the same results. Compared to targeted programs, universal ones tend to serve children from more economically advantaged families who have higher quality childcare options in the absence of UPK. Universal programs also tend to be much larger than targeted programs.[3] It is difficult to maintain quality as programs grow, and larger programs are more likely to generate equilibrium responses (which have hard-to-predict implications) (Gupta et al., 2021).

I begin my analysis by examining short-run pre-K enrollment patterns following Georgia UPK's introduction. Nearly every county enrolled a substantial share of four-year-olds in UPK, but enrollment was higher in more economically disadvantaged counties (even controlling for

[1] In the 2021-22 school year, 35% of all four-year-olds enrolled in state-funded pre-K and 6% enrolled in the federal Head Start program (Friedman-Krauss et al., 2023).

[2] Notably, there is considerable variation in the implementation of UPK across states. For example, Georgia and Alabama require that pre-K classrooms operate for at least 6.5 hours a day for five days a week, whereas Iowa and Vermont only require 10 hours a week. Alabama contributes $10,881 per child for UPK while Georgia contributes only $5,646 (Friedman-Krauss et al., 2023).

[3] In the 2019-20 school year, average enrollment among four-year-olds was 21% in state targeted programs and 64% in state universal programs (Friedman-Krauss et al., 2023).

urbanicity). However, a difference-in-differences (DiD) analysis reveals that counties that enrolled more children in UPK *did not* experience correspondingly larger causal increases in public (or overall) pre-K enrollment. Those types of counties experienced relatively large increases in pre-K enrollment in the control group, suggesting they would have in Georgia too in the absence of UPK. As a result, the net change in public (overall) pre-K enrollment was around 11-15 percentage points (9-11 percentage points) throughout Georgia.

After documenting UPK's net effects on enrollment, I then estimate its long-run intent-to-treat (ITT) impacts on children who were four years old between between 1995 and 2000. The youngest of these cohorts have now reached their late 20s and the oldest their early 30s. Drawing on multiple restricted-use datasets from the U.S. Census Bureau (American Community Survey, SSA Numident file, and the Census Household Composition Key), I use a DiD research design to obtain causal estimates. UPK raised high school graduation rates by 1.5 percentage points (1.7%), reduced SNAP benefit receipt as an adult by 1.9pp (11.1%), and reduced the likelihood of girls having children as teenagers by 1.7pp (10.6%). I also find suggestive evidence that UPK increased bachelor's degree attainment and earnings, but these estimates are imprecise.

Previous long-run UPK studies have focused on a single city (Tulsa or Boston); Georgia provides an opportunity to estimate heterogeneity across a large, diverse state. Mirroring the short-run enrollment effects, I find that UPK's long-run impacts were broadly similar across county types. I do, however, find modest impact heterogeneity by race. White children appear to have benefited more from UPK than Black children, although the differences are imprecise.

Taking a step back, equilibrium responses in the public and private childcare market complicate the interpretation of ITT results. On the demand side, a large public expansion could change community norms about pre-K attendance and affect enrollment in all types of childcare (Rothbart and Morrissey, 2024). On the supply side, an expansion could force competitors to adjust prices and services, or to shut down entirely (Bassok, 2012; Bassok et al., 2014, 2016; Brown, 2019). Using a simple conceptual framework, I show how responses like those can cause a pre-K expansion to affect all children, not just those who enroll in the expanded program. In this case, understanding market responses is key to understanding how a program generates impacts, and whether impacts would replicate in other contexts. Unfortunately, market responses also present issues for methods of estimating treatment effects on the treated (TOT) that are commonly used in the DiD pre-K literature. Dividing an aggregate ITT effect by the net change in pre-K enrollment—which is commonly done—could misattribute impacts on the whole childcare market to only a subgroup of children.

In the Georgia setting, it is an empirical question whether equilibrium responses played a role in UPK's overall impact. I investigate this question by estimating substitution between no pre-K, private pre-K, and public pre-K enrollment following the introduction of UPK. Without market responses, UPK would only induce children to substitute into public pre-K. With market responses, children might substitute between any or all options. Unfortunately, DiD assumptions only identify the *net* effects of UPK on enrollment. The underlying flows between programs—which may cancel out on net—are obscured. To make progress on this issue, I combine a conditional independence assumption with detailed household data from the 2000 census. My estimates suggest market responses generated meaningful substitution between UPK and non-UPK childcare options.

For instance, I estimate that 14% of all four-year-olds in Georgia enrolled in public pre-K but would have enrolled in private pre-K if not for UPK. With this enrollment vacuum in private programs, 5% of all children switched from no pre-K to private pre-K. Separately, the estimates show that Georgia UPK had unintended consequences: 6% of all children would have enrolled in public pre-K (Head Start or a local program) but instead did not enroll in any pre-K, likely because UPK crowded out other public programs. These enrollment patterns, and the others I uncover, are consistent with equilibrium forces playing a large role in UPK's long-run impacts. As a result, I find that TOT effects estimated in the "usual way" could be substantially overstated—by as much as 3 times in a series of plausible scenarios I consider.

Before concluding I conduct a cost-benefit analysis of the introduction of Georgia UPK. Using the marginal value of public funds (MVPF) approach of Kline and Walters (2016) and Cascio (2023), my conceptual framework illustrates that the relevant parameters for welfare analysis are the long-run effects I estimate in the main analysis. Intuitively, an evaluation of a large policy must account for all equilibrium costs and benefits (Finkelstein and Hendren, 2020). My estimates suggest UPK paid for itself in the long run by raising tax revenue and lowering spending on SNAP. However, my earnings impact estimate is imprecise and the MVPF is sensitive to misspecification of this input. To overturn the finding that UPK's benefits exceeded its costs, the earnings impact would have to be less than \$103 (compared to my point estimate of \$1,006).

This paper advances the pre-K literature in three ways. First, alongside research in Tulsa and Boston, I provide some of the first estimates of UPK's long-run impacts in the U.S., including impacts on important outcomes not yet included in this evidence base (teen fertility, government benefit receipt, employment, and earnings).[4] Most of the pre-K literature has examined short- and medium-run impacts, and nearly all long-run research has examined targeted pre-K.[5] My findings generally align with the Tulsa and Boston findings. Using propensity score weighting, Gormley Jr. et al. (2023) find that UPK enrollees in Tulsa in 2005 were more likely to enroll in college than observationally equivalent peers. Gray-Lobe et al. (2023) study the effects of Boston UPK between 1997 and 2003 using a highly credible lottery admissions design. Enrollees were more likely to graduate high school and enroll in college "on-time."

Second, Georgia provides the first opportunity to estimate UPK's long-run impacts across

[4] In contemporaneous ongoing work, Andrew Barr, Jonathan Eggleston, and Alex Smith explore the long-run impacts of universal preschool in Georgia and Oklahoma.

[5] A non-exhaustive list of reduced form papers examining pre-K's short- and medium-run impacts is Berne et al. (2024a); Cascio and Schanzenbach (2013); Cascio (2023); Durkin et al. (2022); Feller et al. (2016); Fitzpatrick (2008); Harden et al. (2023); Kline and Walters (2016); Kose (2023); Weiland and Yoshikawa (2013); Weiland et al. (2020); Yang (2024). A non-exhaustive list of papers that estimate the long-run impacts of targeted programs is Anders et al. (2023); Bailey et al. (2021); Deming (2009); Johnson and Jackson (2019); Ludwig and Miller (2007); Pages et al. (2020); Thompson (2018).

different localities. Some studies of targeted programs have shown that long-run impacts can vary systematically by urbanicity or other local characteristics (Anders et al., 2023; Bailey et al., 2021; Johnson and Jackson, 2019). Relative to the studies in Tulsa and Boston, my investigation of a large, diverse state—and of heterogeneity across localities—may have greater external validity for other states considering UPK or for a federal program.

Third, my paper helps clarify our understanding of public pre-K's impacts in the presence of market adjustments. Researchers have long recognized the importance of general equilibrium thinking for interpreting reduced form estimates (Heckman et al., 1998, 2000). However, the reduced form pre-K literature has often not incorporated equilibrium insights into impact evaluations or welfare analyses.[6] Complications from market adjustments influencing childcare enrollment patterns have been underrecognized. My contribution is clarifying those complications and quantifying their importance.

The paper proceeds as follows. Section 2 provides background on Georgia's UPK program. Section 3 estimates UPK's effects on pre-K enrollment throughout Georgia. In Section 4, I estimate UPK's long-run impacts on children. Section 5 develops the conceptual framework for analyzing public pre-K programs that generate equilibrium market responses. In Section 6 I apply the conceptual framework to the Georgia setting to estimate pre-K substitution patterns. Drawing on results from Sections 3 and 4, I conduct a cost-benefit analysis of Georgia UPK in Section 7 before concluding in Section 8.

2 Background on Georgia UPK

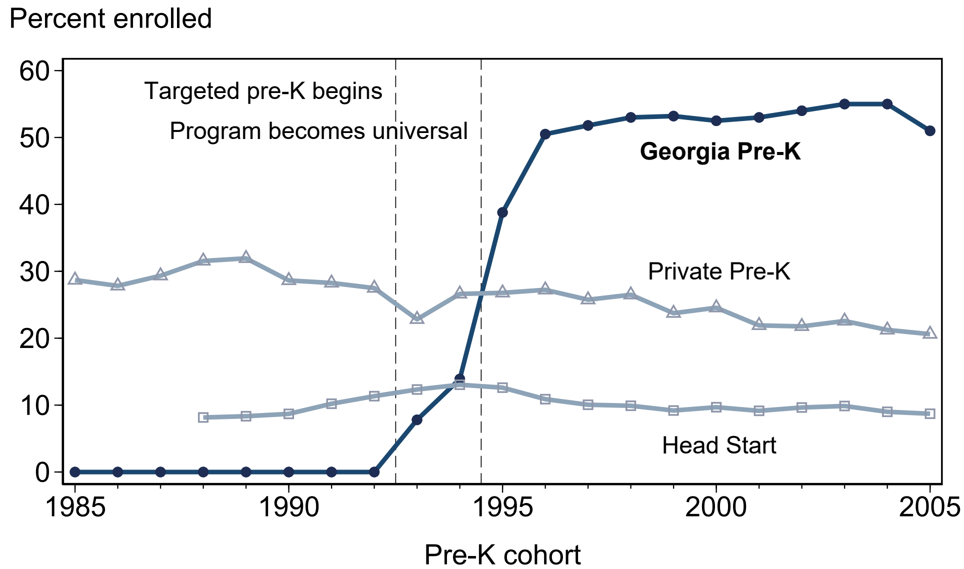
2.1 Program History

In 1993, with funding from a new state lottery game, the Georgia state government created a voluntary, income-targeted pre-K program to improve educational outcomes across the state. As Figure 1 shows, the program served 8% of all four-year-olds in its first year and 14% in its second year. In 1995, Georgia expanded eligibility to all children who turned four by September 1, making Georgia pre-K the first statewide UPK program in the country. Enrollment rose to 39% in the first year of universal eligibility and has hovered around 50%-60% ever since.

Georgia provided funding directly to local providers that met UPK program requirements, which were more stringent than general childcare licensing requirements. Around 90% of all UPK classrooms were run by public school systems (43%) and for-profit companies (47%), but local providers also included some non-profits (Henry et al., 2003a). To receive UPK funding, providers had to offer services for at least 6.5 hours a day, five days a week, during the local school year; classrooms could have no more than 20 students and had to maintain a minimum

[6] Some notable examples of pre-K papers that wrestle with general equilibrium issues are Zerpa (2022), who identifies (strict) assumptions that identify treatment effects on the treated, and Kline and Walters (2016), who account for the possibility that public pre-K programs free up slots in rationed private pre-K programs. Some examples of structural papers that model equilibrium responses in the childcare market include Berlinski et al. (2024), Bodéré (2023), Borowsky et al. (2022), and Griffen (2019).

Figure 1 — Pre-K Enrollment Trends Among Georgia Four-Year-Olds



Notes: The number of children enrolled in Georgia UPK is from the Georgia Department of Early Care and Learning. The number of children enrolled in Head Start is from Head Start Program Information Reports. To estimate shares, I divide by the number of four-year-olds in Georgia in each year as estimated by the National Cancer Institute Surveillance, Epidemiology, and End Results Program. I calculate participation in private pre-K using the October Supplement to the Current Population Survey; to reduce noise, I plot a five-year moving average.

staff-to-child ratio of 1:10; teachers had to be certified in early childhood education; and providers had to use a pre-approved curriculum. Nearly two-thirds of all classrooms used the HighScope curriculum (Bryan and Henry, 1998).

UPK expansion was driven by a combination of new centers and existing (public and private) centers converting to UPK centers (Bassok et al., 2014). There is little data on this, however, and it is unclear what fraction of UPK centers preexisted in another form. Historical accounts do show that conversions frequently involved meaningful programmatic changes to comply with funding requirements, such as adopting universal admissions, community engagement programs, mandatory staff training sessions, funding formulas based on teacher credentials, curricula restrictions, and increased oversight (Raden, 1999).

Administrative statistics on the demographics of children enrolled in Georgia UPK do not exist in the program's early years. The best information comes from a state-representative random sample of 203 classrooms in 1996 (Henry et al., 2003a). In this sample, children enrolled in Georgia UPK were more likely to be non-Hispanic Black, less likely to be non-Hispanic White, and tended to have lower family income compared to the population of four-year-olds in Georgia.

2.2 Program Quality

Overall, the quality of Georgia UPK during my analysis may be described as relatively good for its time. The best evidence on program quality comes from early studies that use the Early Childhood Environment Rating Scale (ECERS), in which scores range from 1 to 7 with 1=inadequate, 3=minimal, 5=good, and 7=excellent. Among a sample of 100 UPK classrooms (that may or may not be representative) in 1997, the average overall ECERS score was 4.66 (SD=0.44)—slightly below "good" (Bryan and Henry, 1998). Comparisons across studies suggest UPK may have been slightly higher quality, on average, than alternative options. Tietze et al. (1996) calculate a mean ECERS score of 4.26 (SD=1.00) across 401 preschool classrooms in four states in 1992. In a sample of 32 Head Start classrooms in the South, Bryant et al. (1994) calculate a mean ECERS score of 4.24 (SD=0.46) in 1990 and 1991.

Digging deeper, early childhood research often distinguishes between two types of program quality: "structural quality," which refers to program inputs like classroom size and per-child spending, and "process quality," which refers to the social, emotional, physical, and instructional aspects of children's day-to-day experiences. As discussed, the Georgia state government guaranteed relatively good structural quality by linking funding to program inputs. On process quality, Georgia UPK may again be described as "good." ECERS has two subscales that most closely approximate process quality: "interactions" and "language-reasoning." The interactions subscale measures the quality of student-student interactions, teacher-student interactions, and general supervision. The language-reasoning subscale measures how well students are encouraged to communicate, develop reasoning skills, and engage with books and other materials. Among a sample of 69 UPK classrooms in 2001, mean scores were 5.5 for interactions and 4.9 for language-reasoning (Henry et al., 2003b).

2.3 Short- and Medium-Run Impacts

Previous studies on the introduction of Georgia UPK have not investigated long-run effects, but they have documented mostly positive effects on short- and medium-run outcomes.[7] Using propensity score matching and data from the 2001 pre-K cohort, Henry et al. (2006) find that UPK was more effective than Head Start for children from families with low income. At the beginning of kindergarten, UPK students outperformed Head Start students on cognitive assessments and teacher evaluations.

A few papers find positive effects on medium-run outcomes using the same difference-in-differences strategy I use. Fitzpatrick (2008) finds that Georgia UPK exposure raised fourth grade math and reading scores and reduced the probability of grade retention. Monnet (2019) and Cascio and Schanzenbach (2013) both exploit the introduction of UPK in Georgia and Oklahoma jointly. Monnet (2019) finds that exposure to UPK reduced the likelihood of being diagnosed with behavioral/conduct problems for children with less educated parents but increased the likelihood for children with more educated parents. Cascio and Schanzenbach (2013) find evidence that UPK exposure improved eighth grade math scores for students from families with

[7] Woodyard et al. (2023) find more mixed evidence for the 2012–2018 pre-K cohorts. They find Georgia UPK had positive effects on test scores in kindergarten, but that those effects faded out in the next few years, becoming negative by fourth grade. However, their results are not directly comparable to results from papers that study the introduction of Georgia UPK; Woodyard et al. (2023) examine a single school district in Metro Atlanta and pre-K cohorts 17 to 23 years after Georgia pre-K became universal.

low income. My paper advances this literature by estimating longer-run effects on academic and non-academic outcomes.

3 Pre-K Enrollment Across Georgia

The first step to understanding UPK's long-run impacts is understanding short-run enrollment patterns. I begin with a descriptive analysis of UPK enrollment throughout Georgia, which gives context to the long-run impacts but is also of interest in its own right given the well-documented disparity in pre-K enrollment between rural and urban areas in the U.S. (Swenson, 2008; Temple, 2009). I then estimate UPK's causal effect on pre-K enrollment, recognizing that pre-K enrollment might have grown in Georgia even without UPK.

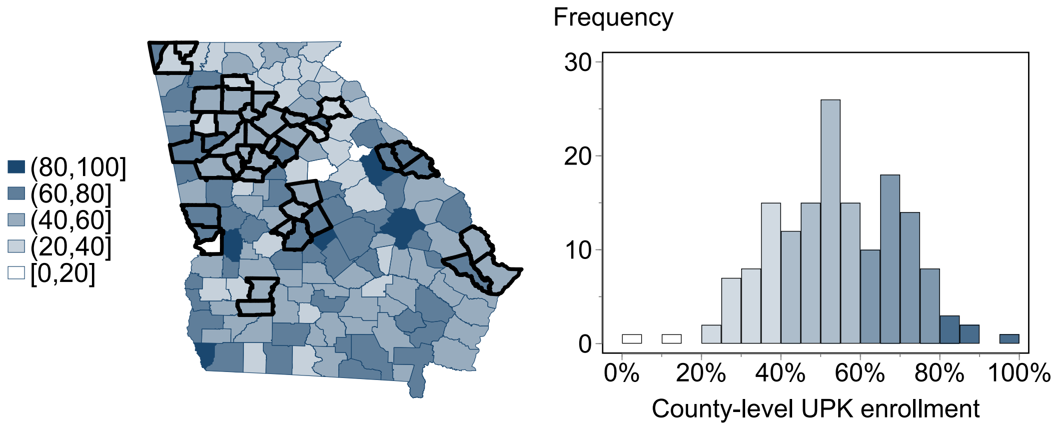
Previous analyses of pre-K enrollment in Georgia have been limited by two data constraints. First, almost all Georgia UPK papers measure pre-K enrollment using the Current Population Survey (CPS), which cannot distinguish UPK from other public pre-K programs (Cascio and Schanzenbach, 2013; Fitzpatrick, 2008; Monnet, 2019). Second, the CPS is not granular enough for sub-state analysis.

I overcome the first constraint using county-level UPK enrollment counts I collect from an old administrative report (Brackett et al., 1999) and UPK funding data at the county$\times$year level shared with me by Georgia's Department of Early Care and Learning. To the best of my knowledge, no previous paper has had enrollment rates specifically for Georgia UPK at sub-state levels. I overcome the second data constraint using the restricted 1990 and 2000 long-form censuses, in which I observe each child's county of residence and enrollment in private or public pre-K. The census data are substantially larger than the CPS and give me precision to estimate enrollment rates at various sub-state levels. For more detail on these datasets, see Appendix B.

Using the county-level UPK data, Figure 2 shows that UPK enrollment rates varied considerably across counties. In 1998, the fourth year of universal eligibility, county enrollment rates were centered around 50% but generally ranged from 30% to 80%. Even at the lower end, enrollment was remarkably high. One reason UPK was able to expand so quickly is that Georgia allowed public school systems *and* community-based organizations to operate UPK classrooms.

To examine variation across counties more systematically, I conduct a simple regression analysis that correlates county-level UPK enrollment rates with county-level characteristics (measured in the 1990 census before Georgia began funding pre-K). The full set of regression results is in Appendix Table A.1. Univariate regressions show that UPK enrollment rates were higher, on average, in non-metropolitan counties, counties with less pre-K enrollment (public+private) prior to UPK, and counties that were more economically disadvantaged on other dimensions (lower income, less educational attainment, and more single mother households)—counties that likely had untapped demand for preschool. In multivariate regressions that control for county characteristics simultaneously, the strongest predictors of UPK enrollment are baseline pre-K enrollment and single mother households.

Figure 2 — Share of Four-Year-Olds Enrolled in Georgia UPK in 1998, by County



Notes: In the map, thick black borders indicate metropolitan counties. Enrollment counts are obtained from Brackett et al. (1999) and converted to shares using estimates of the number of four-year-olds in each county in 1998 from the Surveillance, Epidemiology, and End Results Program at the National Cancer Institute.

Importantly, UPK enrollment did not translate to new pre-K enrollment one-for-one. To estimate the causal effects of UPK on pre-K enrollment, I estimate 2×2 DiD models. The 1989 pre-K cohort (from the 1990 census) is the "before" group, and the 1999 pre-K cohort (from the 2000 census) is the "after" group. Pre-K age children in Georgia form the treatment group, and pre-K age children in states without state-funded pre-K in 1999 form the control group.[8] Letting E_{ict} be an indicator for pre-K enrollment for child i of pre-K cohort t residing in county c, I estimate the following model:

$$E_{ict} = \eta UPK_i + \gamma_c + \lambda_t + \varepsilon_{ict}, \tag{1}$$

where UPK_i is a binary indicator for being exposed to Georgia UPK, γ_c is a vector of county fixed effects, and λ_t is a vector of pre-K cohort fixed effects. The effect of UPK on enrollment is η. I estimate the model separately for private pre-K, public pre-K, and any pre-K (private or public). Public pre-K captures UPK as well as Head Start and local publicly funded pre-K programs. Private pre-K captures tuition-based pre-K programs.

A causal interpretation requires a standard parallel trends assumption, which I evaluate using event study models and CPS data. (I cannot assess pre-trends with the census data because I observe enrollment in only one pre-treatment year.) See Appendix C for more details on this analysis. Appendix Figures C.1 and C.2 show that pre-K enrollment was extremely similar in Georgia and in the control group—in trends and levels—prior to UPK.

[8]The 18 states with no state pre-K program in 1999 are Alabama, Alaska, Florida, Hawaii, Idaho, Indiana, Minnesota, Mississippi, Montana, Nebraska, Nevada, New Hampshire, North Carolina, North Dakota, Rhode Island, South Dakota, Utah, and Wyoming.

Table 1 — Net Effects of Georgia UPK on Pre-K Enrollment, by County Type (DiD Estimates)

	Any Pre-K	Private Pre-K	Public Pre-K	Any Pre-K	Private Pre-K	Public Pre-K
	All Counties					
UPK impact	0.097**	−0.032**	0.129***			
(standard error)	(0.047)	(0.016)	(0.047)			
Impact in %	31.0%	−13.0%	29.3%			
Counterfactual mean	0.313	0.247	0.441			
Observations		541,000				
	Below Median Pre-K Enrollment			Above Median Pre-K Enrollment		
UPK impact	0.119**	−0.017	0.136**	0.088**	−0.037*	0.126***
(standard error)	(0.050)	(0.018)	(0.055)	(0.042)	(0.019)	(0.044)
Impact in %	31.2%	−10.1%	30.2%	30.9%	−13.3%	28.8%
Counterfactual mean	0.381	0.168	0.451	0.285	0.278	0.437
Observations		143,000			398,000	
	Non-Metropolitan			Metropolitan		
UPK impact	0.114*	−0.035*	0.149***	0.088**	−0.030	0.118***
(standard error)	(0.057)	(0.017)	(0.056)	(0.042)	(0.019)	(0.043)
Impact in %	31.2%	−21.5%	31.6%	31.1%	−10.2%	28.0%
Counterfactual mean	0.365	0.163	0.472	0.283	0.295	0.422
Observations		210,000			331,000	
	Below Median HH Income			Above Median HH Income		
UPK impact	0.142**	−0.035	0.177*	0.087*	−0.032*	0.119***
(standard error)	(0.064)	(0.022)	(0.089)	(0.043)	(0.017)	(0.043)
Impact in %	38.3%	−24.5%	36.5%	29.1%	−11.8%	27.7%
Counterfactual mean	0.371	0.143	0.485	0.299	0.272	0.429
Observations		140,000			401,000	

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: Author's calculations using data from the 1990 and 2000 decennial censuses. Inference is conducted using the Ferman and Pinto (2019) method. Each "counterfactual mean" is the average of observed enrollment among individuals exposed to Georgia UPK minus the estimated treatment effect. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701 and CBDRB-FY25-0013). Observation counts are rounded to comply with Census Bureau disclosure avoidance rules.

The DiD results are in Table 1. At the state level, the introduction of Georgia UPK reduced private pre-K enrollment by 3.2pp (13.0%), raised public pre-K enrollment by 12.9pp (29.3%), and raised overall pre-K enrollment by 9.7pp (31.0%). The bulk of the overall increase came from public pre-K enrollment, as expected. The 12.9pp net increase in public pre-K enrollment is substantial, but much smaller than the UPK enrollment rate in 1999 (53.2%). Part of this gap is due to preexisting centers converting to UPK centers, as discussed in Section 2. Further, public pre-K enrollment rose in the control group during this period, suggesting it would have in Georgia too in the absence of UPK.

It is important to note that these net changes in pre-K enrollment potentially mask offsetting inflows and outflows. As one example, I show in Appendix H that Georgia UPK reduced enrollment in Head Start. Some children who would have enrolled in Head Start likely enrolled in UPK instead, but this program substitution would not register as a net change in public or any pre-K enrollment. I dig deeper into this topic of childcare substitution in Section 5.

Table 1 also shows heterogeneity in enrollment effects along three key county-level dimensions: baseline pre-K enrollment, urbanicity, and household income.[9] These characteristics are related to children's counterfactual childcare options and the quality of other investments in children, which have been shown to moderate pre-K's effects in other contexts (Johnson and Jackson, 2019; Kline and Walters, 2016). For urbanicity, I identify metropolitan and non-metropolitan counties using the 1990 urban-rural classification scheme of the National Center for Health Statistics. For county-level pre-K enrollment and household income, I split counties into above and below median groups. For more detail, see Appendix B.

Although different counties enrolled very different shares of children in UPK, the causal net changes in pre-K enrollment were remarkably similar throughout Georgia. The types of counties with greater UPK enrollment were the types of counties that expanded local public pre-K in the control group states—but to a lesser degree. As a result, UPK raised public pre-K enrollment by aroud 11pp-15pp across county types. That effect was somewhat greater in counties with below median household income (17.7pp on average), but the difference is not statistically significant. UPK generally reduced private pre-K enrollment by 2pp-4pp and raised any pre-K enrollment by around 9pp-11pp.

4 Long-Run Impacts of Georgia UPK

4.1 Data and Sample

To estimate Georgia UPK's long-run impacts, I combine information from multiple restricted-use datasets within the U.S. Census Bureau data infrastructure. Primarily, I use annual (2005–2022) waves of the American Community Survey (ACS), a large cross-sectional survey representative of the U.S. resident population. The ACS forms the basis for my primary sample, which I use

[9] These county characteristics are correlated but distinct. See Appendix Table A.2 for a full set of cross-tabulations.

to estimate impacts on all outcomes other than having a child as a teenager. For teen fertility I construct a second sample that draws on additional Census Bureau data. I discuss this secondary sample at the end of this section and in Appendix B.

The main advantages of the ACS are its large sample size and that it has detailed information on adult outcomes, including educational attainment, government benefit receipt, employment, and earnings. The main disadvantage of the ACS is that I do not observe the same person as an adult and as a child. To obtain information about childhood circumstances, I merge people's county of birth onto the data from the Social Security Administration (SSA) Numident file.[10] Then, using birth county as the linking variable, I merge on county-level characteristics from the 1990 decennial census (such as urbanicity, median household income, and pre-K enrollment rate). Although I cannot observe details about people's particular families when they were children, the birthplace information adds useful context and allows me to investigate impact heterogeneity within Georgia.

The sample with data on birth county characteristics constitutes 81% of the full sample. Attrition occurs primarily because the Census Bureau cannot perfectly assign every person a unique identifier for linking between datasets, although 1.7% of the sample is also lost because of missing data on county characteristics.[11] I use the full sample to estimate overall impacts and demographic subgroup impacts, but analyses for *geographic* subgroups are limited to individuals with observed birth county characteristics.

To facilitate the difference-in-differences analysis, I restrict the sample to children whose pre-K year was around the introduction of Georgia UPK. Specifically, I focus on children in the 1987–2000 pre-K cohorts. Cohorts are defined using each child's exact birthdate and Georgia's September 1 kindergarten birthday cutoff. For example, children (in any state) who turn four between September 2, 1990 and September 1, 1991 are categorized as being in the 1991 pre-K cohort. See Appendix Table A.3 for the sample size of each pre-K cohort and the range of ages at which I observe each cohort.

The ACS data also have everyone's state of birth, which I use to define the DiD treatment and control groups. The treatment group consists of children born in Georgia. The control group consists of children born in states with no more than a very small state pre-K program during the analysis. Specifically, I use the 23 states with no more than five percent of four-year-olds enrolled in state-funded pre-K in 2001 (Friedman-Krauss et al., 2023).[12] This group includes five other states from the southeast (Alabama, Florida, Mississippi, North Carolina, and Tennessee).

[10]The raw Numident file contains birthplace locations that are usually more granular than counties. I gratefully draw on work by Taylor et al. 2016, who algorithmically match places to counties.

[11]9.1% of the full sample is lost because of unassigned unique identifiers. Another 8.0% is lost because of conflicting values for one's state of birth in the ACS and Numident, likely because of a mistake assigning an identifier. I defer to the self-reported birth state in the ACS, which requires dropping those with conflicting values from the birth county sample. Lastly, 1.7% of the full sample is lost because of missing data on birth county characteristics, even though birth county itself merges onto the ACS successfully.

[12]The 23 states control group states are Alabama, Alaska, Florida, Hawaii, Idaho, Indiana, Iowa, Minnesota, Mississippi, Missouri, Montana, Nebraska, Nevada, New Hampshire, New Mexico, North Carolina, North Dakota, Pennsylvania, Rhode Island, South Dakota, Tennessee, Utah, and Wyoming.

See Appendix Figure A.1 for a map of the full control group. Note that I include states with very small state pre-K programs because asymptotic inference methods require a sufficient number of clusters (i.e., states), but I show in Section 4.5 that the results are not sensitive to dropping control states with state-funded pre-K.

Table 2 presents summary statistics for the primary sample and birth county subsample. Children born in Georgia and the control group are mostly very similar, although Georgia children are more likely to be raised by a single mother and substantially more likely to be Black. The birth county subsample is extremely similar to the full sample, which suggests that differential selection into the subsample is not a large concern.

Lastly, because the ACS is not well-suited for measuring teen fertility (i.e., having a child as a teenager), I construct a second sample that draws on additional data sources. This sample uses the 2000 long-form census and the 2005-2022 ACS as a base and incorporates data on the timing of births from the Census Household Composition Key (CHCK), an underutilized dataset containing the near-universe of births in the U.S. (Genadek et al., 2021). See Appendix B for more information on the CHCK data and the construction of the teen fertility sample.

4.2 Outcome Measures

Teen fertility. The first outcome of interest is a binary indicator for becoming a parent as a teenager. Unlike the other outcomes in my analysis, this outcome cannot continue to evolve because the 1987–2000 pre-K cohorts have all completed their teenage years. Relatedly, for this outcome only, I do not drop anyone from the sample due to their age in the ACS or census because the CHCK data contain birth information regardless of when an individual is surveyed. Throughout the paper, I show results separately by sex because girls are much more likely to have children in their teens and because teen fertility has disproportionate consequences for girls.

High school graduation. To measure high school graduation, I construct a binary indicator that equals one for those who graduated high school or received an equivalent diploma (e.g., GED). When estimating impacts on high school graduation, I restrict the sample to individuals surveyed at 20 years or older. Nearly everyone who graduates high school does so by age 20 (see Appendix Figure A.2).

Bachelor's degree attainment. For bachelor's degree attainment and all remaining outcomes, I restrict the sample to those surveyed at 24 years or older. For people who begin college at 18 years old, this age matches the commonly-reported 6-year graduation rate statistic. Some people receive bachelor's degrees later in life, but the vast majority of graduates finish by 24 (see Appendix Figure A.2).

SNAP benefit receipt. In the ACS, receipt of Supplemental Nutrition Assistance Program (SNAP) benefits is a household-level question. My measure of SNAP receipt equals one if anyone in a person's household received SNAP benefits in the 12 months preceding the survey. Note that the SNAP sample is 6% smaller than the bachelor's degree sample—despite using the same age restriction—because respondents are only asked about SNAP receipt if they are in their long-term

Table 2 — Analysis Sample Summary Statistics, 1987-2000 Pre-K Cohorts

	Control Group		Georgia	
	Full Sample	Birth County Sample	Full Sample	Birth County Sample
Individual demographics				
Female	0.49	0.50	0.48	0.50
Non-Hispanic White	0.79	0.82	0.65	0.68
Non-Hispanic Black	0.11	0.10	0.30	0.28
Hispanic	0.05	0.04	0.03	0.03
Other race/ethnicity	0.05	0.04	0.02	0.02
Birth county characteristics (1990 Census)				
Metropolitan	-	0.72	-	0.73
White share	-	0.83	-	0.64
Black share	-	0.10	-	0.33
Household speaks English	-	0.90	-	0.93
Median household income	-	$28,100	-	$28,200
Graduated high school (25+)	-	0.77	-	0.72
Attained BA degree (25+)	-	0.19	-	0.21
Has single mother (0-5)	-	0.15	-	0.25
Not enrolled in pre-K (4-5)	-	0.50	-	0.49
Enrolled in private pre-K (4-5)	-	0.19	-	0.20
Enrolled in public pre-K (4-5)	-	0.31	-	0.31
Individual long-run outcomes				
Teen father	0.05	0.05	0.07	0.07
Teen mother	0.13	0.13	0.17	0.17
Graduated high school	0.92	0.93	0.89	0.90
Attained BA degree	0.35	0.37	0.32	0.34
Receive SNAP benefits	0.14	0.14	0.18	0.18
Employed or in school	0.79	0.81	0.74	0.77
Observations	2,506,000	2,037,000	192,000	153,000

Notes: All statistics in this table are means. The control group consists of the 23 states named in Section 4.3 and shown in Appendix Figure A.1. "Teen father" and "teen mother" are calculated using the secondary fertility sample, and all other statistics are calculated using the primary sample. I impose a minimum age restriction of 24 for "attained BA degree," "receive SNAP benefits," and "employed or in school," as discussed in Section 4.2, but all other (non-fertility) statistics restrict the sample to people 20 or older when surveyed (including the observation counts). The county of birth characteristics are derived from the 1990 census and then merged onto the main sample. The parentheticals next to some of the variable names indicate the age of sample members in the 1990 census used to calculate the county-level statistics. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701 and CBDRB-FY25-0013). Observation counts are rounded to comply with Census Bureau disclosure avoidance rules.

housing unit at the time of the survey.

Being employed or in school. The ACS asks respondents whether they've done any work for pay or profit in the last week. I construct a measure of employment that equals one if a respondent answers affirmatively or if they report being enrolled in school. I account for school enrollment because many individuals in their late 20s and early 30s enroll in graduate school and other degree granting and non-granting programs.

Earnings. The ACS asks respondents about their earnings in the 12 months preceding the survey. Using this information, I construct three earnings-related outcome measures: 1) whether a person had positive earnings in the previous year, 2) earnings in the previous year (including zeros and negative earnings), and 3) the log of earnings in the previous year (excluding zeros and negative earnings). All earnings outcomes are converted to 2020 U.S. dollars using the CPI-U.

4.3 Research Design

Throughout this section, the treatment of interest is *exposure* to Georgia UPK during one's pre-K year. This notion of treatment recognizes that any child in Georgia could have been affected by UPK regardless of their own enrollment. I define exposure with a binary indicator for being born in Georgia and belonging to the 1995 pre-K cohort or a later one. This definition may not perfectly capture the idea of exposure since it is based on state of birth rather than state of residence at age four. It has the advantage, though, of avoiding bias from families moving to Georgia because of UPK. Ultimately, the two definitions would likely produce similar results since 86% of children born in Georgia reside there at age four (in the 2000 long-form census).

To estimate long-run impacts, I compare the outcomes of exposed and non-exposed children in a DiD framework. The first "difference" is a comparison of children born in Georgia who turned four before vs. after the introduction of UPK. The second "difference" is a comparison of children born in Georgia (the treatment group) and children born in other states (the control group).

Letting Y_{ict} denote an adult outcome for individual i born in county c and in pre-K cohort t, my primary model specification is:

$$Y_{ict} = \beta_0 + \beta_1 UPK_i + \beta_2 TPK1_i + \beta_3 TPK2_i + \Pi \mathbf{X}_i + \gamma_c + \lambda_t + \varepsilon_{ict}, \tag{2}$$

where UPK_i is a binary indicator for being exposed to Georgia UPK; $TPK1_i$ and $TPK2_i$ are binary indicators for being exposed to Georgia's targeted pre-K program in one of the two years it existed; γ_c is a vector of birth county fixed effects; and λ_t is a vector of pre-K cohort fixed effects.[13] $\mathbf{X}_i$ is a covariate vector that contains race/ethnicity and sex indicators and, when the outcome variable is not teen fertility, a cubic polynomial for age when surveyed.[14] The effect of

[13] The 17.1% of individuals without an observed birth county are assigned to a within-birth-state residual county for the purpose of estimating birth county fixed effects.

[14] I include the polynomial for "age when surveyed" because older pre-K cohorts are observed in the data at older ages. However, the results are not sensitive to excluding the age controls (see Section 4.5).

interest is β_1. Note that issues with two-way fixed effects models under DiD assumptions are not a concern in this setting because treatment adoption is not staggered, all control units are never treated, and treatment never turns off once on.

Inference is challenging in this setting because there are too few states (i.e., clusters) in the analysis for conventional cluster-robust asymptotic inference. With only 24 clusters, the standard cluster-robust variance estimator would over-reject null hypotheses. Moreover, some inference methods that perform well with a small number of clusters, like the Wild cluster bootstrap, would not perform well here because there is only one treated state. To avoid over-rejecting, I use a method developed by Ferman and Pinto (2019) for situations with very few treated units. The method is an extension of the cluster residual bootstrap and the Conley and Taber (2011) method. At a high level, the idea is to estimate residual variation in the treated group using a model trained on control group clusters, allowing for arbitrary within-cluster correlation. Following Ferman and Pinto, I model heteroskedasticity generated by variation in state size (allowing sample sizes to vary within states across cohorts). Ferman and Pinto show in simulations that the method performs well with 1 treated unit and 24 control units.

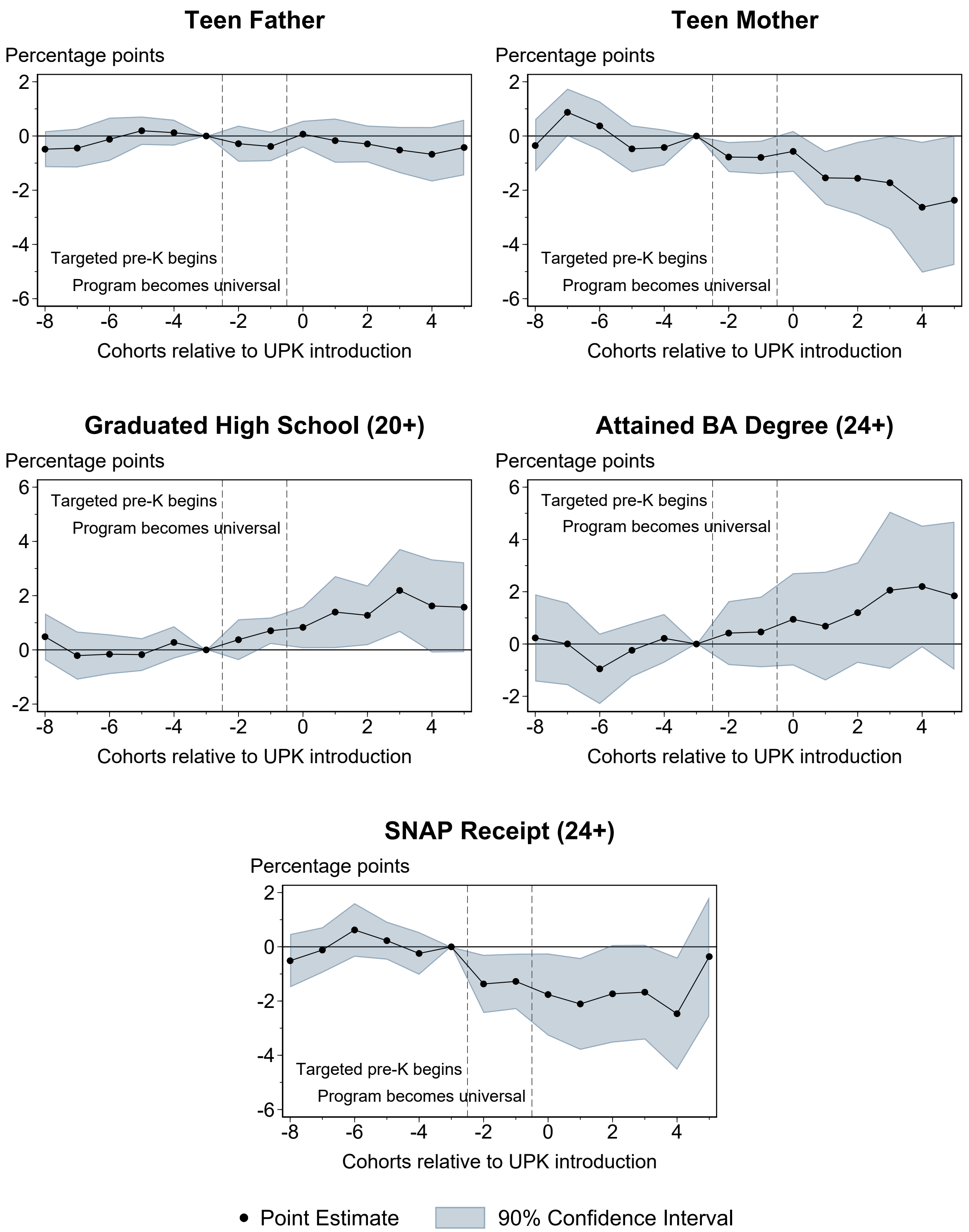
To interpret DiD estimates as causal, three assumptions must be met. The first is that the outcomes in the treatment and control groups would have followed parallel trends in the absence of UPK. I provide evidence supporting this assumption in Section 4.4 using dynamic event study models. The second assumption is that there were no effects on pre-K cohorts not exposed to treatment (”no anticipation”). Such effects would contaminate the counterfactual and bias the post-treatment effects. To satisfy this assumption, I include separate effects (β_2 and β_3) in equation 2 for the two cohorts (1993 and 1994) exposed to Georgia's short-lived targeted pre-K program.

The third assumption is sometimes referred to as ”no concurrent shocks.” The assumption is that no other policies or shocks differentially affected pre- and post-treatment cohorts. Any differential shock that occurred before the measurement of long-run outcomes could be problematic. To my knowledge, there are no shocks that might threaten a causal interpretation. Notably, Cascio and Schanzenbach (2013) make the same assumption to estimate the impact of UPK on eighth grade test scores, although the risk of conflation is smaller in their case because they examine medium-run outcomes.

4.4 Long-Run Impact Estimates

I first evaluate the plausibility of the parallel trends assumption by examining dynamic event study models. Beginning with the non-labor market outcomes, Figure 3 presents estimates from models that are equivalent to equation 2 except that they include a separate effect for each cohort (omitting the 1992 cohort). The estimates for cohorts -8 through -3 provide suggestive evidence on the parallel trends assumption, the estimates for cohorts -2 and -1 are impact estimates for targeted pre-K in Georgia, and the estimates for cohorts 0 through 5 trace out Georgia UPK's impacts for exposed cohorts.

Figure 3 — Georgia UPK's Long-Run Effects I (Event Studies)



Notes: Author's calculations from estimating event study versions of equation 2. Inference is conducted using the Ferman and Pinto (2019) method. Cohort 0 refers to the 1995 pre-K cohort. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701 and CBDRB-FY25-0013).

Table 3 — Georgia UPK's Long-Run Effects I (DiD Estimates)

	Teen Father	Teen Mother	Graduated High School	Attained BA Degree	SNAP Receipt
UPK impact	−0.003	−0.017**	0.015**	0.015	−0.019*
(standard error)	(0.003)	(0.009)	(0.008)	(0.013)	(0.010)
[p-value]	[0.429]	[0.019]	[0.043]	[0.315]	[0.060]
Impact in %	−4.5%	−10.6%	1.7%	4.9%	−11.1%
Counterfactual mean	0.067	0.160	0.885	0.308	0.171
Age restriction	None	None	20+	24+	24+
Observations	2,508,000	2,433,000	2,698,000	1,906,000	1,786,000

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

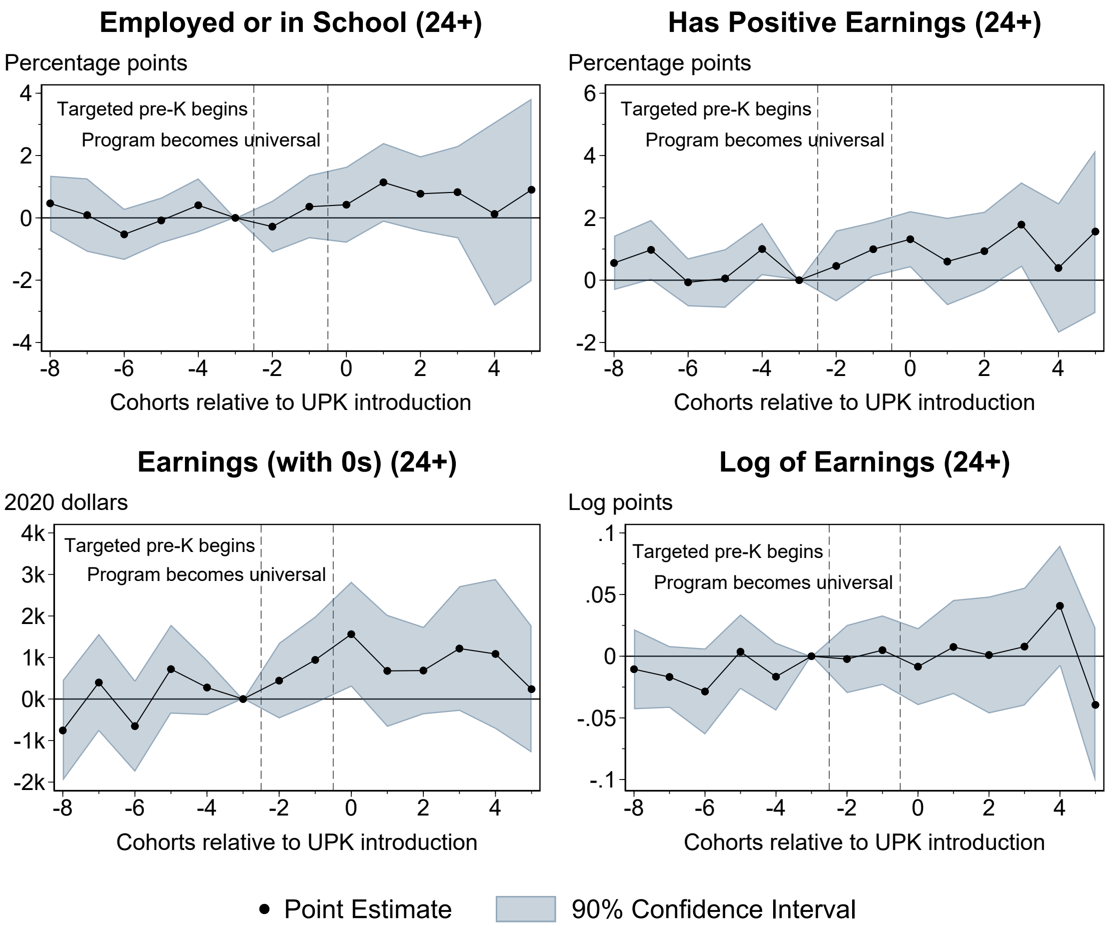
Notes: Author's calculations from estimating equation 2. Inference is conducted using the Ferman and Pinto (2019) method. Each "counterfactual mean" is the average of observed outcomes among individuals exposed to Georgia UPK minus the estimated treatment effect. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701 and CBDRB-FY25-0013). Observation counts are rounded to comply with Census Bureau disclosure avoidance rules.

The pre-treatment estimates in Figure 3 are all statistically insignificant and the trends are generally flat, supporting the parallel trends assumption. The post-treatment estimates indicate increases in educational attainment and reductions in female teen fertility and SNAP receipt. Reassuringly, the pattern of growing impacts generally mirrors growth in UPK enrollment over time. For some outcomes—SNAP receipt in particular—the impacts of Georgia's targeted pre-K program are large relative to UPK's impacts given UPK's greater enrollment. That result is consistent with the targeted pre-K program generating less crowd-out (as seen in Appendix Figures C.2 and H.1) and serving children with greater need.

Table 3 presents the DiD estimates. I find that exposure to Georgia UPK reduced teen mother rates by 1.7 percentage points (pp), a 10.6% reduction off a counterfactual base of 16.0%. Conversely, the effect on teen father rates is small and statistically insignificant. UPK exposure also increased high school graduation rates by 1.5pp (1.7%) and lowered SNAP receipt by 1.9pp (11.1%). Both magnitudes are economically meaningful, especially noting that they are effects of exposure, not enrollment. I also estimate that UPK raised bachelor's degree attainment by 1.5pp (4.9%), a considerable impact, but this estimate is less precise and not statistically different than zero.

It would also be of interest to define treatment as UPK *enrollment* and estimate impacts specifically for children induced to enroll. It is common in the DiD pre-K literature to estimate treatment effects on the treated by dividing the aggregate effect by changes in pre-K enrollment. However, interpreting such estimates as the impact of attending UPK could be wrong, potentially very wrong, if there are broader shifts in the childcare market stemming from the policy. I return to this issue in great detail in Sections 5 and 6.

Figure 4 — Georgia UPK's Long-Run Effects II (Event Studies)



Notes: Author's calculations from estimating event study versions of equation 2. Inference is conducted using the Ferman and Pinto (2019) method. Cohort 0 refers to the 1995 pre-K cohort. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701 and CBDRB-FY25-0013).

Table 4 — Georgia UPK's Long-Run Effects II (DiD Estimates)

	Employed or in School	Has Positive Earnings	Earnings (with 0s)	Log of Earnings
UPK impact	0.005	0.006	1,006	0.014
(standard error)	(0.009)	(0.007)	(644)	(0.019)
[p-value]	[0.488]	[0.405]	[0.100]	[0.529]
Impact in %	0.7%	0.7%	3.7%	-
Counterfactual mean	0.754	0.809	27,110	10.03
Age restriction	24+	24+	24+	24+
Observations	1,906,000	1,906,000	1,906,000	1,601,000

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: Author's calculations from estimating equation 2. Inference is conducted using the Ferman and Pinto (2019) method. Each "counterfactual mean" is the average of observed outcomes among individuals exposed to Georgia UPK minus the estimated treatment effect. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701 and CBDRB-FY25-0013). Observation counts are rounded to comply with Census Bureau disclosure avoidance rules.

Turning to labor market impacts, Figure 4 shows the event study results and Table 4 presents the DiD estimates. The pre-treatment coefficients in the event studies are noisy but not particularly concerning for the parallel trends assumption. Beginning with earnings (including zeros and negative earnings), the event study shows a clear upward trend after Georgia started funding pre-K. The DiD estimate is that UPK raised individual earnings by \$1,006 a year, on average, although the estimate is just shy of statistical significance at conventional levels. Taken at face value, the point estimate implies a 3.7% increase off a counterfactual base of \$27,110.

The overall earnings impact combines impacts on multiple margins—employment, hours worked, and wages. Some of these margins are testable in the data. The event study for having positive earnings suggests UPK may have increased people's likelihood of having earnings, but the DiD estimate is small and insignificant. Conversely, the DiD estimate for log earnings suggests UPK raised earnings (among those with earnings) by 1.4%, but the event study is less convincing. The estimates are not statistically significant in either case. Taken together, it is difficult to say definitively that UPK raised earnings, and if so, via which margin.

4.5 Robustness Checks

In this section I conduct a series of robustness checks to test the sensitivity and validity of the estimates in Tables 3 and 4. The results are generally very supporting. For a more detailed discussion, see Appendix D.

In the first set of checks, I repeat the main DiD analysis using five alternative regression specifications and samples. Specifically, I deviate from the main analysis by 1) removing all covariates from the model, 2) imposing a maximum age for sample inclusion of 32 when surveyed, 3) imposing a minimum age for sample inclusion of 25 when surveyed, 4) restricting the control group to the 18 states with no state pre-K program in 1999, and 5) replacing birth county fixed effects with birth state fixed effects. In alternative 2, imposing a maximum age addresses the fact that older cohorts are surveyed at older ages than younger cohorts. In alternative 4, the more restrictive control group reduces the number of clusters, which is not ideal for inference, but provides a potentially cleaner counterfactual. Across the board, the alternative estimates are highly similar to the primary estimates (see Appendix Table D.1).

For the second robustness check I re-estimate impacts using the synthetic control method (SCM). To avoid overfitting the pre-treatment data, I extend the sample back to the 1980 pre-K cohort for this analysis. Unfortunately, the CHCK birth data is unavailable for earlier cohorts. I therefore create a proxy for female teen fertility in the ACS using adults' and children's ages. Although the measure is imperfect, I nevertheless obtain very similar impact estimates.

Appendix Tables D.2 and D.3 present the weights used to construct the synthetic Georgias. The outcome trends for Georgia and the synthetic Georgias are shown in Appendix Figures D.1 and D.2. For the non-labor market outcomes, the impact estimates are highly similar to the main DiD estimates: -1.0pp (p-value $= 0.109$) for female teen fertility, 1.4pp ($p = 0.068$) for graduating high school, 1.6pp ($p = 0.056$) for attaining a bachelor's degree, and -1.6pp ($p = 0.082$) for

SNAP receipt. Unlike in the main analysis, the SCM estimate for bachelor's degree attainment is statistically significant. For the labor market outcomes, the SCM estimates are not quite as consistent. The SCM estimate for overall earnings ($251) is much smaller than the DiD estimate ($1,006). I find some evidence of positive impacts on having earnings and on log earnings, but neither estimate is significant.

For the third robustness check I examine whether Georgia UPK's impacts could be explained by demographic changes rather than early childhood education per se. Using the birth county subsample, I first regress each outcome on individual and birth county covariates, using only the control group. Then, I use the estimated coefficients to obtain predicted outcomes for Georgia and the control group. Finally, I estimate DiD models with birth state fixed effects, no covariates other than age when surveyed, and the predicted values as the dependent variables. Non-zero estimates could suggest that the introduction of Georgia UPK coincided with compositional changes that, on their own, affected long-run outcomes. One might expect compositional changes if families moved to Georgia (or moved within Georgia) specifically because of UPK. However, as Appendix Table D.4 shows, I find no evidence that demographic changes drive the main estimates.

My last robustness check tests whether changes in living arrangements could explain the SNAP impact (-1.9pp). Recall that the SNAP indicator in the ACS equals one if any person in one's household receives benefits. The estimated reduction in SNAP receipt could therefore reflect causal changes in the households children reside in as adults. A decrease in adult SNAP receipt might occur, for instance, if UPK made children more likely to live with their parents in their late 20s or early 30s. To investigate this, I estimate DiD models for the following two outcomes: 1) being the designated householder in the ACS, and 2) having a household member who is at least 18 years older (a proxy for parents). Reassuringly, I obtain small and statistically insignificant effects for both outcomes: -0.4pp ($p = 0.738$) for being the householder and -0.6pp ($p = 0.731$) for having older household members. (See Appendix Figure D.3 for analogous event study results.)

4.6 Long-Run Impact Heterogeneity

This section investigates heterogeneity in Georgia UPK's long-run impacts, first by county-level characteristics and then by sex and race. I focus on the same county types I examined in the enrollment analysis, defined by baseline pre-K enrollment, urbanicity, and income level. In each analysis I limit the sample to the subgroup of interest and re-estimate equation 2. Given the findings in the previous section, I defer the labor market outcomes to Appendix Tables A.4 and A.5. I find moderate heterogeneity by race, and some indications of heterogeneity across county types, but the differences are imprecise and not particularly large.

Beginning with baseline pre-K enrollment, Table 5 shows that impacts in below median counties are large in levels: -2.4pp for female teen fertility, 2.4pp for high school graduation, and -2.6pp for SNAP receipt. In percentage terms, however, the impacts for below vs. above median counties are quite similar: −10.9% vs. −8.9% for teen motherhood, 2.8% vs. 1.4% for high school graduation, and −12.6% vs. −10.2% for SNAP receipt.

For urbanicity, Table 5 shows that impacts on teen fertility are similar in levels but slightly greater in percentage terms for children born in metropolitan areas: −1.5pp (−10.9%) vs. −1.6pp (−7.6%). Impacts on SNAP receipt are greater in levels and percentage terms in metropolitan areas: −2.2pp (−13.8%) vs. −0.9pp (−4.4%). High school graduation impacts are the same in levels and percentages. Mirroring the urbanicity results, children from higher income counties experienced greater impacts on teen fertility and SNAP in percentage terms, while high school graduation impacts are very similar in levels and percentages.

Larger percentage impacts in metropolitan and higher income counties would be consistent with those areas being better positioned to attract and retain effective teachers. They would also be consistent with complementarities between preschool enrollment and other investments in children (Johnson and Jackson, 2019). However, the differences in the estimates across counties are modest and not statistically significant. A safer conclusion may be that Georgia UPK appears to have had positive and broadly similar effects across different types of counties.

This finding is rationalized by the similarity in UPK's net effects on pre-K enrollment throughout Georgia. Recall that UPK raised public pre-K enrollment and reduced private pre-K enrollment by similar amounts across county types. In theory, of course, enrollment changes of the same magnitude could produce different impacts in different counties. To a first approximation, however, the similarity in enrollment effects provides a compelling explanation for the similarity in long-run impacts.

Table 6 shows how impacts varied by race and sex. Note that I focus on non-Hispanic White and non-Hispanic Black children because other groups have sample sizes too small for generating informative estimates. For female teen fertility, impacts are clearly driven by White girls. The point estimate for Black girls is actually positive, but indistinguishable from 0. Impacts on high school graduation rates tend to be larger for White children than Black children, and for White boys in particular (1.8pp). The differences in impacts are modest in percentage terms though. Lastly, the SNAP impacts are greater in levels for Black boys and girls, but greater in percentage terms for White boys and girls due to their lower rates of benefit receipt in the absence of UPK.

Overall, White children experienced somewhat greater impacts than Black children. Interestingly, White children are comparatively more likely to live in Georgia's non-metropolitan areas—areas that perhaps experienced slightly *smaller* impacts on female teen fertility and SNAP receipt. This suggests that the heterogeneity by race is not driven by differences in residential patterns. Unfortunately, I lack the statistical power to estimate heterogeneity by race within county types.

Table 5 — Georgia UPK's Long-Run Effects, by County Type I

	Teen Father	Teen Mother	Graduated High School	Attained BA Degree	Receive SNAP
			Panel A. Below Median Pre-K Enrollment		
UPK impact	−0.005	−0.024***	0.024**	0.025	−0.026***
(standard error)	(0.003)	(0.010)	(0.011)	(0.018)	(0.011)
Impact in %	−5.6%	−10.9%	2.8%	12.7%	−12.6%
Counterfactual mean	0.090	0.220	0.862	0.197	0.207
Observations	484,000	469,000	425,000	311,000	301,000
			Panel B. Above Median Pre-K Enrollment		
UPK impact	−0.001	−0.013	0.013*	0.017	−0.017*
(standard error)	(0.003)	(0.008)	(0.008)	(0.015)	(0.011)
Impact in %	−1.6%	−8.9%	1.4%	4.9%	−10.2%
Counterfactual mean	0.061	0.146	0.899	0.348	0.166
Observations	2,024,000	1,964,000	1,766,000	1,262,000	1,220,000
			Panel C. Non-Metropolitan		
UPK impact	0.001	−0.016*	0.015	0.016	−0.009
(standard error)	(0.003)	(0.010)	(0.011)	(0.015)	(0.010)
Impact in %	1.2%	−7.6%	1.7%	7.7%	−4.4%
Counterfactual mean	0.085	0.211	0.867	0.208	0.205
Observations	689,000	672,000	604,000	444,000	431,000
			Panel D. Metropolitan		
UPK impact	−0.003	−0.015*	0.015**	0.020	−0.022*
(standard error)	(0.004)	(0.009)	(0.007)	(0.014)	(0.012)
Impact in %	−5.2%	−10.9%	1.7%	5.5%	−13.8%
Counterfactual mean	0.058	0.138	0.903	0.366	0.160
Observations	1,818,000	1,762,000	1,587,000	1,129,000	1,090,000
			Panel E. Below Median Household Income		
UPK impact	0.004	−0.015	0.013	0.000	−0.009
(standard error)	(0.005)	(0.011)	(0.011)	(0.021)	(0.015)
Impact in %	4.7%	−7.0%	1.5%	0.0%	−4.0%
Counterfactual mean	0.086	0.214	0.869	0.223	0.225
Observations	524,000	509,000	464,000	337,000	325,000
			Panel F. Above Median Household Income		
UPK impact	−0.002	−0.015*	0.015*	0.020	−0.019**
(standard error)	(0.003)	(0.009)	(0.008)	(0.017)	(0.010)
Impact in %	−3.2%	−10.1%	1.7%	5.9%	−11.6%
Counterfactual mean	0.062	0.149	0.897	0.341	0.164
Observations	1,984,000	1,924,000	1,727,000	1,237,000	1,196,000

*** p<0.01, ** p<0.05, * p<0.1

Notes: Author's calculations from estimating equation 2. Inference is conducted using the Ferman and Pinto (2019) method. Each "counterfactual mean" is the average of observed outcomes among individuals exposed to Georgia UPK minus the estimated treatment effect. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701 and CBDRB-FY25-0013). Observation counts are rounded to comply with Census Bureau disclosure avoidance rules.

Table 6 — Georgia UPK's Long-Run Impacts, by Race×Sex I

	Teen Fertility	Graduated High School	Attained BA Degree	SNAP Receipt
	Panel A. White Girls			
UPK impact	−0.014***	0.009**	0.018	−0.012
(standard error)	(0.006)	(0.005)	(0.016)	(0.009)
Impact in %	−10.9%	1.0%	4.3%	−9.2%
Counterfactual mean	0.128	0.929	0.422	0.131
Observations	1,907,000	1,028,000	738,000	724,000
	Panel B. White Boys			
UPK impact	−0.004	0.018***	0.018	−0.010
(standard error)	(.003)	(0.008)	(0.019)	(0.007)
Impact in %	−7.7%	2.0%	6.0%	−10.2%
Counterfactual mean	0.052	0.892	0.300	0.098
Observations	1,972,000	1,070,000	765,000	715,000
	Panel C. Black Girls			
UPK impact	0.008	−0.004	0.012	−0.018
(standard error)	(0.008)	(0.010)	(0.008)	(0.026)
Impact in %	4.2%	−0.4%	5.1%	−4.7%
Counterfactual mean	0.192	0.910	0.234	0.379
Observations	317,000	160,000	106,000	101,000
	Panel D. Black Boys			
UPK impact	0.000	0.007	0.011	−0.016
(standard error)	(0.007)	(0.020)	(0.014)	(0.022)
Impact in %	0.0%	0.9%	8.1%	−6.6%
Counterfactual mean	0.094	0.792	0.135	0.243
Observations	321,000	179,000	118,000	81,000

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: Author's calculations from estimating equation 2. Inference is conducted using the Ferman and Pinto (2019) method. Each "counterfactual mean" is the average of observed outcomes among individuals exposed to Georgia UPK minus the estimated treatment effect. Note that I focus on non-Hispanic White and non-Hispanic Black children because other groups have sample sizes too small for generating informative estimates. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701 and CBDRB-FY25-0013). Observation counts are rounded to comply with Census Bureau disclosure avoidance rules.

5 Childcare Substitution in Theory

If treatment is defined by pre-K *enrollment*, then the long-run impacts in the previous section are intent-to-treat (ITT) effects, which are uninformative on their own about how impacts were generated. A massive program like Georgia UPK, which enrolled more than half of all four-year-olds, can reshape the whole childcare market (Bassok, 2012; Bassok et al., 2014, 2016; Brown, 2019). In this case, it is difficult to determine whether impacts on children stem directly from program enrollees, spillovers on non-enrollees, or both. This distinction may not matter for an aggregate cost-benefit analysis, but it may for policy decisions.

Consider a policymaker in another state choosing between early childhood policy options. The specifics of how Georgia UPK generated impacts may affect their decision. If impacts were driven by children induced to enroll in UPK, policymakers may want to emulate the specific program ingredients that made Georgia UPK effective. As one counterexample, if impacts were driven by children switching from no pre-K to private pre-K—perhaps because UPK freed up private slots—then another option, such as subsidizing private pre-K, might be more appealing.

In this section, I use a simple conceptual framework to consider the role of equilibrium responses in the childcare market.[15] The framework highlights how market adjustments can cause children to substitute in and out of any given childcare option, which in turn affects their long-run outcomes. In addition to being an important mechanism to understand, the framework illustrates how spillovers impede our ability to estimate impacts on children who substitute into pre-K. After discussing these issues theoretically, I use the framework to empirically investigate substitution patterns in Georgia in Section 6.

5.1 Framework Setup

Consider a simple market with three childcare options: public pre-K (PUB), private pre-K (PRI), and no pre-K (NP). Suppose there is only one type of public pre-K and that it is rationed (i.e., demand exceeds supply). PRI includes all private pre-K programs. "No pre-K" includes informal arrangements such as care by a parent or relative and formal non-pre-K arrangements such as care by a nanny.

A public pre-K expansion occurs when the government increases the number of slots it funds. A large expansion could have important consequences for the whole childcare market. On the supply side, private pre-K providers and non-pre-K providers (e.g., nannies) may re-optimize prices and quality to remain competitive. Some providers may be forced out of business. On the demand side, changes in the market may affect households' perceptions of various types of care.[16] I will refer to these supply and demand responses collectively as "equilibrium responses" or "market adjustments."

[15]For a slightly more formal derivation of the results in this section, see Appendix E.

[16]Rothbart and Morrissey (2024) find that the expansion of public pre-K in Virginia increased enrollment in *private* pre-K. They hypothesize that the expansion changed community norms about pre-K enrollment.

Figure 5 — Substitution and Impact Matrices for an Arbitrary PUB Expansion

Substitution Matrix

Without PUB Expansion \ With PUB Expansion	NP	PRI	PUB
NP	$S^{NP \to NP}$	$S^{NP \to PRI}$	$S^{NP \to PUB}$
PRI	$S^{PRI \to NP}$	$S^{PRI \to PRI}$	$S^{PRI \to PUB}$
PUB	$S^{PUB \to NP}$	$S^{PUB \to PRI}$	$S^{PUB \to PUB}$

Impact Matrix

Without PUB Expansion \ With PUB Expansion	NP	PRI	PUB
NP	$\delta^{NP \to NP}$	$\delta^{NP \to PRI}$	$\delta^{NP \to PUB}$
PRI	$\delta^{PRI \to NP}$	$\delta^{PRI \to PRI}$	$\delta^{PRI \to PUB}$
PUB	$\delta^{PUB \to NP}$	$\delta^{PUB \to PRI}$	$\delta^{PUB \to PUB}$

Notes: Each cell in the substitution matrix represents a potential enrollment outcome with and without PUB expansion. As an example, the top right cell gives the share of children who would enroll in NP without expansion and in PUB with expansion. $S^{NP \to PUB}$ is the share of all children with these potential outcomes. Each cell in the impact matrix gives the average impact for children of a particular substitution type. For example, $\delta^{NP \to PUB}$ gives the average impact of PUB expansion among $NP \to PUB$ children.

To examine the effects of an expansion we must define potential outcomes for enrollment and human capital. Define $D_i(1)$ and $D_i(0)$ as child i's potential enrollment with and without public pre-K expansion, with $D_i(1), D_i(0) \in \{NP, PRI, PUB\}$. Define $Y_i(1)$ and $Y_i(0)$ as child i's potential human capital (e.g., educational attainment) with and without public pre-K expansion.

Households may be fully partitioned based on potential enrollment with and without a public pre-K expansion. Define $S^{A \to B} \equiv E[D_i(0) = A, D_i(1) = B]$ as the share of children induced by expansion to substitute from program A to program B. The left half of Figure 5 visualizes the nine possible households types in a substitution matrix. Each row indicates a household's potential enrollment *without* expansion, and each column indicates a household's potential enrollment *with* expansion. For example, the top right cell represents children who would enroll in no pre-K without expansion and in public pre-K with expansion.

The color coding in Figure 5 differentiates between three substitution categories. The first category, which I'll call "program stayers," is in red. Program stayers are households that would enroll in program A in either state of the world. These households either always prefer A or prefer public pre-K but are not awarded slots. In the local average treatment effect (LATE) framework, these households are always- and never-takers. The second category, "substitution into public," is in green and includes households induced by the expansion to enroll in public pre-K. In the LATE framework, if treatment is public pre-K enrollment, then these are compliers.

The third substitution category in Figure 5, "equilibrium-driven substitution," is in blue. In the language of the LATE framework, these households are defiers. Without equilibrium responses the population would consist entirely of program stayers and substitution into PUB; the red and green cells would sum to one and the remaining cells would be empty. Intuitively, the blue cells would violate the independence of irrelevant alternatives. Using $PRI \to NP$ as an example, expanding public pre-K should not cause any household that would have enrolled in private pre-K to enroll in no pre-K if prices, quality, and demand-side factors stay constant.

However, with equilibrium responses, children may substitute between programs in any direction. Independence of irrelevant alternatives no longer holds because changes in prices, quality, or other factors can make any childcare option "relevant."

One final note is that I have so far ignored the possibility that private pre-K programs could also be rationed. In theory, prices may equate supply and demand. In reality, however, private programs may be rationed. This provides an additional explanation for substitution from no pre-K and public pre-K into private pre-K. If expansion causes children to substitute from private pre-K to public pre-K, then other children may fill the freed-up private spots. This type of substitution into private pre-K can occur even without other equilibrium responses in the childcare market. However, because it ultimately has a similar spillover effect, I will continue to use "equilibrium responses" as an umbrella term.

5.2 Impacts of Expanding Public Pre-K

The matrix on the right side of Figure 5 depicts the impact of public pre-K expansion on human capital for each substitution type. Define $\delta^{A\rightarrow B} \equiv E[Y_i(1) - Y_i(0)|D_i(0) = A, D_i(1) = B]$ as the average impact among children who substitute from A to B. Note that these impacts are not primitives or policy-invariant; each $\delta^{A\rightarrow B}$ is determined by the set of $A \rightarrow B$ children and by other market and program characteristics. A different public pre-K expansion, even in the same place, would generally produce different values for each $S^{A\rightarrow B}$ and $\delta^{A\rightarrow B}$.

The overall effect of a public pre-K expansion is $\text{ITT} \equiv E[Y_i(1) - Y_i(0)]$. In the simplest case with no equilibrium responses—and no peer effects—the only children affected by public pre-K expansion are those induced to enroll in it.[17] Intuitively, there is no equilibrium-driven substitution and no mechanism for impacts on program stayers. Defining $\Delta^{A\rightarrow B} \equiv S^{A\rightarrow B} \times \delta^{A\rightarrow B}$ as the *population-level* impact of expansion for $A \rightarrow B$ children, the ITT effect is:

$$\text{ITT} = E[Y_i(1) - Y_i(0)] = \underbrace{\Delta^{PRI\rightarrow PUB} + \Delta^{NP\rightarrow PUB}}_{\text{Substitution into PUB}}. \tag{3}$$

In contrast, in the general case with equilibrium responses, the ITT effect is:

$$\begin{aligned}\text{ITT} = E[Y_i(1) - Y_i(0)] = {} & \underbrace{\Delta^{PRI\rightarrow PUB} + \Delta^{NP\rightarrow PUB}}_{\text{Substitution into PUB}} \\ & + \underbrace{\Delta^{PRI\rightarrow PRI} + \Delta^{NP\rightarrow NP} + \Delta^{PUB\rightarrow PUB}}_{\text{Program stayers}} \\ & + \underbrace{\Delta^{PUB\rightarrow PRI} + \Delta^{PUB\rightarrow NP} + \Delta^{PRI\rightarrow NP} + \Delta^{NP\rightarrow PRI}}_{\text{Equilibrium-driven substitution}}.\end{aligned} \tag{4}$$

[17] Peer effects are a mechanism outside of market adjustments and childcare substitution in which public pre-K expansion may affect all four-year-olds (Neidell and Waldfogel, 2010; Carrell et al., 2018). Non-enrollees may be affected if they interact with public pre-K compliers outside of pre-K or in school in subsequent years. This framework does not preclude the existence of peer effects, but it is not a focus. The impact estimates in Section 4 implicitly capture peer effects and other mechanisms not highlighted in this framework.

Any substitution group can experience impacts in this case. Clearly, children who "switch" programs may be affected. Perhaps less clear is that program stayers may also be affected. For instance, non-public childcare providers may respond to public expansion by changing quality levels. Henry and Gordon (2006) find that increased competition from Georgia UPK made pre-K programs more productive (in terms of raising test scores). Conversely, pre-K expansions typically increase demand for pre-K teachers, which could raise costs (teacher wages) for all programs and/or force them to hire less qualified teachers (Borowsky et al., 2022). In general, program stayers could experience positive or negative impacts.

For good reason, evaluations of pre-K programs with individual-level randomization receive great attention (Bruhn and Emick, 2023; Burchinal et al., 2024). However, equation 4 highlights a limitation of evaluations, such as lottery studies, that estimate effects only for program compliers. These evaluations miss spillover effects on non-compliers when there are equilibrium responses in the childcare market. Depending on the sign of spillovers, the direct effect on compliers could overstate or understate overall program impact, painting an incomplete picture of a program's full costs and benefits.

5.3 Treatment Effects on the Treated

For some research questions, the relevant parameter is not the ITT effect but rather the treatment effect on the treated (TOT). In settings with aggregate-level variation and population data, however, the ITT is the more easily identified parameter. It is thus common in the DiD pre-K literature to approximate TOT effects with a Wald estimator by dividing ITT effects by changes in pre-K enrollment.

Scaling an ITT estimate requires assumptions about who is affected by a public pre-K expansion and which programs children substitute between. This idea is not new, but across the pre-K literature the required assumptions are often underrecognized.[18] Here I lay out precisely what is estimated when the (often strong) assumptions are not met. In Section 6 I attempt to quantify the bias by estimating a complete substitution matrix.

Equation 4 illustrates that in the presence of equilibrium responses it is not clear how to define "treated" or whether anyone is "untreated". Given this ambiguity, different papers use different definitions of treatment. One common definition is "children induced to enroll in public pre-K." Thus, many papers estimate TOT effects by dividing ITT effects by the net change in public enrollment (Cascio and Schanzenbach, 2013; Cascio, 2023; Kose, 2023; Thompson, 2018). In the simplest case with no equilibrium responses (or peer effects), this calculation recovers the true TOT effect:

$$\frac{\text{ITT}}{\text{Net Change in PUB}} = \frac{S^{NP\to PUB}\delta^{NP\to PUB} + S^{PRI\to PUB}\delta^{PRI\to PUB}}{S^{NP\to PUB} + S^{PRI\to PUB}}. \tag{5}$$

[18] A notable exception is Zerpa (2022), who explicitly enumerates the assumptions required in her setting to identify or bound TOT effects.

However, in many cases it may be unreasonable to assume that there are no equilibrium responses. In the most general case, with equilibrium responses, this TOT calculation yields the following:

$$\frac{\text{ITT}}{\text{Net Change in PUB}} = \frac{\sum_{A\in\mathcal{J}}\sum_{B\in\mathcal{J}} S^{A\to B} \times \delta^{A\to B}}{S^{NP\to PUB} + S^{PRI\to PUB} - S^{PUB\to NP} - S^{PUB\to PRI}} \tag{6}$$

where $\mathcal{J} \equiv \{NP, PRI, PUB\}$. This expression is difficult to interpret. The numerator is driven by all nine substitution types. Scaling misattributes economy-wide effects to only a subset of children. Moreover, if equilibrium responses cause some children to switch *out* of public pre-K, then the denominator does not represent public pre-K compliers as intended. In the LATE framework, this is the issue of the existence of defiers. Why might children substitute out of public pre-K? One reason is private pre-K rationing; some children may substitute from public to private if the expansion frees up private spots.

Generally, the interpretability of TOT estimates calculated this way hinges on the plausibility of a "no equilibrium response" assumption. In some cases it may be reasonable to assume equilibrium effects are small or non-existent, but in other cases—particularly for large expansions—that assumption may be unrealistic.

A second approach in the literature is to define "treated" as children induced to enroll in *any type* of pre-K (combining public and private). Some papers estimate TOT effects for this group by scaling ITT effects by the net change in (any) pre-K enrollment (Bailey et al., 2021; Baker et al., 2008; Monnet, 2019). One motivation for this approach is the idea that switching from private to public pre-K should not be considered a switch from "untreated" to "treated." Another motivation is that this approach captures substitution from no pre-K to private pre-K due to private rationing.

In the general case with childcare market adjustments, this TOT calculation faces the same issues as equation 6. The numerator may contain impacts from all nine substitution types, and the denominator may not capture the intended treated population. Furthermore, this definition of treatment has an issue in the denominator that is not solved by assuming away equilibrium responses. Observe that the net change in any pre-K is:

$$S^{NP\to PUB} + S^{NP\to PRI} - S^{PUB\to NP} - S^{PRI\to NP}. \tag{7}$$

Substitution between private and public—in either direction—does not appear in this expression (the TOT denominator) because it does not contribute to net changes in any pre-K enrollment. But it would appear in the TOT numerator if it contributes to the ITT effect. Thus, this TOT calculation requires that children do not shift between different types of pre-K or that those shifts have no impact. These assumptions may be valid for a marginal pre-K expansion but are less realistic for large ones.

6 Childcare Substitution in Georgia

6.1 Childcare Substitution

Whether Georgia UPK led to equilibrium-driven substitution is an empirical question. In this section, I deepen our understanding of Georgia UPK's long-run impacts by estimating its full substitution matrix. I find that UPK generated meaningful substitution between pre-K arrangements in all directions. This suggests that 1) focusing only on enrollees could miss an important part of UPK's impact, and 2) conventional TOT estimates would likely be biased.

Identifying Georgia UPK's substitution matrix is a standard causal inference problem. For children exposed to UPK, we observe enrollment in a world with UPK, $D_i(1)$, but not in a world without UPK, $D_i(0)$. Estimating the full substitution matrix requires estimating the share of children observed in program B who *would have* enrolled in program A if not for UPK. With those shares it is straightforward to compute each $S^{A \to B}$ in the substitution matrix.

As in the earlier enrollment analysis, I draw on data from the long-form 2000 census, in which I observe pre-K enrollment for the 1999 pre-K cohort. I again use the 18 states that had no state-funded pre-K program in 1999 as a control group. Unlike in the earlier analysis, identification requires a different assumption than parallel trends, which identifies *net* enrollment changes but not underlying inflows and outflows. Instead, I make a conditional independence assumption: I assume observationally equivalent children in Georgia and the control group would enroll in the same type of pre-K, on average, in a world without UPK.[19] This is a strong assumption, but it is strengthened by the rich set of predictive variables in the census.

To estimate substitution, I first fit a multinomial logit model, using only the control group, with enrollment in no pre-K, private pre-K, and public pre-K as outcomes. The predictor variables include a large set of children's individual and household characteristics, as well as county variables that are predictive of one's local pre-K market (including county-level pre-K enrollment rates before UPK). Then, I use the logit estimates to obtain predicted values for each child in Georgia. Summing the predicted probability of enrolling in A across all children enrolled in B, and normalizing by the population, gives the final estimate for $S^{A \to B}$. For more details on the analysis, see Appendix F.

The conditional independence assumption necessary for a causal interpretation cannot be tested directly, meaning the results should be interpreted with caution. However, I do find that the logit model is highly predictive of enrollment in a world without UPK. In Appendix F I use Southeastern states in the control group as a testing ground, on the premise that they most closely represent Georgia's counterfactual childcare market. I first estimate the logit model on all other control states, and then I predict enrollment in the Southeastern states. The predictions are highly accurate, on average (see Appendix Table F.1).

A second piece of evidence for the logit model's predictiveness is that the aggregate coun-

[19]Formally, I assume $E[D_i(0) = A | X_i, GA_i = 0] = E[D_i(0) = A | X_i, D_i(1) = B, GA_i = 1]$ for all $A, B \in \{NP, PRI, PUB\}$, where X_i is a vector of covariates and GA_i indicates residing in Georgia.

Figure 6 — Georgia UPK Substitution Estimates

		With UPK			
		NP	**PRI**	**PUB**	
Without UPK	**NP**	**8%**	**5%**	**15%**	28%
	PRI	**3%**	**11%**	**14%**	28%
	PUB	**6%**	**8%**	**30%**	44%
		17%	24%	59%	

Notes: This figure depicts estimates of substitution between pre-K options. For example, the top right cell gives the estimated share of children in Georgia who substituted from no pre-K to public pre-K because of UPK. The numbers in the margins are row and column sums. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701).

terfactual enrollment rates it predicts for Georgia are very close to the counterfactual enrollment rates implied by the DiD estimates in Section 3. I obtain the latter by subtracting the DiD estimates from the observed enrollment rates below the matrix in Figure 6. Across analyses, the counterfactual enrollment rates are 28% vs. 27.7% for no pre-K, 28% vs. 26.2% for private pre-K, and 44% vs. 46.1% for public pre-K.

The substitution estimates are shown in Figure 6. Nearly half (49%) of all four-year-olds are program stayers (red cells) and 29% substitute from no pre-K or private pre-K into public pre-K (green cells). The remaining 22% represent "equilibrium-driven substitution" (blue cells), or substitution that occurs due to equilibrium market adjustments. The size of this group suggests that focusing only on program enrollees could miss a substantial portion of UPK's overall impact.

The largest single cell is the 30% of children who enroll in public pre-K with and without UPK. This estimate is driven by localities in the control group creating public programs in the absence of a state program. Therefore, in this setting, some public program stayers substitute between local public programs, Head Start, and UPK. Differences between those programs provide a new reason this substitution group may experience non-zero impacts ($\delta^{PUB \to PUB} \neq 0$).

Figure 6 reveals that Georgia UPK had unintended consequences: some children who would have enrolled in some form of pre-K were induced to not enroll at all. 6% of all children substituted from public pre-K to no pre-K, and 3% substituted from private pre-K to no pre-K. These substitutions are most easily explained by UPK forcing some programs to close. I find evidence consistent with that explanation by examining Head Start. Appendix H shows with an event study that UPK reduced Head Start enrollment by 5.3pp (36.4%) for the 1999 cohort. If some Head Start centers were forced to close, some would-be Head Start families may not have wanted or may not have been awarded spots in UPK.

Returning to the short-run enrollment analysis in Section 3, the substitution estimates add important context to the estimates of *net* changes in pre-K enrollment. The DiD estimate was that UPK reduced private pre-K enrollment by 3.2pp on net. Without market responses, UPK's only effect on enrollment would be shifting children from alternative childcare options into UPK. The DiD estimate would then imply that private pre-K was the counterfactual option for only 6% of UPK enrollees (3.2/52.3). That share is improbably small in light of analogous causal estimates from other UPK programs, hinting at the existence of equilibrium-driven substitution (Humphries et al., 2024; Weiland et al., 2020).

Indeed, the more granular substitution estimates reveal that private pre-K was the counterfactual enrollment option for 23.7% (14/59) of public pre-K enrollees. This estimate is much more in line with causal estimates from other contexts. Further, 50.8% of children enrolled in public pre-K would have enrolled in public pre-K even without UPK, and the remaining 25.4% would not have enrolled in any pre-K.[20]

What then explains the small net reduction in private pre-K enrollment? Figure 6 shows that the net change conceals offsetting inflows and outflows. 17% of children would have enrolled in private pre-K but instead enrolled in no pre-K (3%) or public pre-K (14%). On the other side of the ledger, 13% of children were induced by market adjustments to enroll in private pre-K, filling most of the vacated spots. Due to a lack of data on childcare centers in the 1990s, I can only speculate about the causes of those flows. The outflows from private to public—which constitute 82% of the outflows—are easily explained by families preferring the free UPK option. The inflows are consistent with UPK freeing up spots in rationed private programs, or with price or quality adjustments attracting new families into private programs. Either way, this result highlights that net changes can be misleading for inferences about deeper substitution patterns.

6.2 Substitution-Specific Impacts

We would ideally like to estimate Georgia UPK's full impact matrix too. Unfortunately, impacts on specific substitution groups are not identified without significantly more structure. That said, the substitution and ITT impact estimates in this paper do impose useful constraints. For a given long-run outcome, a weighted sum of the sub-impacts—with substitution shares as weights—must equal the ITT effect. This narrows the set of possible sub-impacts considerably.

To explore the range of possible impact matrices, I construct six plausible scenarios. Figure 7 depicts the first and the third. The full six are in Appendix G. In each scenario, the sub-impacts aggregate to 1.5pp, my estimate of UPK's overall effect on high school graduation. The scenarios are hypothetical, but they illustrate some general takeaways.

I defer a full discussion of each scenario's construction to Appendix G. Briefly, however, in each scenario I make the simplifying assumption that the off-diagonal cells in the impact matrix are symmetric but opposite in sign. In other words, if moving from A to B increases a child's likelihood of graduating high school by Xpp, then moving from B to A decreases their likelihood by Xpp. Then, in each scenario I make assumptions about whether private or public pre-K is more effective and whether children who do not switch programs experience impacts. Scenario

[20]These estimates are for all public pre-K enrollees rather than UPK enrollees specifically, but UPK comprised around 90% of public pre-K enrollment in Georgia in 1999.

Figure 7 — Hypothetical Impact Matrices for High School Graduation Impacts

Scenario I

Without UPK \ With UPK	NP	PRI	PUB
NP	0.0pp	13.6pp	13.6pp
PRI	-13.6pp	0.0pp	0.0pp
PUB	-13.6pp	0.0pp	0.0pp

Scenario III

Without UPK \ With UPK	NP	PRI	PUB
NP	0.0pp	9.3pp	7.7pp
PRI	-9.3pp	1.0pp	-1.5pp
PUB	-7.7pp	1.5pp	2.0pp

	Scenarios					
	I	II	III	IV	V	VI
More effective	=	=	*PRI*	*PRI*	*PUB*	=
Program stayer effects	0	+	+	0	+	−

Notes: This figure depicts two hypothetical scenarios that are consistent with my estimates for UPK's ITT effect on high school graduation and UPK's substitution matrix. In each matrix, each cell gives the average impact for children of a particular substitution type. For example, the top right cell of each matrix gives the average impact for children in Georgia who substituted from no pre-K to public pre-K because of UPK. In the table, "more effective" indicates whether private or public pre-K is assumed to have larger impacts, or whether they are assumed to be equal. The "program stayer effects" row indicates whether the diagonal cells in the impact matrix are positive, negative, or zero. See Appendix G for the exact impact matrix in each scenario. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701).

I is a benchmark case in which (i) program stayers do not experience any impact, and (ii) private and public pre-K are equally effective. Given these assumptions, the entire matrix is identified, and one can back out that the non-zero impacts are all 13.6pp in magnitude.

However, prior research suggests Scenario I may be unrealistic. In some scenarios, I assume positive impacts for private and public pre-K stayers, reflecting evidence from Henry and Gordon (2006) that competition from UPK made pre-K programs in Georgia more effective. Positive impacts for public pre-K stayers also reflect evidence that UPK had greater short-run impacts on children in Georgia than Head Start (Henry et al., 2006). Differential effectiveness for private and public reflect large differences in quality that have been documented across programs (Chaudry et al., 2021). The table at the bottom of Figure 7 summarizes the assumptions in each scenario.

The first takeaway from this exercise is that a wide range of sub-impacts is consistent with the aggregate ITT impact. To give one example, in this small set of scenarios not designed to maximize variation, the impact of substitution from no pre-K to any pre-K ranges from 6.5pp to 17.4pp. The second takeaway is that we cannot rule out that equilibrium-driven impacts were a meaningful part of Georgia UPK's overall impact. In Scenario III, for example, those who substitute from no pre-K or private pre-K to public pre-K raise the population-level high school graduation rate by only 0.9pp. The ITT effect reaches 1.5pp because of net positive impacts experienced in the rest of the market.

Table 7 — Six Hypothetical Scenarios of High School Graduation Impacts

	TOT Estimate (Scaled ITT)	"True" TOT Effect by Scenario					
		I	II	III	IV	V	VI
Public pre-K compliers	10.0	7.1	3.7	3.3	6.2	4.0	9.0
Any pre-K compliers	13.6	13.6	7.2	8.1	15.4	6.5	17.4
More effective	N/A	=	=	*PRI*	*PRI*	*PUB*	=
Program stayer effects	N/A	0	+	+	0	+	−

Notes: All impacts are in percentage points. Public pre-K compliers are those induced by UPK to substitute from no pre-K or private pre-K to public pre-K. Any pre-K compliers are those induced by UPK to substitute from no pre-K to private pre-K or public pre-K. "More effective" indicates whether private or public pre-K is assumed to have larger impacts, or whether they are assumed to be equal. The "program stayer effects" row indicates whether the diagonal cells in the impact matrix are positive, negative, or zero. See Appendix G for the exact impact matrix in each scenario. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701).

The next question, then, is whether TOT estimators commonly used in the DiD literature produce reasonable estimates (see Section 5.3). I investigate this question in Table 7 using the six hypothetical impact scenarios. Taking my substitution estimates and the hypothetical impacts as given, I compute the "true" average impact for two groups: 1) public pre-K compliers (children who substitute from no pre-K or private pre-K to public pre-K), and 2) any pre-K compliers (children who substitute from no pre-K to any pre-K). Then, I perform the type of calculation common in the reduced form pre-K literature. I divide the ITT estimate by: 1) the net change in public pre-K enrollment, or 2) the net change in any pre-K enrollment.

In Scenario I, the scaled ITT (10.0pp) overestimates the "true" impact for public pre-K compliers (7.1pp). Intuitively, the upward bias comes from misattributing system-wide impacts to too small a group. Public pre-K compliers are 29% of the population, but the net change in public enrollment is only 15pp. On the other hand, the scaled ITT perfectly obtains the "true" impact for any pre-K compliers (13.6pp). This is because Scenario I meets the conditions laid out in Section 5.3 that rationalize this estimator: no impacts for children who substitute between private and public pre-K or for program stayers.

As I relax these assumptions, the bias tends to grow. The scaled ITTs range from 1.1 to 3.0 times the "true" impact for public pre-K compliers and 0.8 to 2.1 times the "true" impact for private pre-K compliers. Some patterns emerge. Positive impacts on program stayers are especially problematic. As Scenario II shows, positive impacts on program stayers require that the other sub-impacts fall (in magnitude) to keep the ITT effect constant. Thus, the true TOT effects become smaller, and the scaled ITTs become more upwardly biased. By comparison, assuming that private and public pre-K have different impacts has a smaller effect on the bias, as shown in Scenarios III, IV, and V.

Scenario VI assumes program stayers experience small negative impacts, which may happen, for example, if greater pre-K enrollment throughout the market forces centers to hire less qualified teachers to fill spots. Holding the ITT effect constant, negative impacts on program stayers require larger impacts in the rest of the matrix. In this example, the impact on private pre-K compliers grows by enough to make the scaled ITT a slight *underestimate*. Conversely, the upward bias for public pre-K compliers is not fully offset, but it does come close. Further reducing the upward bias would require larger negative impacts on program stayers, which would then require implausibly large magnitudes for the other sub-impacts.

Another result is that the upward bias is greater for public pre-K compliers than for any pre-K compliers in most scenarios. This is driven by what is defined as a direct effect versus a spillover, because spillovers bias scaled ITTs. When public pre-K compliers are the group of interest, impacts from switching from no pre-K to private pre-K ($\Delta^{NP\rightarrow PRI}$) are spillovers. When any pre-K compliers are the group of interest, impacts from switching from private pre-K to public pre-K ($\Delta^{PRI\rightarrow PUB}$) are spillovers. If the former exceed the latter ($\Delta^{NP\rightarrow PRI} > \Delta^{PRI\rightarrow PUB}$), which is likely true in many settings, spillovers will be greater when treatment is defined as public pre-K. Thus, the public pre-K TOT estimate will likely have a greater upward bias, matching the intuition for defining treatment as any pre-K.

To summarize, it is difficult to draw firm conclusions about sub-impacts in Georgia. Many plausible scenarios are consistent with my aggregate estimates. That said, given the evidence in Georgia for positive impacts on program stayers, the average impacts for public pre-K compliers and any pre-K compliers are likely less than the scaled ITTs (10.0pp and 13.6pp).

What can we conclude about the broader literature? In settings with the potential for equilibrium market adjustments, scaled ITT estimates may be substantially biased. Further, scaled ITTs will tend to be overestimates (in absolute value) for program compliers unless there are large impacts in the opposite direction for non-compliers. Of course, the amount of substitution between childcare types and the size/direction of impacts on non-compliers is highly context dependent. To determine whether equilibrium responses are an important mechanism, and whether they impede TOT estimation, researchers studying large pre-K expansions should investigate impacts on the childcare market and on granular substitution patterns.

7 Cost-Benefit Analysis

Before concluding, I use my estimates from the rest of the paper to conduct a cost-benefit analysis of the introduction of Georgia UPK. I do so using the marginal value of public funds (MVPF) approach. Following Hendren (2016), the MVPF is defined simply as the ratio between a policy's marginal benefits and its net cost to government, where the net cost includes direct costs and any fiscal externalities the policy generates.

Building on Kline and Walters (2016) and Cascio (2023), I estimate marginal benefits using long-run impacts on earnings and government benefit receipt, as well as short-run impacts on childcare expenditures. I depart from their analyses, though, by using ITT estimates rather than TOT estimates. As my conceptual framework highlights, for a large pre-K expansion with equilibrium responses, a complete accounting of costs and benefits must include the direct effects on program compliers *and* indirect effects on non-compliers (Finkelstein and Hendren, 2020).

Intuitively, then, ITT effects are the relevant parameters for this analysis.[21]

I write the MVPF of the introduction of Georgia UPK as:

$$\text{MVPF} = \frac{\rho(1-\tau)\text{ITT}^{Y} + \rho\varphi^{\text{SNAP}}\text{ITT}^{\text{SNAP}} - \kappa^{\text{PRI}}\text{NC}^{\text{PRI}}}{\varphi^{\text{PUB}}\text{NC}^{\text{PUB}} + \varphi^{\text{HS}}\text{NC}^{\text{HS}} - \rho\tau\text{ITT}^{Y} + \rho\varphi^{\text{SNAP}}\text{ITT}^{\text{SNAP}}}. \tag{8}$$

Starting with the numerator, the first term is the present discounted value of UPK's effect on long-run earnings. ITT^{Y} is UPK's effect on earnings, τ is a marginal tax rate, and ρ is a discount factor. The second term is the (discounted) loss in SNAP benefits children receive as adults. ITT^{SNAP} is UPK's effect on the share of children who receive benefits, and φ^{SNAP} is the average benefit amount. The final term is savings to families from reduced enrollment in private pre-K. NC^{PRI} is the net change in private pre-K enrollment, and κ^{PRI} is average household spending on private pre-K.

The first two terms in the denominator represent UPK's effect on government pre-K spending. The NC terms give the net change in enrollment in non-Head Start public pre-K (PUB) and in Head Start (HS). The φ terms give the average cost (to government) of each program. I separate Head Start from other public pre-K because it is notably more expensive. I assume UPK and local public programs cost the same. I do not include cost savings from reduced spending on private pre-K subsidies due to a lack of data on subsidization rates. This makes the MVPF estimate conservative. Lastly, $\rho\tau\text{ITT}^{Y}$ and $\rho\varphi^{\text{SNAP}}\text{ITT}^{\text{SNAP}}$ represent long-run effects on government tax receipt and benefit spending.[22] For more information on my exact calibration, see Appendix I. Where possible, I closely follow Cascio (2023).

A few details should be noted. First, equation 8 does not require information about the underlying substitution and substitution-specific impacts in Georgia. The estimates from this paper that I plug in are all well-identified aggregate DiD estimates. Second, my estimates for earnings and SNAP receipt are only through age 32, the oldest age of treated children in my sample. Longer-lasting impacts on earnings would make my MVPF estimate a lower bound. The implications of longer-lasting SNAP effects are less clear because they reduce both benefits and costs. Third, this MVPF calculation does not account for many other outcomes pre-K has been shown to affect. Most notably, previous analyses have demonstrated large cost savings to government from reductions in crime, not to mention the harder-to-quantify social benefits of less crime (Belfield et al., 2006).

Taking the point estimates in this paper as given, the introduction of Georgia UPK has marginal benefits to individuals worth $\$1,722$ and a net cost to government of $-\$154$. Strikingly, the negative net cost suggests UPK fully paid for itself by raising tax revenue and lowering spending on SNAP and Head Start. Hendren and Sprung-Keyser (2020) define the MVPF as

[21] In the context of an experiment that held the broader childcare market constant, Kline and Walters (2016) show that TOT effects are the relevant parameters for welfare analysis.

[22] Note that there is no strong evidence that Georgia UPK raised parents' earnings on average. Positive impacts on parents' earnings would raise program benefits and lower costs via additional tax collections. In this case, my MVPF estimate would be overly conservative.

"infinite" in this case. Researchers have estimated infinite MVPFs for other early childhood interventions too, such as Head Start and Medicaid expansion to pregnant women and infants (Hendren and Sprung-Keyser, 2020).

However, the MVPF estimate is highly sensitive to the long-run earnings parameter, which I estimate imprecisely. Recall that my primary estimate is that UPK exposure raised annual earnings by $\$1,006$ (SE $= 644$). How would the MVPF estimate change if the true earnings impact were smaller? UPK continues to have an infinite MVPF as long as the earnings impact exceeds \$637. Further, the MVPF exceeds one (meanings benefits exceed costs) as long as the earnings impact is greater than \$102. Thus, although I cannot obtain a precise estimate, it appears likely that the pecuniary benefits of introducing Georgia UPK exceeded the costs.

8 Conclusion

Universal pre-K has become increasingly common in states and cities across the U.S., and it continues to be a popular policy proposal at all levels of government. Because UPK programs are relatively young, very little is known about their long-run impacts. This paper addresses that gap by estimating the long-run effects of Georgia's first-in-the-nation statewide UPK program.

I show that exposure to Georgia UPK made children less likely to have children as teenagers, more likely to graduate from high school, and less likely to receive SNAP benefits as adults. Impacts on adult earnings are imprecise, but the point estimates suggest there may have been positive average effects. I find only modest impact heterogeneity across counties in Georgia, suggesting that UPK was similarly effective throughout the state. Importantly, this includes rural areas where the provision of pre-K has historically been more challenging.

My results echo previous studies of Georgia UPK, which find positive effects through fourth and eighth grade. Notably, our studies find consistent evidence of positive effects across multiple datasets. More broadly, estimates of UPK's long-run effects in Boston, Tulsa, and now Georgia all demonstrate that UPK can improve children's life trajectories.

This paper also considers the role of equilibrium market adjustments in response to public pre-K expansions. My estimates suggest market forces in Georgia caused children to substitute between no pre-K, private pre-K, and public pre-K in all directions following the introduction of UPK—consistent with market forces playing a large role in UPK's overall effects.

Recent papers outside of Georgia have also found evidence of equilibrium responses to public pre-K. Rothbart and Morrissey (2024) find that public pre-K expansion in Virginia caused an increase in *private* pre-K enrollment. They hypothesize that the public expansion changed community norms about pre-K attendance. In Michigan, Berne et al. (2024b) find evidence that the introduction of Transitional Kindergarten drew some children away from the state's income-targeted pre-K program, but that those spots were filled by other children.

Unfortunately, equilibrium responses in the childcare market present challenges for estimating TOT effects when treatment is defined by enrollment. I show empirically that the standard approach in the DiD pre-K literature may yield substantially biased estimates, with the bias depending on the size and direction of equilibrium-driven substitution and spillover impacts on non-enrollees. In the Georgia setting, scaling ITT effects by net changes in pre-K enrollment would likely *overestimate* the true effects on pre-K compliers.

Despite this limitation, population ITT effects are often the parameter of interest for policymakers. For instance, a cost-benefit analysis of a large pre-K expansion should account for impacts on program compliers and equilibrium-driven impacts on non-compliers—both of which are captured by ITT effects. This argues for more quasi-experimental evaluations of large program expansions and for cluster-randomized controlled trials that randomize at the level of a childcare market. Studies with individual-level randomization can capture many elements of a scaled-up program, but equilibrium market responses are difficult to replicate (Gupta et al., 2021). DiD analyses with population data require stronger assumptions, but they succinctly capture impacts on compliers and non-compliers. This type of holistic evidence is critical as policymakers decide whether to expand public pre-K.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller, "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association*, 2010, *105* (490), 493–505.

Anders, John, Andrew C. Barr, and Alexander A. Smith, "The Effect of Early Childhood Education on Adult Criminality: Evidence from the 1960s through 1990s," *American Economic Journal: Economic Policy*, 2023, *15* (1), 37–69.

Bailey, Martha, Shuqiao Sun, and Brenden Timpe, "Prep School for Poor Kids: The Long-Run Impacts of Head Start on Human Capital and Economic Self-Sufficiency," *American Economic Review*, 2021, *111* (12), 3963–4001.

Baker, Michael, Jonathan Gruber, and Kevin Milligan, "Universal Child Care, Maternal Labor Supply, and Family Well-Being," *Journal of Political Economy*, 2008, *116* (4), 709–745.

Bassok, Daphna, "Competition or Collaboration? Head Start Enrollment During the Rapid Expansion of State Pre-kindergarten," *Educational Policy*, 2012, *26* (1), 96–116.

__ , Luke C. Miller, and Eva Galdo, "The effects of universal state pre-kindergarten on the child care sector: The case of Florida's voluntary pre-kindergarten program," *Economics of Education Review*, 2016, *53*, 87–98.

__ , Maria Fitzpatrick, and Susanna Loeb, "Does State Preschool Crowd-Out Private Provision? The Impact of Universal Preschool on the Childcare Sector in Oklahoma and Georgia," *Journal of Urban Economics*, 2014, *83*, 18–33.

Belfield, Clive, Milagros Nores, Steve Barnett, and Lawrence Schweinhart, "The High/Scope Perry Preschool Program: Cost-Benefit Analysis Using Data from the Age-40 Followup," *Journal of Human Resources*, 2006, *41* (1), 162–190.

Berlinski, Samuel, Maria Marta Ferreyra, Luca Flabbi, and Juan David Martin, "Childcare Markets, Parental Labor Supply, and Child Development," *Journal of Political Economy*, 2024, *132* (6), 2113–2177.

Berne, Jordan, Brian Jacob, Tareena Musaddiq, Anna Shapiro, and Christina Weiland, "The Effect of Early Childhood Programs on Third-Grade Test Scores: Evidence from Transitional Kindergarten in Michigan," *AEA Papers and Proceedings*, 2024, *114*, 480–485.

Berne, Jordan S., Katia Cordoba Garcia, Brian A. Jacob, Tareena Musaddiq, Samuel Owusu, Anna Shapiro, and Christina Weiland, "Transitional Kindergarten: The New Kid on the Early Learning Block," Technical Report EdWorkingPaper: 24-921, Annenberg Institute at Brown University 2024.

Bodéré, Pierre, "Dynamic Spatial Competition in Early Education: an Equilibrium Analysis of the Preschool Market in Pennsylvania," Technical Report Working Paper November 2023.

Borowsky, Jonathan, Jessica H. Brown, Elizabeth E. Davis, Chloe Gibbs, Chris M. Herbst, Aaron Sojourner, Erdal Tekin, and Matthew J. Wiswall, "An Equilibrium Model of the Impact of Increased Public Investment in Early Childhood Education," Technical Report Working Paper 30140, National Bureau of Economic Research, Cambridge, MA June 2022.

Brackett, Margret, Gary Henry, and Jeanie Weathersby, "Report on the Expenditure of Lottery Funds Fiscal Year 1999," Technical Report, The Council for School Performance, GA 1999.

Brown, Jessica H., "Does Public Pre-K Have Unintended Consequences on the Child Care Market for Infants and Toddlers?," Technical Report, Working Paper December 2019.

Bruhn, Jesse and Emily Emick, "Lottery evidence on the impact of preschool in the United States: A review and meta-analysis," Technical Report 2023.20, Blueprint Labs 2023.

Bryan, Kris and Gary Henry, "Quality of Georgia's Pre-Kindergarten Program: 1997-98 School Year," Technical Report, Council for School Performance, Atlanta, GA October 1998.

Bryant, Donna, Margaret Burchinal, Lisa Lau, and Joseph Sparling, "Family and Classroom Correlates of Head Start Children's Developmental Outcomes," *Early Childhood Research Quarterly*, 1994, *9*, 289–309.

Burchinal, Margaret, Anamarie Whitaker, Jade Jenkins, Drew Bailey, Tyler Watts, Greg Duncan, and Emma Hart, "Unsettled science on longer-run effects of early education," *Science*, 2024, *384* (6695), 506–508.

Carrell, Scott E., Mark Hoekstra, and Elira Kuka, "The Long-Run Effects of Disruptive Peers," *American Economic Review*, 2018, *108* (11), 3377–3415.

Cascio, Elizabeth, "Early Childhood Education in the United States: What, When, Where, Who, How, and Why," Technical Report Working Paper 28722, National Bureau of Economic Research, Cambridge, MA April 2021.

—, "Does Universal Preschool Hit the Target? Program Access and Preschool Impacts," *Journal of Human Resources*, 2023, *58* (1), 1–42.

— **and Diane Schanzenbach**, "The Impacts of Expanding Access to High-Quality Preschool Education," *Brookings Papers on Economic Activity*, 2013, *44* (2), 127–178.

Chaudry, Ajay, Taryn Morrissey, Christina Weiland, and Hirokazu Yoshikawa, *Cradle to kindergarten: A new plan to combat inequality*, New York, NY: Russell Sage Foundation, 2021.

Chernozhukov, Victor, Kaspar Wüthrich, and Yinchu Zhu, "An exact and robust conformal inference method for counterfactual and synthetic controls," *Journal of the American Statistical Association*, 2021, *116* (536), 1849–1864.

Conley, Timothy and Christopher Taber, "Inference with "Difference in Differences" with a Small Number of Policy Changes," *The Review of Economics and Statistics*, 2011, *93* (1), 113–125.

Currie, Janet and Douglas Almond, "Human capital development before age five," in David Card and Orley Ashenfelter, eds., *Handbook of Labor Economics*, Vol. 4, Elsevier, 2011, pp. 1315–1486.

Deming, David, "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start," *American Economic Journal: Applied Economics*, 2009, *92* (4), 111–134.

Doudchenko, Nikolay and Guido Imbens, "Balancing, Regression, Difference-In-Differences and Synthetic Control Methods: A Synthesis," Technical Report Working Paper 22791, National Bureau of Economic Research, Cambridge, MA October 2016.

Duncan, Greg J. and Katherine Magnuson, "The Nature and Impact of Early Achievement Skills, Attention Skills, and Behavior Problems," in Greg J. Duncan and Richard J. Murnane, eds., *Whither opportunity?: Rising inequality, schools, and children's life chances*, 1 ed., New York, NY: Russell Sage Foundation, 2011, chapter 3, pp. 47–69.

Durkin, Kelley, Mark Lipsey, Dale Farran, and Sarah Wiesen, "Effects of a Statewide Pre-Kindergarten Program on Children's Achievement and Behavior Through Sixth Grade," *Developmental Psychology*, 2022, *58* (3), 470–484.

Feller, Avi, Todd Grindal, Luke Miratrix, and Lindsay C. Page, "Compared to what? Variation in the impacts of early childhood education by alternative care type," *Annals of Applied Statistics*, 2016, *10* (3), 1245–1285.

Ferman, Bruno and Cristine Pinto, "Inference in Differences-in-Differences with Few Treated Groups and Heteroskedasticity," *The Review of Economics and Statistics*, 2019, *101* (3), 452–467.

— **and** —, "Synthetic Controls with Imperfect Pretreatment Fit," *Quantitative Economics*, 2021, *12* (4), 1197–1221.

Finkelstein, Amy and Nathaniel Hendren, "Welfare Analysis Meets Causal Inference," *Journal of Economic Perspectives*, 2020, *34* (4), 146–167.

Fitzpatrick, Maria, "Starting School at Four: The Effect of Universal Pre-Kindergarten on Children's Academic Achievement," *The B.E. Journal of Economic Analysis Policy*, 2008, *8* (1), 1–40.

Friedman-Krauss, Allison H., Steven Barnett, Katherine S. Hodges, Karin A. Garver, G.G. Weisenfeld, Beth Ann Gardiner, and Tracy Merriman Jost, "The State of Preschool 2022: State Preschool Yearbook," Technical Report, National Institute for Early Education Research, New Brunswick, NJ 2023.

Genadek, Katie, Joshua Sanders, and Amanda Jean Stevenson, "Measuring U.S. Fertility using Administrative Data from the Census Bureau," Technical Report ADEP-WP-2021-02, U.S. Census Bureau July 2021.

Gormley Jr., William T., Sara Amadon, Katherine Magnuson, Amy Claessens, and Douglas Hummel-Price, "Universal Pre-K and College Enrollment: Is There a Link?," *AERA Open*, 2023, *9.*

Gray-Lobe, Guthrie, Parag Pathak, and Christopher Walters, "The Long-Term Effects of Universal Preschool in Boston," *The Quarterly Journal of Economics*, 2023, *138* (1), 363–411.

Griffen, Andrew S., "Evaluating the effects of childcare policies on children's cognitive development and maternal labor supply," *Journal of Human Resources*, 2019, *54* (3), 604–655.

Gupta, Snigdha, Lauren H. Supplee, Dana Suskind, and John A. List, "Failed to Scale: Embracing the Challenge of Scaling in Early Childhood," in John A. List, Dada Suskind, and Lauren H. Supplee, eds., *The Scale-Up Effect in Early Childhood and Public Policy: Why Interventions Lose Impact at Scale and What We Can Do About It*, New York, NY: Routledge, 2021, pp. 1–21.

Harden, Brenda J., Bonnie E. Brett, Jacquelyn T. Gross, Christina Weiland, Jordan Berne, Elisa L. Klein, and Christy Tirrell-Corbin, "Benefits of pre-kindergarten for children in Baltimore, MD," *Early Childhood Research Quarterly*, 2023, *64*, 1–12.

Heckman, James J., Lance Lochner, and Christopher R. Taber, "General equilibrium treatment effects: A study of tuition policy," Technical Report Working Paper 6426, National Bureau of Economic Research, Cambridge, MA February 1998.

Heckman, James, Neil Hohmann, Jeffrey Smith, and Michael Khoo, "Substitution and Dropout Bias in Social Experiments: A Study of an Influential Social Experiment," *The Quarterly Journal of Economics*, 2000, *115* (2), 651–694.

Hendren, Nathaniel, "The Policy Elasticity," *Tax Policy and the Economy*, 2016, *30*, 51–89.

— and Ben Sprung-Keyser, "A Unified Welfare Analysis of Government Policies," *Quarterly Journal of Economics*, 2020, *135* (3), 1209–1318.

Henry, Gary, Craig Gordon, and Dana Rickman, "Early Education Policy Alternatives: Comparing Quality and Outcomes of Head Start and State Prekindergarten," *Educational Evaluation and Policy Analysis*, 2006, *28* (1), 77–99.

— , — , Laura Henderson, and Bentley Ponder, "Georgia Pre-K Longitudinal Study: Final Report, 1996-2001," Technical Report, Andrew Young School of Policy Studies, Georgia State University, Atlanta, GA May 2003.

— , Laura Henderson, Bentley Ponder, Craig Gordon, Andrew Mashburn, and Dana Rickman, "Report on the Findings from the Early Childhood Study: 2001-02," Technical Report, Andrew Young School of Policy Studies, Georgia State University, Atlanta, GA August 2003.

Henry, Gary T. and Craig S. Gordon, "Competition in the sandbox: A test of the effects of preschool competition on educational outcomes," *Journal of Policy Analysis and Management*, 2006, *25* (1), 97–127.

Humphries, John Eric, Christopher Neilson, Xiaoyang Ye, and Seth D. Zimmerman, "Parents' Earnings and the Returns to Universal Pre-Kindergarten," Technical Report Working Paper 33038, National Bureau of Economic Research, Cambridge, MA October 2024.

Johnson, Rucker and C. Kirabo Jackson, "Reducing Inequality through Dynamic Complementarity: Evidence from Head Start and Public School Spending," *American Economic Journal: Economic Policy*, 2019, *11* (4), 310–349.

Kline, Patrick and Christopher Walters, "Evaluating Public Programs with Close Substitutes: The Case of Head Start," *The Quarterly Journal of Economics*, 2016, *131* (4), 1795–1848.

Kose, Esra, "Public investments in early childhood education and academic performance: Evidence from Head Start in Texas," *Journal of Human Resources*, 2023, *58* (6), 2042–2069.

Lohman, Judith, "Kindergarten Entrance Age," Technical Report 2000-R-1188, Connecticut General Assembly December 2000.

Ludwig, Jens and Douglas L. Miller, "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design," *The Quarterly Journal of Economics*, 2007, *122* (1), 159–208.

Monnet, Jessica, "The effect of preschool participation on intellectual and behavioral disorder diagnoses: Evidence from surveys on children's health," *Economics of Education Review*, 2019, *68*, 136–147.

Nandi, Preetha, Michael Kramer, and Melissa Kottke, "Changing disparities in teen birth rates and repeat birth rates in Georgia: implications for teen pregnancy prevention," *Contraception*, 2019, *99* (3), 175–178.

Neidell, Matthew and Jane Waldfogel, "Cognitive and Noncognitive Peer Effects in Early Education," *The Review of Economics and Statistics*, 2010, *92* (3), 562–576.

Nielsen, Eric, "The Income-Achievement Gap and Adult Outcome Inequality," *Journal of Human Resources*, 2023.

Pages, Remy, Dylan Lukes, Drew Bailey, and Greg Duncan, "Elusive Longer-Run Impacts of Head Start: Replications Within and Across Cohorts," *Educational Evaluation and Policy Analysis*, 2020, *42* (4), 471–492.

Raden, Anthony, "Universal Prekindergarten in Georgia: A Case Study of Georgia's Lottery-Funded Pre-K Program," Technical Report, Foundation for Child Development 1999.

Rothbart, Michah and Taryn Morrissey, "Beyond the Classroom: The Broader Effects of Public Preschool Expansion on Children's Early Educational Experiences and Early Literacy Skills," Technical Report 250, Center for Policy Research, Syracuse, NY 2024.

Swenson, Kendall, "Child Care Arrangements in Urban and Rural Areas," Technical Report, U.S. Department of Health and Human Services, Washington, DC June 2008.

Taylor, Evan, Bryan Stuart, and Martha Bailey, "Summary of Procedure to Match NUMIDENT Place of Birth County to GNIS Places," Technical Report 1284 Technical Memo 2, U.S. Census Bureau May 2016.

Temple, Judy A., "Rural Gaps in Participation in Early Childhood Education," *Journal of Agricultural and Applied Economics*, 2009, *41* (2), 403—-410.

Thompson, Owen, "Head Start's Long-Run Impact: Evidence from the Program's Introduction," *Journal of Human Resources*, 2018, *53* (4), 1100–1139.

Tietze, Wolfgang, Debby Cryer, Joachim Bairrão, Jesús Palacios, and Gottfried Wetzel, "Comparisons of Observed Process Quality in Early Child Care and Education Programs in Five Countries," *Early Childhood Research Quarterly*, 1996, *11*, 447–475.

Weiland, Christina and Hirokazu Yoshikawa, "Impacts of a Prekindergarten Program on Children's Mathematics, Language, Literacy, Executive Function, and Emotional Skills," *Child Development*, 2013, *84* (6), 2112–2130.

__, Rebecca Unterman, Anna Shapiro, Sara Staszak, Shana Rochester, and Eleanor Martin, "The Effects of Enrolling in Oversubscribed Prekindergarten Programs Through Third Grade," *Child Development*, 2020, *91* (5), 1401–1422.

Woodyard, Henry, Tim Sass, and Ishtiaque Fazlul, "Assessing the Benefits of Education in Early Childhood: Evidence from a Pre-K Lottery in Georgia," Technical Report EdWorkingPaper: 23-880, Annenberg Institute at Brown University November 2023.

Yang, Hyunwoo, "The Effects of Wisconsin's Universal Prekindergarten Program on Third-Grade Academic Achievement," *American Educational Research Journal*, 2024, pp. 403—-410.

Zerpa, Mariana, "Short and Medium Run Impacts of Preschool Education: Evidence from State Pre-K Programs," Technical Report, Working Paper September 2022.

Appendix A Supplemental Tables and Figures

Table A.1 — County-Level Predictors of UPK Enrollment in 1995-2000

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IV)
Metropolitan	−0.052*							0.044	−1.673
(standard error)	(0.029)							(0.043)	(1.548)
Log(HH income)		−0.206***						0.036	
(standard error)		(0.058)						(0.114)	
BA share (25+)			−0.429*					0.284	0.239
(standard error)			(0.221)					(0.313)	(0.394)
Single mother share (0-5)				0.604***				0.881***	0.847***
(standard error)				(0.151)				(0.285)	(0.293)
Log(pop. age 0-5)					−0.028**			−0.028	−0.023
(standard error)					(0.012)			(0.019)	(0.020)
Black share						0.240**		−0.116	−0.123
(standard error)						(0.098)		(0.151)	(0.153)
Pre-K enrollment (4-5)							−0.427**	−0.530**	
(standard error)							(0.174)	(0.206)	
Metro*Log(HH income)									0.121
(standard error)									(0.139)
Non-metro*Log(HH income)									−0.049
(standard error)									(0.142)
Metro*Pre-K enrollment									−0.558
(standard error)									(0.393)
Non-metro*Pre-K enrollment									−0.525**
(standard error)									(0.215)
Observations	954	954	954	954	954	954	954	954	954

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: This table presents estimates from regressions of county-level UPK enrollment rates on county-level characteristics (as measured in the 1990 census). I discuss the construction of the county-level UPK enrollment rates in Appendix B. Metropolitan and non-metropolitan status are binary indicators, and all other independent variables are continuous. The parentheticals next to some of the variable names indicate the age of sample members in the 1990 census used to calculate the county-level statistics. Each of Georgia's 159 counties is included in the sample six times, once for each year between 1995 and 2000. Standard errors are clustered at the county level to account for serial correlation. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701 and CBDRB-FY25-0013).

Table A.2 — Georgia County Cross-Tabulations, 1990 Census Characteristics

Pre-K enrollment rate	**Metropolitan status** Non-metro	Metro
Below median	69	11
Above median	48	31

HH income level	**Metropolitan status** Non-metro	Metro
Below median	76	4
Above median	41	38

Pre-K enrollment rate	**HH income level** Below median	Above median
Below median	48	32
Above median	32	47

Notes: Each table gives the number of counties in Georgia of a particular type. "Pre-K enrollment rate" is the county-level share of children ages four or five enrolled in public or private preschool. "HH income level" is the county-level median household income. Both of these characteristics are measured using 1990 long-form census data, before Georgia introduced state-funded pre-K. Metropolitan status comes from the 1990 urban-rural classification scheme of the National Center for Health Statistics. Metropolitan counties are those within metropolitan statistical areas with at least 50,000 people. All other counties are non-metropolitan. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY25-0013).

Table A.3 — Sample Size by Pre-K Cohort

	Teen Fertility Sample		20+ ACS Sample		
Pre-K Cohort	Control Group	Georgia	Observed Ages	Control Group	Georgia
1987	317,000	22,500	[21,40]	232,000	16,000
1988	313,000	22,500	[20,39]	230,000	16,000
1989	319,000	24,000	[20,38]	232,000	17,000
1990	316,000	24,000	[20,37]	221,000	16,500
1991	316,000	25,000	[20,36]	208,000	16,000
1992	324,000	26,000	[20,35]	200,000	15,500
1993	334,000	27,000	[20,34]	195,000	15,500
1994	343,000	28,500	[20,33]	188,000	14,500
1995	340,000	27,500	[20,32]	174,000	13,500
1996	337,000	27,500	[20,31]	158,000	13,000
1997	325,000	27,500	[20,30]	140,000	11,000
1998	328,000	27,500	[20,29]	125,000	10,000
1999	331,000	28,000	[20,28]	109,000	8,900
2000	331,000	28,000	[20,27]	93,000	7,700

Notes: The two samples are described in Section 4.1. The control group consists of the 23 states named in Section 4.3 and shown in Appendix Figure A.1. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY25-0013). Observation counts are rounded to comply with Census Bureau disclosure avoidance rules.

Table A.4 — Georgia UPK's Long-Run Effects, by County Type II

	Employed or in School	Has Positive Earnings	Earnings (with 0s)	Log of Earnings
	Panel A. Below Median Pre-K Enrollment			
UPK impact	0.020	0.020	2,328	0.027
(standard error)	(0.012)	(0.011)	(1,379)	(0.033)
Impact in %	2.7%	2.5%	9.8%	0.3%
Counterfactual mean	0.746	0.803	23,650	9.96
Observations	311,000	311,000	311,000	266,000
	Panel B. Above Median Pre-K Enrollment			
UPK impact	0.008	0.003	905	0.020
(standard error)	(0.010)	(0.009)	(811)	(0.020)
Impact in %	1.0%	0.4%	3.1%	0.2%
Counterfactual mean	0.791	0.842	29,350	10.06
Observations	1,262,000	1,262,000	1,262,000	1,088,000
	Panel C. Non-Metropolitan			
UPK impact	0.005	0.010	1,774	0.014
(standard error)	(0.010)	(0.009)	(1,229)	(0.029)
Impact in %	0.7%	1.2%	7.5%	0.1%
Counterfactual mean	0.753	0.806	23,780	9.97
Observations	444,000	444,000	444,000	380,000
	Panel D. Metropolitan			
UPK impact	0.012	0.005	817	0.023
(standard error)	(0.010)	(0.009)	(929)	(0.021)
Impact in %	1.5%	0.6%	2.7%	0.2%
Counterfactual mean	0.794	0.846	30,240	10.08
Observations	1,129,000	1,129,000	1,129,000	974,000
	Panel E. Below Median Household Income			
UPK impact	−0.010	0.002	2,027	0.036
(standard error)	(0.015)	(0.014)	(1,336)	(0.061)
Impact in %	−1.3%	0.2%	8.8%	0.4%
Counterfactual mean	0.757	0.802	23,010	9.93
Observations	337,000	337,000	337,000	284,000
	Panel F. Above Median Household Income			
UPK impact	0.013	0.006	932	0.017
(standard error)	(0.010)	(0.009)	(840)	(0.022)
Impact in %	1.6%	0.7%	3.2%	0.2%
Counterfactual mean	0.788	0.841	29,350	10.07
Observations	1,237,000	1,237,000	1,237,000	1,070,000

*** p<0.01, ** p<0.05, * p<0.1

Notes: Author's calculations from estimating equation 2. Inference is conducted using the Ferman and Pinto (2019) method. Each "counterfactual mean" is the average of observed outcomes among individuals exposed to Georgia UPK minus the estimated treatment effect. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701 and CBDRB-FY25-0013). Observation counts are rounded to comply with Census Bureau disclosure avoidance rules.

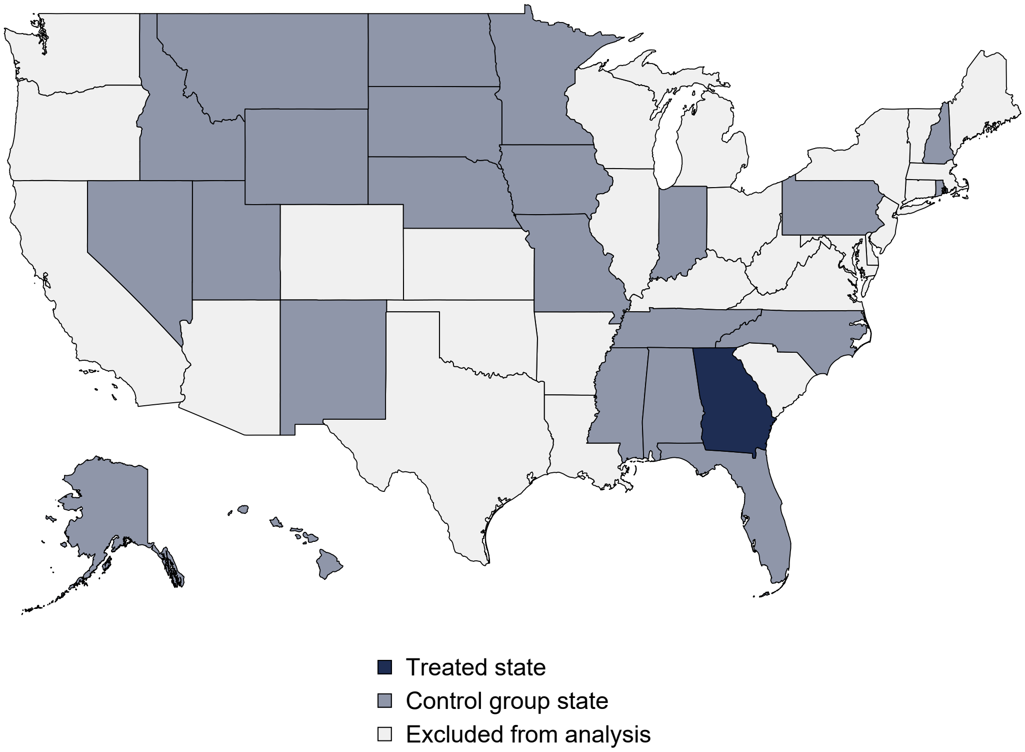
Table A.5 — Georgia UPK's Long-Run Impacts, by Race×Sex II

	Employed or in School	Has Positive Earnings	Earnings (with 0s)	Log of Earnings
	Panel A. White Girls			
UPK impact	0.012	0.004	502	0.021
(standard error)	(0.016)	(0.014)	(860)	(0.033)
Impact in %	1.5%	0.5%	1.9%	0.2%
Counterfactual mean	0.792	0.822	27,130	10.02
Observations	738,000	738,000	738,000	614,000
	Panel B. White Boys			
UPK impact	0.005	0.003	258	0.006
(standard error)	(0.007)	(0.005)	(1,422)	(0.035)
Impact in %	0.6%	0.3%	0.7%	0.1%
Counterfactual mean	0.801	0.867	35,090	10.24
Observations	765,000	765,000	765,000	685,000
	Panel C. Black Girls			
UPK impact	−0.006	0.008	668	0.021
(standard error)	(0.016)	(0.019)	(748)	(0.039)
Impact in %	−0.8%	1.0%	3.4%	0.2%
Counterfactual mean	0.762	0.801	19,890	9.71
Observations	106,000	106,000	106,000	83,500
	Panel D. Black Boys			
UPK impact	0.000	0.000	190	0.002
(standard error)	(0.011)	(0.023)	(1,804)	(0.025)
Impact in %	0.0%	0.0%	1.0%	0.0%
Counterfactual mean	0.544	0.658	18,110	9.78
Observations	118,000	118,000	118,000	75,500

*** p<0.01, ** p<0.05, * p<0.1

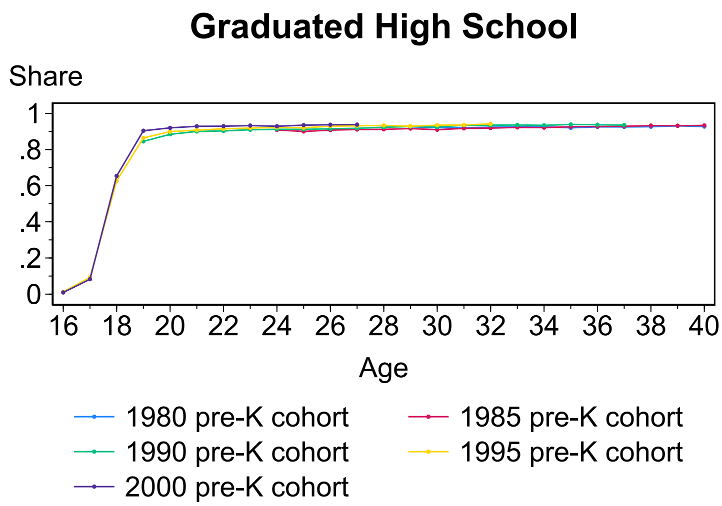
Notes: Author's calculations from estimating equation 2. Inference is conducted using the Ferman and Pinto (2019) method. Each "counterfactual mean" is the average of observed outcomes among individuals exposed to Georgia UPK minus the estimated treatment effect. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701 and CBDRB-FY25-0013). Observation counts are rounded to comply with Census Bureau disclosure avoidance rules.

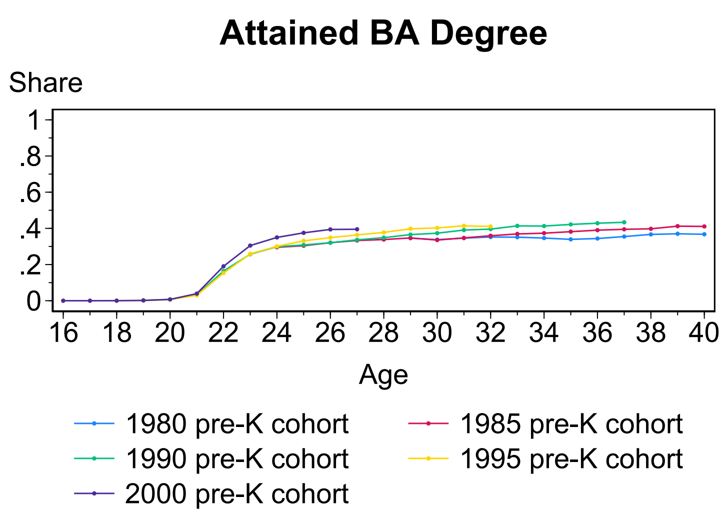
Figure A.1 — Treatment and Control Group States



Notes: The 23 control group states are Alabama, Alaska, Florida, Hawaii, Idaho, Indiana, Iowa, Minnesota, Mississippi, Missouri, Montana, Nebraska, Nevada, New Hampshire, New Mexico, North Carolina, North Dakota, Pennsylvania, Rhode Island, South Dakota, Tennessee, Utah, and Wyoming. These are the states that had no more than a very small statewide pre-K program, if any, during the time of my analysis (Friedman-Krauss et al., 2023).

Figure A.2 — Long-Run Outcome Stabilization, by Age





Notes: Author's calculations using public-use American Community Survey data from 2005 through 2020. The sample includes people born in all U.S. states other than Georgia.

Appendix B Data Details

B.1 1990 and 2000 Census Data

To investigate enrollment, I use data from the 1990 and 2000 long-form censuses, in which I observe the 1989 and 1999 pre-K cohorts. Each long-form census is a one-in-six sample of all U.S. households with detailed demographic and economic information, including preschool enrollment for young children. In both censuses, I observe whether a child is enrolled in no pre-K, private pre-K, or public pre-K.

In the 2000 census, I observe children's exact date of birth, which I use to sort children into exact pre-K cohorts according to the kindergarten cutoff of their state of residence (retrieved from Lohman 2000). For the substitution analysis in Section 6.1, I use exact pre-K cohorts.

Exact birthdays are not observed in the 1990 census data; I only observe a child's age (in whole numbers) at the time of the survey. Thus, I cannot define exact pre-K cohorts in the 1990 census. Instead, I define the pre-K cohort as including all children who are ages four and five. Because the census is conducted in April, children in the 1989 pre-K cohort were four or five at the time of the 1990 census. To verify that the enrollment rate for four- and five-year-olds (collectively) matches the enrollment rate of the exact pre-K cohort, I compare enrollment rates in the 2000 census, where exact birthdays are observed. As Table B.1 shows, in the 2000 census, pre-K enrollment rates in Georgia and in the control group are highly similar for four- and five-year-olds and the exact pre-K cohorts. For consistency in the definition of a cohort across years, I use four- and five-year-olds in the DiD analysis of net enrollment changes in Section 3.

B.2 County-Level UPK Data

In Section 3 I examine UPK enrollment at the county level using data from two sources. First, I obtain the number of children enrolled in UPK in 1998 from an administrative report on the use of Georgia's lottery funds (Brackett et al., 1999). Second, I estimate the number of children enrolled in each county×year between 1995 and 2000 using data on UPK spending shared with me by Georgia's Department of Early Care and Learning (DECAL). To convert county×year level spending to enrollment, I divide by UPK's state×year level annual cost per child (publicly reported by DECAL). To convert enrollment counts into shares, I use publicly available estimates of the number of four-year-olds in each county×year from the Surveillance, Epidemiology, and End Results Program at the National Cancer Institute. The data from the administrative report are used to construct Figure 2, and the data from DECAL are used for the regression analysis in Table A.1.

Note that UPK costs per child may vary somewhat across counties due to the nature of Georgia UPK's funding formula. The state-level cost per child is therefore an imperfect proxy for county-level costs per child. To test the validity of the 1995-2000 estimates derived from the spending data, I compare the estimated 1998 enrollment rates to the "actual" enrollment rates derived from the 1998 administrative report. As B.1 shows, the estimated rates match the actual rates very closely.

B.3 Teen Fertility Sample

To estimate impacts on teen fertility, I combine data from multiple sources. As a base for the sample, I use the 2000 long-form census and the 2005-2022 waves of the ACS. As the precursor to the ACS, the long-form census is representative of the same population and contains nearly identical survey questions. Including the 2000 census in the sample broadens the sample size. Note that I am unable to use the 2000 census in the primary sample because children in the 1987-2000 pre-K cohorts were too young at the time of the 2000 census for long-run outcomes to have materialized. This is not an issue for the teen fertility sample because the data on teen fertility are from an outside source that is not tied to the timing of any survey.

I obtain data on births from the Census Household Composition Key (CHCK), an underutilized dataset derived from administrative birth certificate data. After childbirth, nearly all parents choose to transmit their child's birth certificate data to the SSA. Those data are then turned into a research file, called the CHCK, which contains information on birth timing and parent-child links (Genadek et al., 2021).

To measure teen fertility, I merge data on the timing of births from the CHCK to the census/ACS sample and compare the timing of a birth to a parent's exact birthday. As with county of birth, only individuals with a unique Census identifier can be linked to the CHCK. Therefore, I limit the teen fertility sample to individuals with an identifier.

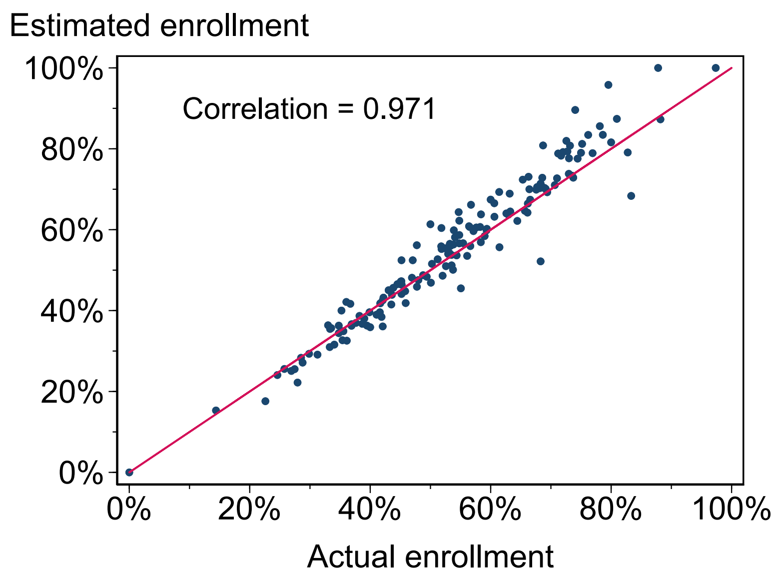
The teen fertility rates I observe in my sample are substantially higher than analogous statistics reported in other sources. According to Nandi et al. (2019), the birth rate among females ages 15-19 in Georgia was 4.5% between 2008 and 2010 (which covers most of the treated cohorts in my analysis). In comparison, I estimate 14.3% of the girls in Georgia's treated cohorts had children as teenagers. The difference may be attributable to a couple factors. First, the CHCK data may contain some slight error in the measurement of fertility because adoptive parents sometimes replace birth parents in Social Security records (Genadek et al., 2021). Second, teen birth rates reported in other sources are typically point-in-time statistics measuring how many teenagers had children in a given year. Conversely, my statistic is a cumulative measure for how many individuals ever had children as a teenager. I am currently investigating these issues to better understand the differing statistics.

Table B.1 — Pre-K Enrollment by Definition of Pre-K Cohort, 2000 Census

	Control Group		Georgia	
	Pre-K Age	Ages 4 and 5	Pre-K Age	Ages 4 and 5
Private pre-K	0.268	0.231	0.233	0.215
Public pre-K	0.411	0.462	0.588	0.569
Observations	122,000	245,000	16,500	32,500

Notes: Author's calculations using data from the 2000 long-form census. The "Pre-K Age" columns estimate enrollment among children exactly in the pre-K cohort based on their state kindergarten cutoff. The "Ages 4 and 5" columns estimate enrollment among children who are four or five years old at the time of the census. The control group consists of the 18 states with no state pre-K program in 1999 (Alabama, Alaska, Florida, Hawaii, Idaho, Indiana, Minnesota, Mississippi, Montana, Nebraska, Nevada, New Hampshire, North Carolina, North Dakota, Rhode Island, South Dakota, Utah, and Wyoming). These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701). Observation counts are rounded to comply with Census Bureau disclosure avoidance rules.

Figure B.1 — Actual and Estimated Georgia UPK Enrollment Rates in 1998, by County



Notes: This figure compares "actual" county-level UPK enrollment rates in 1998 with estimated rates. The "actual" enrollment rates are obtained by dividing county-level enrollment counts (obtained from Brackett et al. (1999)) by estimates of the number of four-year-olds in each county×year from the Surveillance, Epidemiology, and End Results (SEER) Program at the National Cancer Institute. The estimated enrollment rates are obtained by converting county-level UPK spending by state-level per-child costs, and then dividing by the SEER population estimates.

Appendix C Net Effects of Georgia UPK on Pre-K Enrollment

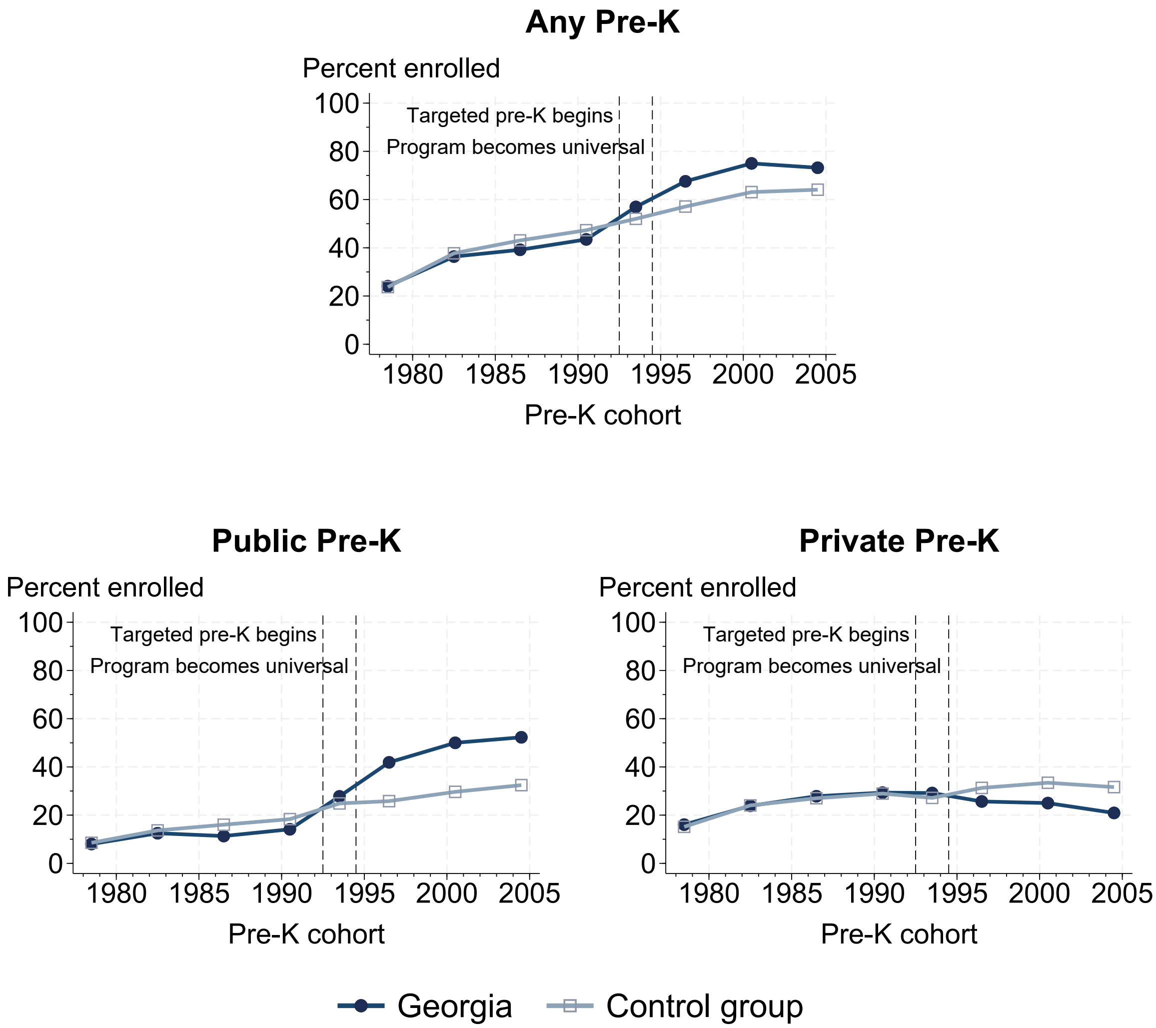
Interpreting the DiD analysis in Section 3 requires a parallel trends assumption. However, I do not observe enrollment trends in the pre-treatment period with the census data because there is only one pre-treatment cohort. Instead, I draw on data from the Current Population Survey (CPS) School Enrollment Supplement, following Cascio and Schanzenbach (2013). As in the long-form census, the CPS contains data on enrollment in no pre-K, private pre-K, and public pre-K. I limit the sample to children who are four years old when surveyed. Because the School Enrollment Supplement is fielded in October, most children in the exact pre-K cohort should be four in October.

The main advantage of the CPS is that I observe enrollment in multiple pre-treatment years, and the main disadvantage is the small sample size. Following Cascio and Schanzenbach (2013), I pool cohorts together to reduce noise from small CPS samples. Each pool contains four cohorts, other than one pool which contains the two cohorts exposed to Georgia's targeted pre-K program.

First, Figure C.1 shows that enrollment in the pre-treatment period was very similar in levels and trends in Georgia and the control group. Next, I estimate event study models with state of residence and pre-K cohort fixed effects, using data from the 1981-2006 pre-K cohorts. As Figure C.2 shows, the pre-treatment coefficients are close to 0 and statistically insignificant without exception. After UPK's introduction, the probability of enrolling in public pre-K rises, and the probability of enrolling in private pre-K falls. On net, enrollment in any pre-K rises.

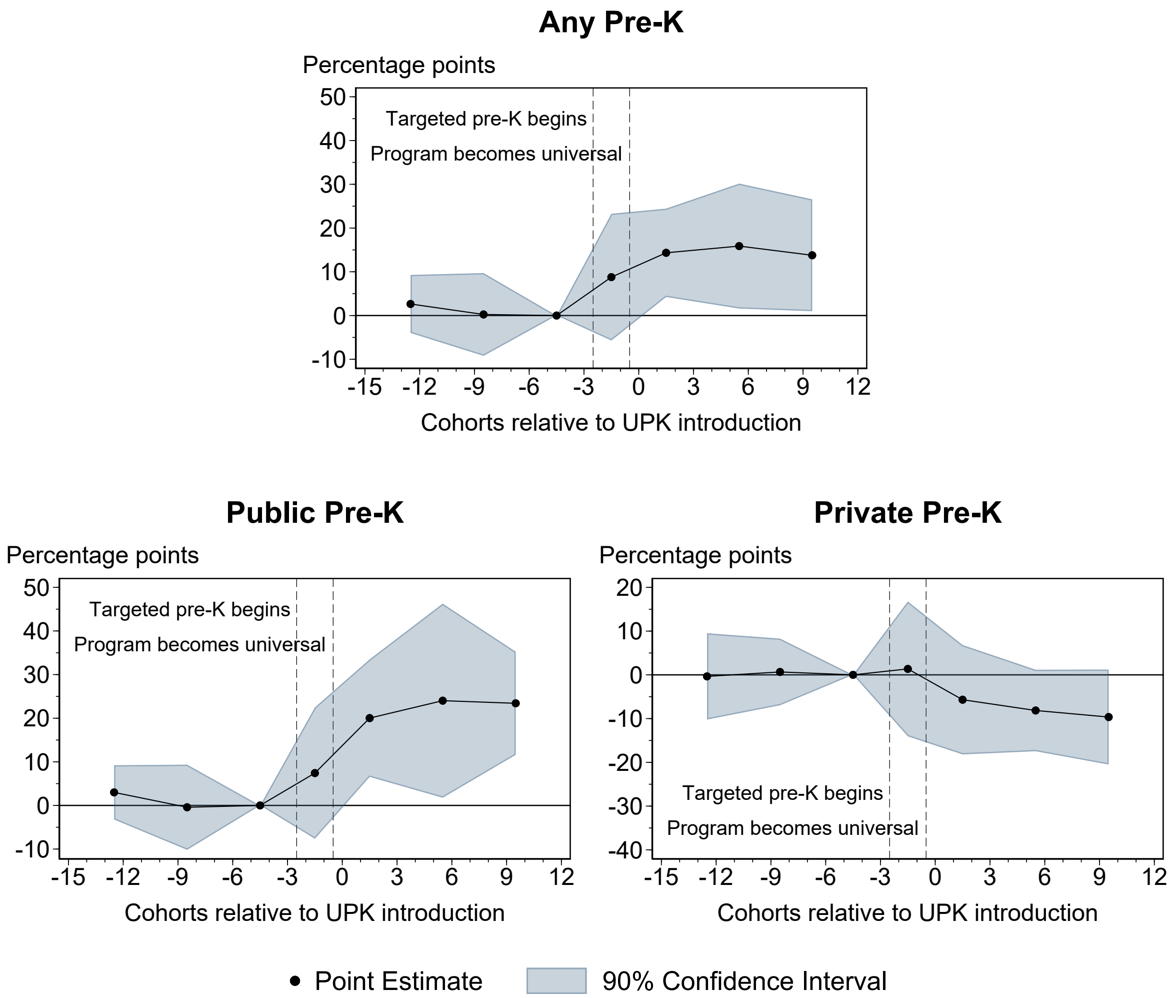
Note that the CPS impact estimates are larger than the corresponding estimates using the long-form census data, although the differences are not statistically significant. I privilege the estimates from the census data due to the its greater sample size and because enrollment rates in the census align better with administratively reported pre-K enrollment rates in Georgia. Nevertheless, I view the CPS analysis as providing strong evidence on *trends* in pre-K enrollment in Georgia and the control group.

Figure C.1 — Pre-K Enrollment Trends (CPS)



Notes: Author's calculations using data from the Current Population Survey. Cohorts are pooled to reduce noise from small samples. Each pool contains four cohorts, other than the pool which contains the two cohorts exposed to Georgia's targeted pre-K program. The estimates are plotted at the midpoint of each pool.

Figure C.2 — Effect of Georgia UPK on Net Changes in Pre-K Enrollment (CPS)



Notes: Author's calculations using data from the Current Population Survey. Inference is conducted using the Ferman and Pinto (2019) method. Cohorts are pooled to reduce noise from small samples. Each pool contains four cohorts, other than the pool which contains the two cohorts exposed to Georgia's targeted pre-K program. The estimates are plotted at the midpoint of each pool.

Appendix D Robustness Checks

D.1 Alternative DiD Specifications

My first set of robustness checks tests the sensitivity of the primary DiD results to alternative regression specifications and samples. The results, shown in Table D.1, are very similar to the primary estimates in Tables 3 and 4.

In the first column of Table D.1, I remove all covariates from the model. This specification relies on an unconditional parallel trends assumption rather than a conditional one.

In the second column, I restrict the non-fertility samples to individuals surveyed in the ACS at 32 or younger. The motivation for this check is that older cohorts in the sample are, by definition, observed at older ages in the ages. Using a maximum age imposes that the sample is observed at more similar ages.

In the third column, I impose a minimum age of 25 for inclusion in the sample. This test ensures the main results are not driven by outcomes during an early part of adulthood that is not representative of later age outcomes.

In the fourth column, I restrict the control group to states with no state pre-K program in 1999: Alabama, Alaska, Florida, Hawaii, Idaho, Indiana, Minnesota, Mississippi, Montana, Nebraska, Nevada, New Hampshire, North Carolina, North Dakota, Rhode Island, South Dakota, Utah, and Wyoming. This is the control group I use in the enrollment analyses (Sections 3 and 6).

In the fifth column, I replace birth county fixed effects with birth state fixed effects.

D.2 Synthetic Control Method

The second set of robustness checks uses the synthetic control method (SCM), which takes a different approach to estimating Georgia's counterfactual than the DiD analysis. SCM constructs a counterfactual by taking a weighted average of "donor pool" states, with weights chosen so that the weighted average closely approximates Georgia's observed outcomes in the pre-treatment period (Abadie et al., 2010). Assuming this "synthetic Georgia" provides an accurate counterfactual, differences in post-treatment outcomes can be interpreted as UPK's causal effects.

This analysis addresses two potential concerns with the DiD analysis. If there are concerns about non-parallel trends in the pre-period, the SCM analysis has a control unit that has parallel trends in the pre-period by construction. Second, conventional inference is a challenge in DiD analyses with one treated unit. The SCM literature has developed alternative approaches for inference that are applicable in settings with only one treated unit.

To implement the SCM analysis, I use pre-K cohorts from 1980 through 2000 since SCM is more reilable with a longer pre-treatment period. I use the primary DiD control group states for the donor pool. Following Doudchenko and Imbens (2016), I account for individual-level covariates by residualizing people's observed outcomes. Using only data on the pre-treatment cohorts (1980-1992), I estimate the following model:

$$Y_{is} = \Pi \mathbf{X}_i + \lambda_s + \epsilon_{is}, \tag{D.1}$$

where the vector $\mathbf{X}_i$ is the same as in equation 2 and λ_s contains state of birth fixed effects. Then, I use the estimated parameters to calculate residuals for the entire sample (pre- and post-treatment). Finally, I aggregate the residuals, rather than the observed outcomes, to the state×cohort level. Throughout the SCM analysis I use the aggregated residuals as the outcome variable. This procedure purges Georgia and the donor pool of differences driven by individual-level covariates.

Notice that the state of birth fixed effects in equation D.1 also eliminate permanent additive differences between Georgia and donor pool states. Demeaning forces SCM to match on trends rather than levels, which is important in this setting because Georgia's educational outcomes tend to be at the bottom of the donor pool distribution (Doudchenko and Imbens, 2016; Ferman and Pinto, 2021). Following the recent SCM literature, I use every pre-treatment outcome as a predictor and no other predictors (Ferman and Pinto, 2021; Doudchenko and Imbens, 2016).

Letting $\tau_{s,t}$ denote the difference between real and synthetic outcomes for cohort t of state s, my estimate of Georgia UPK's ITT effect is $\frac{1}{6}\sum_{t=1995}^{2000} \tau_{Georgia,t}$. Note that the two cohorts exposed to Georgia's short-lived targeted pre-K program are excluded. The SCM results are shown in Figures D.1 and D.2. The weights selected for each state in the donor pool are shown in Tables in D.2 and D.3.

I conduct inference using the conformal approach of Chernozhukov et al. (2021). The basic intuition for this approach is that the distribution of differences between Georgia and synthetic Georgia should be similar in the pre- and post-treatment periods if UPK had no impact. I implement the procedure by permuting estimated blocks of $\tau_{s,t}$ across pre-K cohorts. For each permutation, I calculate the following test statistic: $\frac{1}{\sqrt{6}}|\sum_{t=1995}^{2000} \tau_{Georgia,t}|$. The p-value is the share of permuted test statistics that are greater than the observed test statistic.

D.3 Demographic Changes

This robustness check tests whether the estimated long-run impacts could be driven by demographic changes in Georgia that coincide with, or were perhaps caused by, the introduction of UPK. To test this, I first create summary measures for the observable demographic variables. First, I regress each long-run outcome on individual and birth county covariates, using only the control group. The individual-level independent variables are children's sex and race. The birth county-level independent variables are metropolitan status, log of median household income, share of adults with a bachelor's degree, the share of young children in single mother households, the share of individuals who are Black, and the any pre-K enrollment rate. These birth county characteristics are all measured in the 1990 census. Next, I use the estimated coefficients to obtain predicted outcomes for each individual in the full sample.

I then estimate DiD models with birth state fixed effects, no covariates other than "age when surveyed," and the predicted values as the dependent variables. Non-zero estimates could suggest that the introduction of Georgia UPK coincided with compositional changes that, on their own, affected long-run outcomes. One might expect compositional changes if families moved to Georgia (or moved within Georgia) specifically because of UPK. However, as Table D.4 shows, I find no evidence that demographic changes drive the main estimates. The estimates are small and statistically insignificant.

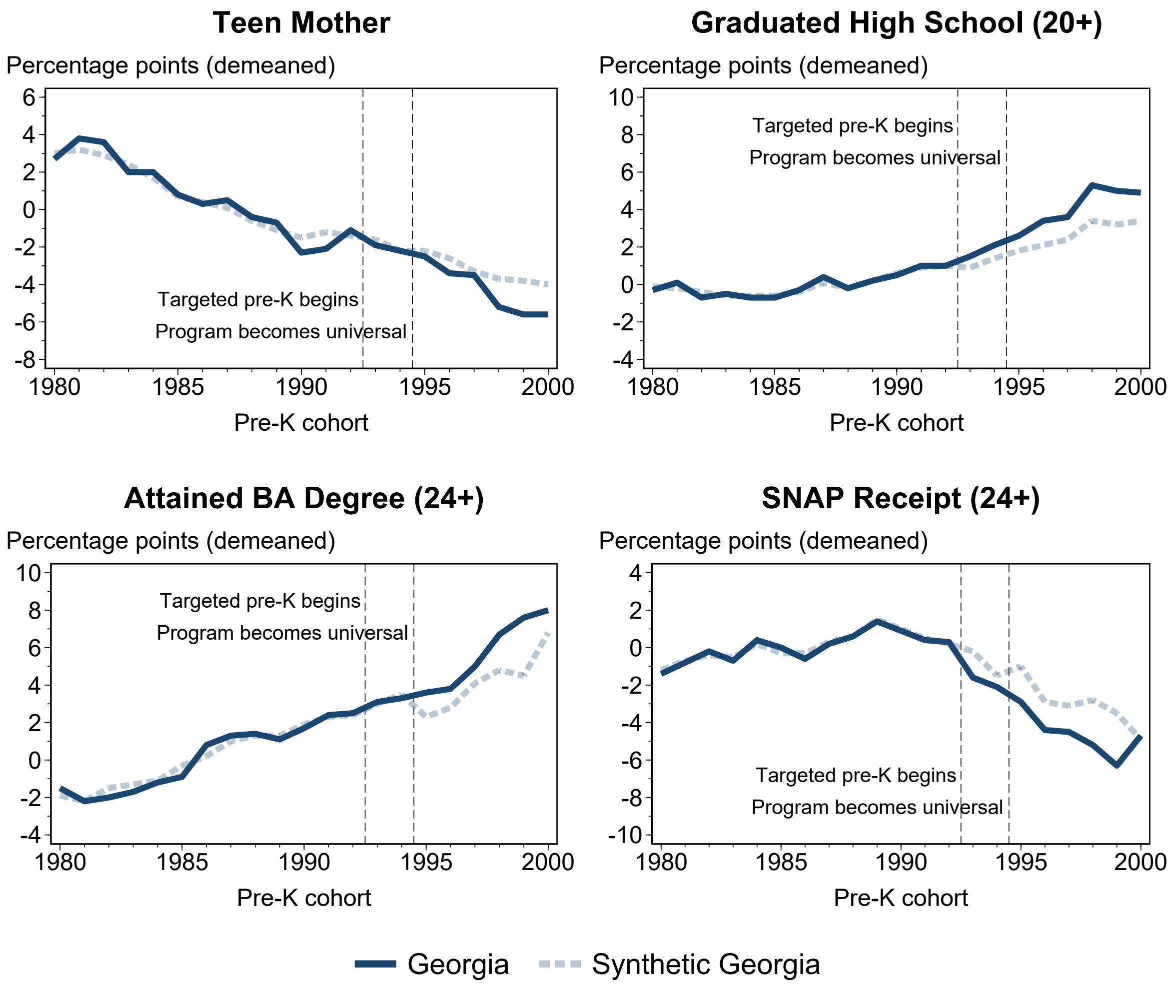
Table D.1 — Robustness of DiD Estimates to Alternative Specifications

	(I)	(II)	(III)	(IV)	(V)
Teen father	-0.002	-	-	-0.002	-0.003
(standard error)	(0.003)	-	-	(0.003)	(0.003)
Teen mother	-0.015**	-	-	-0.015**	-0.018**
(standard error)	(0.009)	-	-	(0.008)	(0.010)
Graduated high school	0.014*	0.018**	0.014*	0.015*	0.015*
(standard error)	(0.008)	(0.009)	(0.008)	(0.009)	(0.009)
Attained BA degree	0.014	0.016	0.016	0.013	0.013
(standard error)	(0.013)	(0.014)	(0.013)	(0.014)	(0.013)
Receive SNAP	-0.018*	-0.022*	-0.016*	-0.016**	-0.020*
(standard error)	(0.011)	(0.012)	(0.015)	(0.011)	(0.009)
Employed or in school	0.004	0.008	0.007	0.003	0.005
(standard error)	(0.009)	(0.010)	(0.009)	(0.009)	(0.009)
Has positive earnings	0.005	0.011	0.006	0.004	0.006
(standard error)	(0.008)	(0.008)	(0.007)	(0.008)	(0.007)
Earnings (with 0s)	1,243	1,094	1,042	923	1,023
(standard error)	(895)	(522)	(666)	(589)	(760)
Log of Earnings	0.016	0.022	0.021	0.011	0.014
(standard error)	(0.020)	(0.021)	(0.019)	(0.019)	(0.020)
Covariates	No	Yes	Yes	Yes	Yes
Age floor	No	No	Yes	No	No
Age ceiling	No	Yes	No	No	No
Place fixed effects	County	County	County	County	State
Control group	Main	Main	Main	Alt	Main

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: Author's calculations. Inference is conducted using the Ferman and Pinto (2019) method. Each "counterfactual mean" is the average of observed outcomes among individuals exposed to Georgia UPK minus the estimated treatment effect. Relative to the primary model and sample, column I drops all covariates from the regression model; column II restricts the sample to individuals 32 and younger when surveyed; column III restricts the sample to individuals 25 and older when surveyed; column IV limits the control group to states that had no state-funded pre-K program by 1999; and column V replaces birth county fixed effects with birth state fixed effects in the regression model. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701 and CBDRB-FY25-0013). Observation counts are rounded to comply with Census Bureau disclosure avoidance rules.

Figure D.1 — Georgia UPK's Long-Run Effects I (SCM)

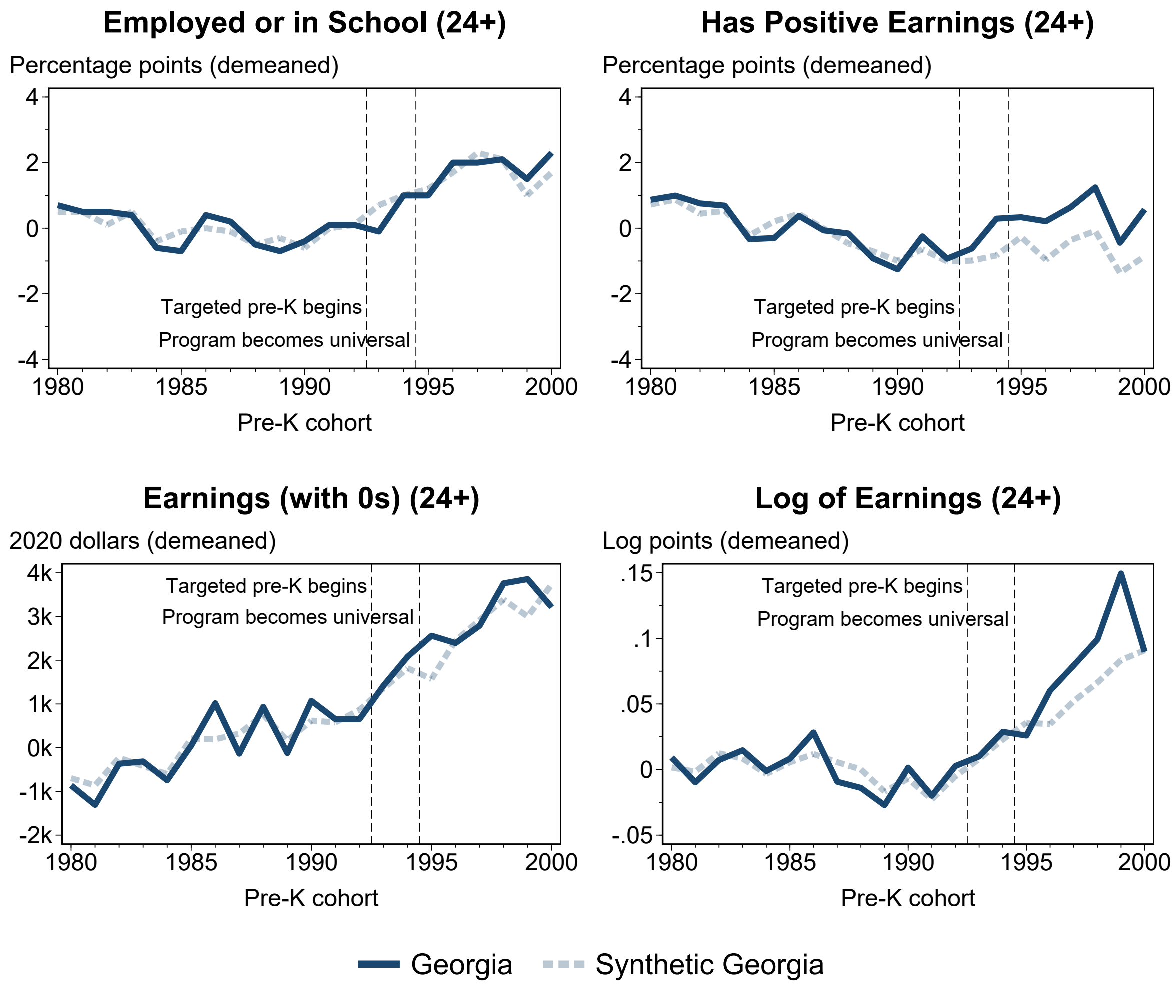


	Teen Mother	Graduated High School	Attained BA Degree	SNAP Receipt
UPK impact	−0.010	0.014*	0.016*	−0.016*
[p-value]	[0.109]	[0.068]	[0.056]	[0.082]

*** p<0.01, ** p<0.05, * p<0.1

Notes: Author's calculations using the synthetic control method, as described in Appendix Section D.2. *P*-values are obtained using the Chernozhukov et al. (2021) conformal inference method. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701).

Figure D.2 — Georgia UPK's Long-Run Effects II (SCM)



	Employed or in School	Has Positive Earnings	Earnings (with 0s)	Log of Earnings
UPK impact	0.002	0.011	251	0.023
[p-value]	[0.358]	[0.585]	[0.982]	[0.190]

*** p<0.01, ** p<0.05, * p<0.1

Notes: Author's calculations using the synthetic control method, as described in Appendix Section D.2. *P*-values are obtained using the Chernozhukov et al. (2021) conformal inference method. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701 and CBDRB-FY25-0013).

Table D.2 — SCM Donor Pool Weights and Pre-Treatment Period Fit

	Teen Mother	Graduated High School	Attained BA Degree	SNAP Receipt
Donor Pool Weights				
Alabama	0.100	0	0	0
Alaska	0	0	0	0
Florida	0.294	0.290	0	0
Hawaii	0	0	0.023	0
Idaho	0	0	0	0
Indiana	0	0	0	0
Iowa	0	0	0	0
Minnesota	0	0	0.406	0
Mississippi	0	0	0.113	0.533
Missouri	0.374	0	0	0
Montana	0	0.057	0	0
Nebraska	0	0	0.056	0
Nevada	0	0.048	0	0
New Hampshire	0.032	0.153	0	0.074
New Mexico	0.037	0.313	0	0.080
North Carolina	0.163	0	0	0
North Dakota	0	0	0	0.157
Pennsylvania	0	0	0.235	0
Rhode Island	0	0.117	0	0
South Dakota	0	0	0	0
Tennessee	0	0.023	0	0
Utah	0	0	0	0.072
Wyoming	0	0	0.168	0.083
Pre-Treatment Period Fit				
RMSE	0.0049	0.0017	0.0034	0.0018

Notes: Author's calculations using the synthetic control method, as described in Appendix Section D.2. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701).

Table D.3 — SCM Donor Pool Weights and Pre-Treatment Period Fit

	Employed or in School	Has Positive Earnings	Earnings (with 0s)	Log of Earnings
Donor Pool Weights				
Alabama	0	0	0.427	0
Alaska	0.022	0.231	0	0.016
Florida	0.282	0	0	0.490
Hawaii	0	0	0	0
Idaho	0	0	0	0.015
Indiana	0	0	0	0
Iowa	0	0	0	0.093
Minnesota	0	0	0	0
Mississippi	0	0.077	0.063	0.242
Missouri	0	0	0	0
Montana	0	0	0	0
Nebraska	0	0	0	0
Nevada	0.029	0	0	0
New Hampshire	0.289	0	0	0
New Mexico	0	0	0	0
North Carolina	0.225	0.274	0.385	0
North Dakota	0	0	0	0
Pennsylvania	0	0	0	0
Rhode Island	0	0.166	0	0.112
South Dakota	0.143	0.092	0	0
Tennessee	0	0	0	0
Utah	0	0	0	0
Wyoming	0.011	0.160	0.125	0.031
Pre-Treatment Period Fit				
RMSE	0.0029	0.0025	347	0.0094

Notes: Author's calculations using the synthetic control method, as described in Appendix Section D.2. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701 and CBDRB-FY25-0013).

Table D.4 — Falsification Tests - Impacts on Predicted Measures (DiD Estimates)

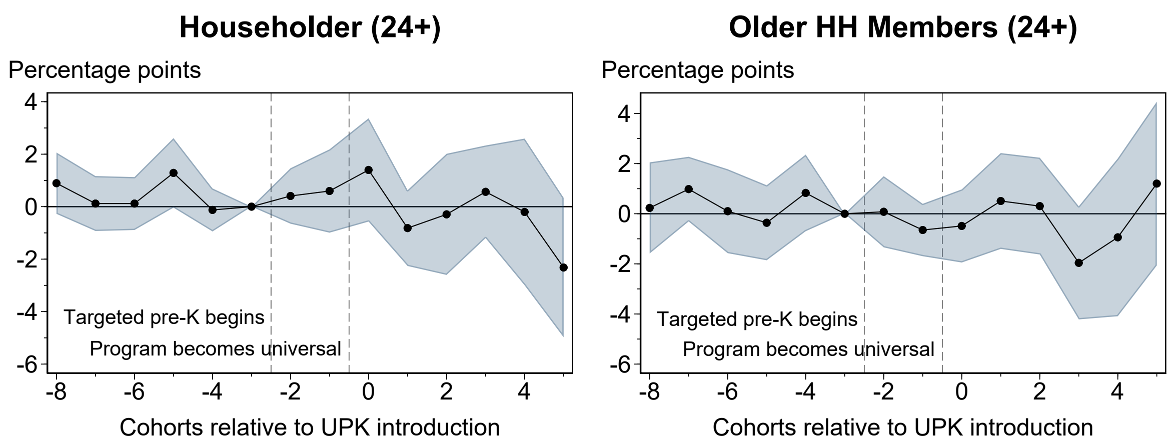
	Teen Father (Predicted)	Teen Mother (Predicted)	Graduated High School (Predicted)	Attained BA Degree (Predicted)	SNAP Receipt (Predicted)
UPK impact	0.000	−0.002	0.000	0.002	−0.004
(standard error)	(0.001)	(0.002)	(0.001)	(0.005)	(0.003)
[p-value]	[0.603]	[0.352]	[0.966]	[0.688]	[0.245]
Counterfactual mean	0.061	0.167	0.899	0.33	0.194
Age restriction	None	None	20+	24+	24+
Observations	2,508,000	2,433,000	2,191,000	1,574,000	1,521,000

	Employed or in School (Predicted)	Has Positive Earnings (Predicted)	Earnings (with 0s) (Predicted)	Log of Earnings (Predicted)
UPK impact	0.002	0.002	304	0.007
(standard error)	(0.002)	(0.002)	(271)	(0.007)
[p-value]	[0.356]	[0.345]	[0.297]	[0.366]
Counterfactual mean	0.775	0.830	27,840	10.02
Age restriction	24+	24+	24+	24+
Observations	1,574,000	1,574,000	1,574,000	1,354,000

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: Author's calculations from estimating equation 2. Inference is conducted using the Ferman and Pinto (2019) method. Each "counterfactual mean" is the average of observed outcomes among individuals exposed to Georgia UPK minus the estimated treatment effect. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701 and CBDRB-FY25-0013). Observation counts are rounded to comply with Census Bureau disclosure avoidance rules.

Figure D.3 — Georgia UPK's Long-Run Effects on Living Arrangements



Notes: Author's calculations from estimating event study versions of equation 2. Inference is conducted using the Ferman and Pinto (2019) method. Cohort 0 refers to the 1995 pre-K cohort. Unlike other analyses in this paper, these are conducted using the public version of the American Community Survey (ACS). The outcome in the first panel is being the designated householder in the ACS. The outcome in the second panel is having a household member who is at least 18 years older.

Appendix E Formalizing the Conceptual Framework

E.1 Setup

Consider an economy with three childcare options: public pre-K (PUB), private pre-K (PRI), and no pre-K (NP). PUB is free to attend, but capacity is limited by a central planner. For each household i, let Z_i be an indicator that equals 1 for households awarded slots in PUB and 0 for households not awarded slots.

Private pre-K providers choose prices and other program characteristics that determine quality to maximize profit. NP providers include some paid arrangements, like home-based care and nannies, so some NP providers also choose prices and quality levels. Let R denote a given regime, which is a collection of all the prices and quality levels in the childcare market.

Households have preferences over childcare arrangements according to the utility function $U_i(j,r)$, where i indexes households, j indexes childcare arrangement, and r indexes regime. Utility depends on both enrollment and regime because some childcare arrangements may be more or less productive at improving human capital, or more or less expensive, in different regimes. (Note that I abstract away from a budget constraint for simplicity.) Households make childcare enrollment decisions D_i to maximize utility given the policy regime R and their individual realization of Z_i:

$$D_i(z,r) \equiv \arg\max_{j \in J} U_i(j,r) \quad \text{with } J = \begin{cases} \{\text{NP, PRI}\} & \text{if } z = 0 \\ \{\text{NP, PRI, PUB}\} & \text{if } z = 1. \end{cases}$$

The choice of childcare enrollment D_i at age four influences children's long-run human capital, Y_i. We can write human capital as $Y_i(D_i(Z_i,R),R)$, or $Y_i(D_i,R)$ for short, because it depends not only on D_i but also on the productivity of D_i in regime R.

A public pre-K expansion is an increase in $E[Z_i = 1]$, the share of households that receive slots in public pre-K. When it expands, private pre-K and NP providers may adjust their price and quality levels. We can then operationalize "equilibrium responses" as a change in policy regime from R to R'. Let Z_i' denote whether a household is awarded a public pre-K slot in regime R'.

E.2 Partitioning Households into Substitution Types

Given a public pre-K expansion, households may be fully partitioned based on the childcare decisions they would make in each regime (given their realizations of Z_i and Z_i'). With three types of childcare arrangements, there are $3^2 = 9$ permutations of childcare arrangements across regimes, or 9 types of substitution households can make. For $A, B \in$ {NP, PRI, PUB}, let $\mathcal{J}$ denote the set of all 9 permutations.

For each $(A,B) \in \mathcal{J}$, let $A \rightarrow B$ be an indicator function that equals 1 for households that enroll in A in regime R and in B in regime R':

$$A \to B \equiv \mathbb{1}\left[D_i(Z_i, R) = A \text{ and } D_i(Z_i', R') = B\right].$$

These are household-level functions, but I omit the i subscript for notational convenience.

Define $\mathrm{S}^{A\to B} \equiv Pr(A \to B)$ as the share of children induced by PUB expansion to substitute from A to B. By the law of iterated expectations, we have:

$$\begin{aligned} &\mathrm{S}^{NP\to NP} + \mathrm{S}^{NP\to PRI} + \mathrm{S}^{NP\to PUB} + \mathrm{S}^{PRI\to NP} + \mathrm{S}^{PRI\to PRI} + \mathrm{S}^{PRI\to PUB} \\ &\quad + \mathrm{S}^{PUB\to NP} + \mathrm{S}^{PUB\to PRI} + \mathrm{S}^{PUB\to PUB} = 1. \end{aligned}$$

Without equilibrium responses in the childcare market, we can eliminate some substitution groups. No equilibrium responses means $R = R'$, so households' rank ordering of childcare arrangements remains constant. The only reason a household would switch programs is if they already preferred PUB and they newly obtain a slot after the PUB expansion. Thus, we have:

$$\mathrm{S}^{NP\to NP} + \mathrm{S}^{NP\to PUB} + \mathrm{S}^{PRI\to PRI} + \mathrm{S}^{PRI\to PUB} + \mathrm{S}^{PUB\to PUB} = 1.$$

E.3 Decomposing the ITT Effect

Recall that an individual's long-run human capital is $Y_i(D_i, R)$. Average human capital in the population is:

$$E\left[Y_i(D_i, R)\right].$$

We can decompose this expectation by household substitution type. By the law of iterated expectations,

$$E\left[Y_i(D_i, R)\right] = \sum_{(A,B)\in\mathcal{J}} E\left[Y_i(A, R) | A \to B\right] \times \mathrm{S}^{A\to B}.$$

Using this expression, the ITT effect of a regime change from R to R' is:

$$\begin{aligned} \text{ITT} &= E\left[Y_i(D_i, R') - Y_i(D_i, R)\right] \\ &= \sum_{(A,B)\in\mathcal{J}} E\left[Y_i(B, R') | A \to B\right] \times \mathrm{S}^{A\to B} - \sum_{(A,B)\in\mathcal{J}} E\left[Y_i(A, R) | A \to B\right] \times \mathrm{S}^{A\to B} \\ &= \sum_{(A,B)\in\mathcal{J}} E\left[Y_i(B, R') - Y_i(A, R) | A \to B\right] \times \mathrm{S}^{A\to B}. \end{aligned}$$

Defining $\delta^{A\to B} \equiv E\left[Y_i(B, R') - Y_i(A, R) | A \to B\right]$, we can write the ITT effect more succintly as:

$$\text{ITT} = \sum_{(A,B)\in\mathcal{J}} \delta^{A\to B} \times \mathrm{S}^{A\to B}.$$

Finally, it is useful for intuition to expand the summation. In this case, for conciseness, we can define $\Delta^{A\to B} \equiv \delta^{A\to B} \times \mathrm{S}^{A\to B}$ and write the ITT as:

$$\begin{aligned}\mathrm{ITT} = {} & \Delta^{PRI\to PUB} + \Delta^{NP\to PUB} \\ & + \Delta^{PRI\to PRI} + \Delta^{NP\to NP} + \Delta^{PUB\to PUB} \\ & + \Delta^{PUB\to PRI} + \Delta^{PUB\to NP} + \Delta^{PRI\to NP} + \Delta^{NP\to PRI},\end{aligned}$$

which is equation 4 in the paper.

Without equilibrium responses, this equation simplifies considerably. The equilibrium-driven substitution terms ($\Delta^{PUB\to PRI}$, $\Delta^{PUB\to NP}$, $\Delta^{PRI\to NP}$, and $\Delta^{NP\to PRI}$) drop out because no household engages in these substitutions. Moreover, if the regime R is held constant, there are no quality changes, so we have $\delta^{PRI\to PRI} = \delta^{NP\to NP} = \delta^{PUB\to PUB} = 0$. Thus, the ITT effect simplifies to:

$$\mathrm{ITT} = \Delta^{PRI\to PUB} + \Delta^{NP\to PUB},$$

which is equation 3 in the paper.

Appendix F Substitution Analysis Details

F.1 Estimation

My analysis follows a two-step procedure. First, using just the control group, I fit a multinomial logit model with enrollment in no pre-K, private pre-K, and public pre-K as the outcome variables. The predictor variables are children's individual and household characteristics, as well as county variables that are highly predictive of the size of one's local pre-K market. The individual and household variables are children's sex and race; age in months; whether a child attended pre-K in the previous year; presence of siblings in the household; presence of a grandparent in the household; maternal age and educational attainment; and paternal educational attainment, employment, and income. The county-level characteristics—measured in the 1990 census—are overall pre-K enrollment rate among four- and five-year-olds, median household income, metropolitan status, Black share, educational attainment among those age 25+, and the share of children ages 0-5 in single mother households.

In the second step, I use the logit results to obtained predicted values for each child in Georgia. Summing the predicted probability of enrolling in A across all children enrolled in B, and normalizing by the population, gives the final estimate for $S^{A \to B}$. By construction, the nine $\hat{S}^{A \to B}$ estimates sum to 1. As one example, to estimate $S^{PRI \to PUB}$ I sum the predicted probability of enrolling in private pre-K across all pre-K-age children in Georgia who actually enrolled in public pre-K. Then I divide that sum by the total number of pre-K-age children in Georgia.

F.2 Diagnostics

The assumptions necessary for estimating substitution cannot be tested directly. The results should therefore be interpreted with some caution. However, I find supportive evidence from two validity checks.

The first validity check assesses the predictiveness of the logit model within the control group as a whole and then in the states most similar to Georgia. If the model predicts well in those states, we can be more confident it predicts counterfactual enrollment in Georgia. In a series of tests, I split the control group into a training sample and a validation sample. I use the following three validation samples: (i) a random half of the control group; (ii) Alabama; and (iii) the joint set of southeastern control group states (Alabama, Mississippi, Florida, and North Carolina). After estimating the model on a training sample, I obtain predicted probabilities for the validation sample.

Table F.1 shows that the model performs very well for each validation sample; children tend to enroll in each program at rates very similar to their predicted probabilities. This test is only suggestive because I can't assess predictiveness conditional on (unobserved) enrollment in a world with UPK, but it is nevertheless encouraging.

The second validity check assesses predictiveness specifically in Georgia. I use the well-identified DiD estimates from Section 3 to produce a benchmark for the logit model predictions.

Subtracting the DiD enrollment effects from Georgia's observed (state-level) enrollment rates, I estimate that enrollment would have been 27.6% in no pre-K, 26.5% in private pre-K, and 45.9% in public pre-K in a Georgia without UPK.

I then estimate those same counterfactual enrollment rates using the substitution analysis methodology. After estimating the multinomial logit model on the full control group, I predict counterfactual probabilities for children in Georgia and aggregate them to the state level. Reassuringly, the rates are very similar to those implied by the DiD estimates: 28% vs. 27.7% for no pre-K, 28% vs. 26.2% for private pre-K, and 44% vs. 46.1% for public pre-K.

Table F.1 — Substitution Model Diagnostics

	No Pre-K		Private Pre-K		Public Pre-K	
Predicted Probability	Share Enrolled	N	Share Enrolled	N	Share Enrolled	N
	Panel A. Validation Sample = Half of Control Group					
$[0.0, 0.1)$	0.07	20,000	0.06	12,500	0.10	250
$[0.1, 0.2)$	0.10	3,600	0.14	16,000	0.17	5,600
$[0.2, 0.3)$	0.22	2,300	0.24	11,000	0.25	16,500
$[0.3, 0.4)$	0.34	6,900	0.34	7,400	0.35	13,000
$[0.4, 0.5)$	0.45	11,500	0.47	5,300	0.45	7,400
$[0.5, 0.6)$	0.56	10,500	0.55	3,800	0.55	6,000
$[0.6, 0.7)$	0.66	4,900	0.62	2,800	0.67	4,600
$[0.7, 0.8)$	0.77	850	0.75	1,700	0.76	3,900
$[0.8, 0.9)$	Omitted	N<100	0.82	600	0.84	2,900
$[0.9, 1.0]$			Omitted	N<100	0.89	400
	Panel B. Validation Sample = Alabama					
$[0.0, 0.1)$	0.07	2,900	0.06	2,300	Omitted	N<100
$[0.1, 0.2)$	0.11	650	0.14	2,500	0.15	1,000
$[0.2, 0.3)$	0.26	450	0.24	1,500	0.22	2,700
$[0.3, 0.4)$	0.36	1,300	0.38	1,000	0.28	1,400
$[0.4, 0.5)$	0.49	1,600	0.48	750	0.44	1,300
$[0.5, 0.6)$	0.62	1,400	0.56	550	0.51	1,100
$[0.6, 0.7)$	0.66	1,000	0.63	400	0.64	700
$[0.7, 0.8)$	0.78	200	0.79	300	0.71	550
$[0.8, 0.9)$			0.86	150	0.83	550
$[0.9, 1.0]$					Omitted	N<100
	Panel C. Validation Sample = Southeastern States					
$[0.0, 0.1)$	0.07	20,500	0.10	15,000	0.12	300
$[0.1, 0.2)$	0.11	3,100	0.15	14,500	0.17	5,900
$[0.2, 0.3)$	0.27	2,500	0.26	9,100	0.25	14,500
$[0.3, 0.4)$	0.34	5,100	0.39	6,200	0.36	12,000
$[0.4, 0.5)$	0.41	9,500	0.50	4,600	0.46	6,600
$[0.5, 0.6)$	0.52	10,500	0.59	3,500	0.51	4,200
$[0.6, 0.7)$	0.61	5,500	0.69	2,600	0.58	4,800
$[0.7, 0.8)$	0.77	1,100	0.78	1,700	0.72	4,900
$[0.8, 0.9)$	Omitted	N<100	0.84	550	0.80	2,800
$[0.9, 1.0]$			Omitted	N<100	0.62	1,300

Notes: Author's calculations from training multinomial logit models on training samples and making predictions on validation samples, as described in Section 6.1. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701). Observation counts are rounded to comply with Census Bureau disclosure avoidance rules.

Appendix G Hypothetical Impact Matrix Scenarios

In this section I provide detail on the construction of the six scenarios used in Section 6.2. Each scenario is depicted in full in Figure G.1. All six scenarios are consistent with 1) my estimate of UPK's substitution matrix, and 2) my estimate of the ITT effect on high school graduation. They are consistent in the sense that a weighted sum of the sub-impacts, with the substitution estimates as weights, equals the high school graduation ITT estimate (1.5pp). Based on previous Georgia UPK research, and given the positive ITT effects in this paper, I focus on scenarios in which children who substitute from no pre-K to some form of pre-K have non-negative impacts. Further, in each scenario I make the simplifying assumption that the off-diagonal cells in the impact matrix are symmetric but opposite in sign. In other words, if moving from A to B increases a child's likelihood of graduating high school by Xpp, then moving from B to A decreases their likelihood by Xpp.

Scenario I. The first matrix represents a benchmark case. I assume private and public pre-K are equally effective (in the sense that $\delta^{NP \to PRI} = \delta^{NP \to PUB}$ and $\delta^{PRI \to PUB} = 0$). I also assume program stayers experience no impacts. Given these assumptions, the entire matrix is identified and one can back out that the non-zero impacts are all 13.6pp. An impact of 13.6pp is possible but somewhat large relative to the literature.

Scenario II. The second matrix relaxes the assumptions of the first matrix by allowing for small impacts on public and private pre-K program stayers. Given evidence from Henry and Gordon (2006) that increased competition made pre-K programs in Georgia more effective, I assume private pre-K stayers experienced a 1pp impact. I assume a slightly larger but still small impact (2pp) for public pre-K stayers given evidence that UPK had greater positive short-run impacts on children in Georgia than Head Start (Henry et al., 2006) and given evidence from Cascio (2023) that universal programs may have greater effects for low-income children than targeted programs. Continuing with the assumptions of $\delta^{NP \to PRI} = \delta^{NP \to PUB}$ and $\delta^{PRI \to PUB} = 0$, the entire impact matrix is identified and the values on the off-diagonals can be backed out. Relative to Scenario I, the inclusion of program stayer impacts causes the off-diagonal impacts to decrease in absolute value. Intuitively, program stayers now account for some of the overall ITT effect, so impacts on non-stayers must be smaller to keep the ITT effect constant.

Scenario III. The third matrix relaxes the assumption in the previous two matrices that private and public pre-K are equally effective. Instead, I assume private pre-K is 20% more effective, in the sense that $\delta^{NP \to PRI} = 1.2 * \delta^{NP \to PUB}$ and $\delta^{PRI \to PUB} = -0.2 * \delta^{NP \to PUB}$. Continuing with the program stayer impacts from Scenario II, the entire impact matrix is identified and the values on the off-diagonals can be backed out. Relative to Scenario II, the impacts for switching between no pre-K and private pre-K increase, and the impacts for switching between no pre-K and public pre-K decrease.

Scenario IV. The fourth matrix examines the effect of allowing private and public pre-K to have different impacts when there are no impacts on program stayers. As in Scenario III, I assume private pre-K is 20% more effective than public pre-K in the sense that $\delta^{NP \to PRI} = 1.2 * \delta^{NP \to PUB}$

and $\delta^{PRI \to PUB} = -0.2 * \delta^{NP \to PUB}$. These assumptions identify the entire matrix. The impacts on children who switch from no pre-K to private pre-K or public pre-K are quite large. However, the implied local average treatment effect for those who switch from no pre-K or private pre-K to public pre-K (6.2pp) is very normal for the literature.

Scenario V. The fifth matrix reintroduces small impacts for program stayers but assumes public pre-K is more effective than private pre-K (in the sense that $\delta^{NP \to PUB} = 1.2 * \delta^{NP \to PRI}$ and $\delta^{PRI \to PUB} = 0.2 * \delta^{NP \to PRI}$). As in Scenarios II and III, the impacts of switching into private or public pre-K are smaller than in Scenarios I and IV because program stayers account for some of UPK's overall impact.

Scenario VI. The sixth matrix considers a situation in which private and public program stayers experience small negative impacts. To isolate the effect of this assumption, I assume private and public pre-K are equally effective. Given the negative impacts in the diagonal cells, the impacts in the off-diagonal cells become larger in magnitude to keep the aggregate ITT effect constant. As in Scenario IV, the off-diagonal magnitudes are quite large, but the implied local average treatment effect for public pre-K compliers is reasonable. Larger negative impacts for program stayers would require larger impacts for substitution into private and public pre-K, which may be difficult to believe.

Figure G.1 — Six Hypothetical Impact Matrices for Georgia UPK

Scenario I

Without UPK \ With UPK	NP	PRI	PUB
NP	0.0pp	13.6pp	13.6pp
PRI	-13.6pp	0.0pp	0.0pp
PUB	-13.6pp	0.0pp	0.0pp

Scenario II

Without UPK \ With UPK	NP	PRI	PUB
NP	0.0pp	7.2pp	7.2pp
PRI	-7.2pp	1.0pp	0.0pp
PUB	-7.2pp	0.0pp	2.0pp

Scenario III

Without UPK \ With UPK	NP	PRI	PUB
NP	0.0pp	9.3pp	7.7pp
PRI	-9.3pp	1.0pp	-1.5pp
PUB	-7.7pp	1.5pp	2.0pp

Scenario IV

Without UPK \ With UPK	NP	PRI	PUB
NP	0.0pp	17.6pp	14.7pp
PRI	-17.6pp	0.0pp	-2.9pp
PUB	-14.7pp	2.9pp	0.0pp

Scenario V

Without UPK \ With UPK	NP	PRI	PUB
NP	0.0pp	5.6pp	6.8pp
PRI	-5.6pp	1.0pp	1.1pp
PUB	-6.8pp	-1.1pp	2.0pp

Scenario VI

Without UPK \ With UPK	NP	PRI	PUB
NP	0.0pp	17.4pp	17.4pp
PRI	-17.4pp	-1.0pp	0.0pp
PUB	-17.4pp	0.0pp	-1.0pp

Notes: This figure depicts six hypothetical scenarios that are consistent with my estimates for UPK's ITT effect on high school graduation and UPK's substitution matrix. In each matrix, each cell gives the average impact for children of a particular substitution type. For example, the top right cell of each matrix gives the average impact for children in Georgia who substituted from no pre-K to public pre-K because of UPK. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701).

Appendix H Impact of Georgia UPK on Head Start Enrollment

In this section, I estimate the impact of Georgia UPK's introduction on Head Start enrollment in Georgia. To do so, I use data on Head Start enrollment from publicly available Head Start Program Information Reports (PIR). I aggregate the PIR enrollment counts to the state level and divide by the number of four-year-olds in each states to obtain enrollment rates. As in Section 3, I obtain data on the number of four-year-olds in each state in each year from the Surveillance, Epidemiology, and End Results Program at the National Cancer Institute.

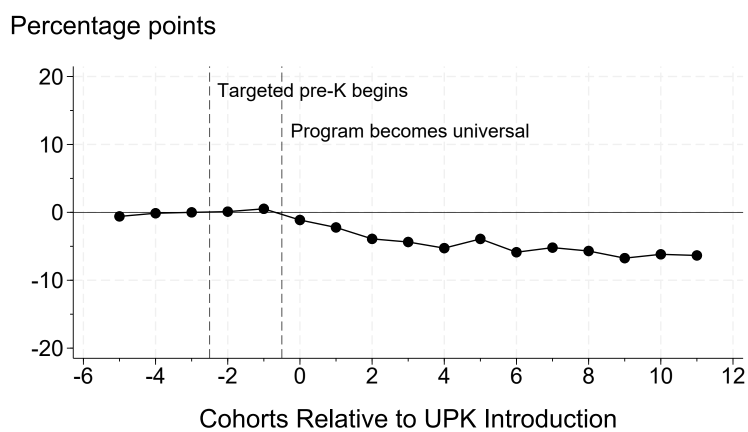
I then conduct an event study analysis in which Head Start enrollment is the outcome variable. As in the previous enrollment analyses, the control group is the set of 18 states with no state-funded pre-K program in 1999. I estimate the following model:

$$HS_{st} = \sum_{\substack{e=-5 \\ e\neq -3}}^{11} \beta^e GA^e + \lambda_s + \gamma_t + \varepsilon_{st}, \tag{H.1}$$

where HS_{st} is enrollment in Head Start for state s and cohort t; λ_s and γ_t are state and cohort fixed effects, respectively; and GA^e is a binary indicator that equals 1 for living in Georgia and being e cohorts away from the 1995 pre-K cohort.

The results are shown in Figure H.1. The point estimates are very close to zero in the pre-treatment period, and then begin to decrease after UPK is introduced. When I pool the 1995-2000 cohorts together and estimate a DiD model, I find that UPK reduced Head Start enrollment by 3.4 percentage points, which is a 25% decrease off a counterfactual base of 13.8%.

Figure H.1 — Effect on Georgia UPK on Head Start Enrollment



Notes: Author's calculations using data from the Head Start Program Information Reports.

Appendix I Details on Cost-Benefit Analysis

A general equation for Georgia UPK's MVPF could be written as:

$$\text{MVPF} = \frac{\rho(1-\tau)\text{ITT}^{Y} + \rho\text{ITT}^{\text{Ben}} - \text{ITT}^{p}}{\varphi^{\text{UPK}}\text{NC}^{\text{UPK}} + \text{ITT}^{\text{Sub}} - \rho\tau\text{ITT}^{Y} + \rho\text{ITT}^{\text{Ben}}}. \tag{I.1}$$

In the numerator, the first term is the present discounted value of UPK's effect on long-run earnings. ITT^{Y} is UPK's effect on earnings, τ is a marginal tax rate, and ρ is a discount factor. The second term, $\beta\text{ITT}^{\text{Ben}}$, is the (discounted) effect of UPK on long-run government benefit receipt. The final term, ITT^{p}, is the overall effect of UPK on households' childcare expenditures—accounting implicitly for all substitution between programs.

The first term in the denominator is the cost to government of providing UPK. NC^{UPK} is the net change in UPK enrollment (which is simply program take-up), and φ^{UPK} the cost per child (which is publicly reported). The second term, ITT^{sub}, is the overall change in the cost to government of subsidizing non-UPK childcare arrangements. As children substitute between programs, government spending on Head Start, local public programs, and childcare subsidies for low-income families in private programs may change. Lastly, $\rho\tau\text{ITT}^{Y}$ and ρITT^{Ben} represent long-run effects on government tax receipt and spending.

Note that there is no strong evidence that Georgia UPK raised parents' earnings. Positive impacts on parents' earnings would raise program benefits and lower costs via additional tax collections. In this case, this MVPF formula would be overly conservative.

In practice, I do not have the data to directly estimate each term in equation I.1. Instead, my estimable analog is:

$$\text{MVPF} = \frac{\rho(1-\tau)\text{ITT}^{Y} + \rho\varphi^{\text{SNAP}}\text{ITT}^{\text{SNAP}} - \kappa^{\text{PRI}}\text{NC}^{\text{PRI}}}{\varphi^{\text{PUB}}\text{NC}^{\text{PUB}} + \varphi^{\text{HS}}\text{NC}^{\text{HS}} + \varphi^{\text{PRI}}\text{NC}^{\text{PRI}} - \rho\tau\text{ITT}^{Y} + \rho\varphi^{\text{SNAP}}\text{ITT}^{\text{SNAP}}}. \tag{I.2}$$

I replace $\rho\text{ITT}^{\text{Ben}}$ with $\varphi^{\text{SNAP}}\text{ITT}^{\text{SNAP}}$, where ITT^{SNAP} is the net change in the share of people receiving SNAP, and φ^{SNAP} is the average benefit amount. I replace ITT^{p} with $\kappa^{\text{PRI}}\text{NC}^{\text{PRI}}$, where NC^{PRI} is the net change in private pre-K enrollment, and κ^{PRI} is average household spending on private pre-K. In the denominator, I replace $\varphi^{\text{UPK}}\text{NC}^{\text{UPK}} + \text{ITT}^{\text{sub}}$ with the net change in Head Start (HS), non-Head Start public pre-K (PUB), and private pre-K (PRI) enrollment multiplied by their average per-child cost to government (φ^{HS}, φ^{PUB}, and φ^{PRI}). I separate Head Start from other public pre-K because it is notably more expensive, and I assume UPK and local public programs have equal costs.

I obtain values for the inputs to equation I.2 from estimates in this paper and external sources, following Cascio (2023) where possible. Each parameter is presented and explained in Table I.1.

For long-run earnings and SNAP benefits, I discount from age 27 to age 4 because the average age of treated cohorts in my analysis is 27 (for these outcomes). I also inflation-adjust the long-run outcomes to 1998 dollars to reflect the middle of the period I examine. Lastly, I multiply these annual impacts by 6.5 because I observe each treated cohort for an average of 6.5 years (for these outcomes) and the annual impacts presumably last for more than one year.

The enrollment analysis in Section 3 uses census data, which does not distinguish between enrollment in Head Start and other public pre-K programs. However, Head Start costs more per child than Georgia UPK. For the 1995-2000 pre-K cohorts, the average cost of Georgia UPK was \$3,678, compared to \$6,358 for Head Start in Georgia in 2001 (in 1998 dollars). To account for this difference, I estimate the net effect of UPK's introduction on Head Start enrollment in Section Appendix H. This net effect is the relevant parameter for estimating the reduction in spending on Head Start, but estimating the net effect on spending for other public pre-K programs is more challenging. The net increase in public pre-K enrollment, estimated in Section 3, conceals substitution from Head Start to UPK. To be conservative, I add the entire Head Start estimate to the overall public pre-K estimate. Relatedly, in the absence of good data on the cost to government of subsidizing private pre-K for four-year-olds in Georgia, I take the conservative approach of assuming zero cost savings from reductions in these subsidies.

Table I.1 — Inputs to Georgia UPK's MVPF Estimate

Description	Symbol	Estimate	Source/Note
Discount factor	ρ	0.51	Following Cascio (2023), I use a discount rate of 3%. I discount from age 27 to 4.
Marginal tax rate	τ	0.2	Following Cascio (2023)
ITT effect on earnings	ITT^{Y}	$4,118$	Earnings estimate from Table 4 ($1,006), inflation-adjusted to 1998 dollars, and multiplied by 6.5 because impacts last multiple years.
ITT effect on SNAP receipt	ITT^{SNAP}	-0.019	Table 3
Average SNAP benefit	φ^{SNAP}	$12,101$	Average monthly household benefit amount in 2018 ($239), converted to annual, inflation-adjusted to 1998 dollars, and multiplied by 6.5 because impacts last multiple years. Source: USDA.
Cost to gov't of non-HS public pre-K	φ^{PUB}	$3,678$	Average cost of Georgia UPK from 1995-2000. Source: Georgia DECAL.
Cost to gov't of Head Start	φ^{HS}	$6,358$	Average cost of Head Start in 2001 in Georgia, inflation-adjusted to 1998 dollars. Source: NIEER (2003).
Cost to gov't of private pre-K	φ^{PRI}	0	Assuming 0 yields a lower bound.
Cost to families of private pre-K	κ^{PRI}	$4,966$	Following Cascio (2023), lower bound estimate from Laughlin (2013).
Net change in non-HS public pre-K enrollment	NC^{PUB}	0.163	Net change in public pre-K from Table 1 plus change in HS enrollment from Appendix H.
Net change in Head Start enrollment	NC^{HS}	-0.034	Appendix H
Net change in private pre-K enrollment	NC^{PRI}	-0.032	Table 1

Notes: Inflation adjustments are made using the consumer price index for all urban consumers (CPI-U), obtained from the Federal Reserve Bank of Minneapolis. These estimates have been approved for release by the U.S. Census Bureau (CBDRB-FY24-P3054-R11701).