

Does Public Pre-K Have Unintended Consequences on the Child Care Market for Infants and Toddlers?*

Jessica H. Brown [†]

December 30, 2019

Abstract

I estimate the impact of public pre-kindergarten for 4-year-olds on the provision of private child care for younger children by considering New York City's 2014 Universal Pre-K expansion. Private child care facilities often care for children from infancy or toddlerhood through pre-K. A public option for older children could therefore affect availability, prices, or quality of care for younger children. This effect could be positive or negative depending on the structure of the child care market, the design of the public pre-K program, and parent preferences. I use a panel dataset covering all licensed child care facilities in New York City and a difference-in-differences strategy that compares changes over time for neighborhoods with more versus fewer new public pre-K sites. I estimate that the public pre-K program reduced the capacity for children younger than 2 years old at private child care centers by 2,700 seats. The entire decrease in capacity occurs in areas with high poverty, and this decline was not offset by an increase in provision in the home day care market. In complementary analysis, I find a within-center increase in public complaints and inspection violations for day care centers that are closer to new public pre-K sites, suggesting a decrease in quality due to the increased competition from public pre-K. A back-of-the-envelope calculation indicates that for every seven 4-year-olds who shifted from day care centers to public pre-K, there was a reduction of one day care center seat for children under the age of 2.

JEL Classification: H44, H75, I21, I28, J13

Keywords: Child Care, Early Childhood Education, Education Policy.

*I am grateful to my advisor, Alan Krueger, for his guidance and support throughout this project. I am also grateful to Leah Boustan and Will Dobbie for their advice and encouragement. I benefited from helpful comments from Janet Currie, Bill Evans, Chloe Gibbs, Felipe Goncalves, Bo Honoré, Adam Kapor, Ilyana Kuziemko, Alex Mas, Steve Mello, and seminar participants at the University of Notre Dame and Princeton University. All errors are my own.

[†]Department of Economics, Industrial Relations Section, Louis A. Simpson International Building, Princeton University, Princeton, NJ 08544. Email: jbrown2@princeton.edu

1 Introduction

In the past 15 years, enrollment in public pre-kindergarten programs has nearly doubled. By 2016-17, state-funded pre-K served about 1.5 million children, enrolling over 30 percent of 4-year-olds and accounting for over \$7.6 billion in spending (Friedman-Krauss *et al.*, 2017). Expansions are ongoing and remain part of conversations at the national level, with President Obama’s “Preschool for All” initiative, and the local level, with recent expansions in Philadelphia, Seattle, and Albuquerque. The expansions are driven by research on the benefits of quality pre-kindergarten, particularly for disadvantaged children, and hopes of closing achievement gaps already present when children enter kindergarten.

Although free public pre-K expands access to early childhood education, it also draws some children to the public sector who would have attended private pre-K otherwise (Cascio and Schanzenbach 2013; Fitzpatrick 2010). On the private market, these 4-year-olds are often cared for in facilities that also care for younger children. Crowd out due to the public expansion could therefore affect the availability, price, and/or quality of child care for younger children between the ages of 0 and 3. García *et al.* (2018) show that investments in human capital development have larger returns for younger ages, suggesting that care environments for children under age 4 may be even more influential than at age 4. Changes in the child care market could thus have a significant impact on the human capital accumulation of affected children.

Without empirical evidence, it is not clear whether public pre-K programs would increase or decrease availability of private child care for younger children. If 4-year-olds shift to the public market from the private market, private day care facilities may be negatively affected by the lost revenue. Reduced 4-year-old enrollment may be especially problematic for the private sector because centers typically cross-subsidize the very labor-intensive and costly care of infants and toddlers with tuition dollars from less costly older children.¹ If the most profitable children leave, day care centers may need to raise prices, shut down, or reduce quality.

On the other hand, if the supply of buildings suitable for child care is relatively inelastic, if 4-year-olds shift out of private care, it could free up classrooms that could be converted

¹For example, in New York City, licensing requirements allow one teacher to care for twelve 4-year-olds but only four infants. Day cares charge more for infants but not enough to make up for the cost difference. For a more detailed explanation, see Section 3.

for use by younger children, which could increase availability of care for younger children. Additionally, because public pre-K only operates during school hours, it may increase demand for before- and after-school care. The government’s emphasis on quality early childhood education could also increase parents’ demand for quality child care at younger ages, shifting their demand curve through a “crowd in” effect.²

Center-based care is often the highest-quality form of non-parental care available, particularly for children from lower-income families, so a reduction in availability would be likely to shift children to lower-quality care at a critical time for human capital development (Bassok *et al.* 2016a; Bernal and Keane 2010; Loeb *et al.* 2007). A decrease in availability or increase in price of child care could also decrease female labor force participation (Cascio 2009b; Baker *et al.* 2008; Gelbach 2002; Berger and Black 1992).

This paper uses a quasi-experiment created by New York City’s 2014 Universal Pre-K (UPK) expansion to estimate the effect of new public pre-K sites on availability and quality of care for younger children. In 2014, NYC added over 25,000 new free, full-day public pre-K seats and increased the number of pre-K sites by over 400 on net. The spatial variation in the location of new sites leads some neighborhoods to be more exposed to the expansion than others.³ To exploit this variation in exposure, I use a difference-in-differences estimator to compare changes in capacity and quality in the child care market before and after 2014 in areas that were more exposed to UPK to areas that were less exposed.

New York City is an ideal setting to examine this question for several reasons. First, due to the large scale of the roll out, there are over 400 new site “experiments” to analyze, many more than would exist in a typical city. Relatedly, the scope of the program, which enrolls 70,000 students each year, means the effects of the program are independently interesting. To give a sense of scale, there are more 4-year-olds in public pre-K in NYC than there are students in the entire Boston Public School system. Second, the structure of the New York

²There are a number of other forces that could generate positive or negative effects on the child care market for younger children. These include general equilibrium labor market effects and parental income effects. Also, many public pre-K programs partner with day care centers to provide some of the pre-K classroom space, which could induce entry or reduce exit of the centers. See Section 3 for additional discussion.

³Parents prefer to send their children to schools that are closer compared to ones that are farther away, and this preference is especially strong for the youngest children (Dinerstein and Smith, 2014). Therefore, the distribution of sites across the city leads to differential exposure to UPK across neighborhoods.

City licensure data allows me to specifically identify the effect of the roll out on younger children, in contrast to data available for previous studies that only allowed for analysis of the market summed across all ages (Bassok *et al.* 2014; Bassok *et al.* 2016b). NYC requires separate licenses for classrooms with children under the age of 2 years and classrooms with children ages 2 years to 5 years, so I am able to observe and analyze capacity separately for this younger age group. Finally, NYC is in the process of rolling out UPK for 3-year-olds. Analyzing the response of the child care market to the first expansion is an important step in predicting what might happen to the child care market in light of the additional expansion.

In order to take advantage of spatial variation in exposure, it is necessary to define a relevant market for each UPK. In addition, because the outcome of interest is the stock of day care centers, it is necessary to define areas over which to measure the stock. This area could be the same or different from the market for the UPK. There is no standard reduced-form approach for defining this area or the market for a UPK in this type of setting.⁴ I approach this problem by dividing NYC into a grid of tessellated hexagons and allowing the child care market in a given hexagon to be affected by new UPK sites in that hexagon and in neighboring hexagons.⁵ There are several advantages to this approach. It is able to handle the spatial density of treatment by aggregating the number of sites to the hexagon level. Also, allowing UPK sites to affect neighboring hexagons implicitly allows child care markets to overlap. Furthermore, the method allows for straightforward robustness checks of the sensitivity of results to geographic boundaries through changing the size of the hexagons and shifting the grid. Sensitivity to boundaries is something that is often ignored but has been shown to be important in other contexts (Foote *et al.*, 2017).

I have three sets of findings. First, the expansion of UPK decreased availability of center-based care for infants and toddlers. A new UPK site reduces capacity at day care centers for children under the age of 2 by 1.6 seats in an average hexagon with an area of ten blocks. Overall, I estimate that there would be 20% more seats in day care centers for children under

⁴Two approaches for defining markets in the literature include the “ring method” and using pre-existing administrative boundaries. The ring method is not preferred here because the spatial density of treatment causes the rings to overlap substantially. Using administrative boundaries often leads to areas that are irregularly-shaped, and there is no natural way to test robustness to them. See Section 5 for more details.

⁵Hexagons are preferred to squares because all of a hexagon’s neighbors have the same relationship to the central hexagon, making it more natural to pull in “neighbor effects.” Squares have both same-side and diagonal neighbors.

age 2 in the absence of the Universal Pre-K program. This decrease in child care center seats for younger children does not seem to be offset by an increase in capacity in licensed home day care centers. However, the analysis of the home day care market can only be suggestive as I do not observe the actual age distribution of the children cared for in this market.

Second, heterogeneity analysis reveals that the entire decline in day care center capacity occurs in poorer areas of the city. I split the hexagons into two groups based on the median fraction of people living below 200% of the federal poverty line and find no effect on total capacity in the wealthier hexagons. A likely explanation for the result that the decrease in capacity is entirely in poorer areas is that the price elasticity of demand for child care is higher in poorer areas. Now the effect of the decreased capacity for younger children on total welfare depends critically on parents' outside option for child care, and in poorer areas, children are likely crowded out to home day cares or informal family and friends care. Based on previous literature, children crowded out of the center-based market to home day cares or other informal settings are likely to experience environments that are less conducive to human capital development (Bassok *et al.* 2016a; Bernal and Keane 2010; Loeb *et al.* 2007). Given that returns to human capital investment are generally the highest at younger ages (García, Heckman, Leaf and Prados, 2018), this change could harm their future outcomes.

Third, I analyze inspections violations as a proxy for safety and quality changes. Some of the most common violations include failing to conduct on-time background checks, missing staff or child medical clearances, and not maintaining the physical facility. When a day care faces pressure due to a loss of 4-year-olds, they may attempt to reduce their costs by lowering quality. I find that within-center, the number of inspections violations increases by 13% at day care centers that share a 10-block area hexagon with a new UPK site. They are also more likely to do poorly enough on the initial inspection to trigger a follow-up inspection. Public complaints against centers near new UPK sites also increased, suggesting that these changes in inspection outcomes point to observable decreases in quality. These results suggest that centers are reducing costs as a result of the competition from public pre-K in a way that affects the quality of care available on the private market.

These empirical results build on the literature studying crowd out in early childhood education. So far, the literature has focused on the direct crowd-out effect of 4-year-olds switching from private pre-K to public pre-K. Crowd out is of interest because it lowers the expected return to public pre-K. Significant crowd out of private pre-K attendance has been

documented for some demographic groups by [Cascio and Schanzenbach \(2013\)](#). Additionally, [Fitzpatrick \(2010\)](#) finds that the total increase in pre-K attendance in response to Universal Pre-K expansions is smaller in urban areas, suggesting that there is more crowd out in areas like NYC where there are more potential private market substitutes.⁶ I contribute to this literature by expanding the discussion to younger children who may be affected by public pre-K's crowd out of private child care facilities. Just as direct crowd out of 4-year-olds lowers the expected return to public pre-K, if public pre-K causes younger children to be crowded out to lower quality care, it would also lower the expected return.

Most closely related to my work, there are two other studies that estimate the effect of public pre-K on the supply of child care. [Bassok, Fitzpatrick and Loeb \(2014\)](#) and [Bassok, Miller and Galdo \(2016b\)](#) find that public pre-K expands, or at least holds fixed, the size of the child care market. This net expansion is likely due to an increase in the number of seats for 4-year-olds due to the new program, and data available for these previous studies did not allow for disaggregating the change in provision by age group. Considering the total capacity across all ages potentially masks decreases in capacity at younger ages caused by public pre-K expansions. Although they could not separately estimate supply effects for younger children, using state-level ACS data, [Bassok, Miller and Galdo \(2016b\)](#) do estimate that reported preschool attendance for 3-year-olds declines, indicating that separating increases in provision for 4-year-olds from changes in provision at younger ages is important.

I advance this research in three ways. First, my data allow me to separately analyze the effect of public pre-K on child care supply for younger children between the ages of 0 and 2. I am the first to study the unintended consequences of UPK on this non-targeted age group. I also suggest additional mechanisms through which this effect may operate. Second, my identification strategy allows me to examine the sizeable heterogeneity in the supply response by the level of poverty in the neighborhood. Understanding this heterogeneity is important for understanding welfare implications because the next best option for care will be much different for wealthier families than for poorer families. Third, I am the first to consider

⁶There are additional studies of crowd out in early childhood education more broadly. For example, when public kindergarten was introduced to public schools, it crowded out enrollment in private kindergarten and Head Start ([Cascio, 2009a](#)). [Bassok \(2012\)](#) studies the crowd out effect of public pre-K on another public alternative, Head Start, and finds that public pre-K reduced Head Start enrollment for 4-year-olds and increased enrollment for younger children. [Kline and Walters \(2016\)](#) study substitutes for Head Start and develop a framework for evaluating the effect of public programs with close substitutes.

the unintended impact of UPK on the safety and quality of child care supplied in the private sector. For this analysis, I use longitudinal administrative inspections and complaints data, and to my knowledge, I am the first to use data of this kind in any context.⁷

I also add to a small literature on the effect of public policies on day care center quality. Previous papers have studied whether policies directly targeting improved day care center quality have been successful (Hotz and Xiao 2011; Blau 2007; Chipty and Witte 1999; Chipty 1995). I contribute to this literature by studying the unintended consequences on child care quality of a public policy that does not directly target this outcome.⁸

Furthermore, I add nuance to the discussion of child care pricing. Prior research estimates how child care prices respond to the price of inputs (Helburn and Howes, 1996) and the child care tax credit (Rodgers, 2018), along with estimating the price elasticity of demand (Blau and Robins, 1988) and the the elasticity of demand for quality (Blau and Mocan, 2002) and how these factors have affected the evolution of prices over time (Herbst, 2018). This literature models the price of child care as one singular price. I advance this literature by considering differences in prices and profit margins across age groups. These differences are important for thinking about the impact of a public policy that provides a child care substitute for just one age group. I provide empirical evidence showing that day care centers use infants as loss leaders and cross-subsidize care for younger children with tuition from older children. I also provide a model to rationalize this pricing behavior. I believe I am the first to document this fact within the economics literature.

More broadly, my work is connected to research on the effects, often unintended, of public programs on non-targeted individuals. Public interventions often target a subset of the population served by private markets and can therefore have spillovers on the rest of the market. For example, the government provides vouchers and housing targeted toward the subset of the market comprised of low-income households. But housing vouchers increase rents for non-voucher recipients (Susin, 2002) and government-built low-income housing affects nearby housing prices (Diamond and McQuade, 2016). Medicare provides health insurance for the elderly but affects private-pay prices (Clemens and Gottlieb, 2017) and spending (Finkelstein and McKnight, 2005) for the non-elderly. I contribute to this literature by evaluating the

⁷Doromal *et al.* (2018) do use *cross-sectional* inspections data in North Carolina to explore differences in safety across provider types.

⁸Although UPK was intended to provide quality early childhood education in public pre-K classrooms, the rest of the private child care market was not targeted.

impact of public pre-K, a substitute for child care for 4-year-olds, on other ages served in the child care market.

The remainder of the paper is structured as follows. Section 2 provides additional institutional details about New York City’s UPK expansion and on the landscape of licensed child care facilities in the city. Section 3 introduces a conceptual framework for understanding UPK’s effect on child care facilities that serve both younger and older children. The data are described in Section 4. The empirical strategy is further described in Section 5, and results are presented in Section 6. Section 7 concludes.

2 Institutional Background

2.1 Background on Universal Pre-K in New York City

Universal Pre-K (UPK) launched in New York City in 2014 and provides the option of free public pre-K to all 4-year-old residents of NYC for 6 hours and 20 minutes per day and 180 days per year. From 1998 to 2014, NYC ran a much smaller public pre-K program with only 19,000 children enrolled in full-day pre-K in the 2013-14 school year. In the 2014-15 school year, the first year of the expansion, nearly 52,000 students were enrolled, representing 60% of the future first grade public school enrollment for this cohort.⁹ The program further expanded to enroll 67,000 students in 2015-16, the first year the city was able to guarantee every 4-year-old a spot, and its enrollment and capacity has held steady since then.

NYC’s pre-K program is “universal” in the sense that all children are eligible if they are NYC residents and have turned or will be turning 4 in that calendar year. Universal programs are in contrast to “targeted” pre-K programs that are targeted to more vulnerable children, and only children who meet certain income limits, for example, are eligible. Parents who want their child to participate fill out an application ranking their preferred schools. There are no restrictions on where they can apply, though preference is given for students with siblings at the school, and for public schools, children who are zoned for that school. Next priority is given to children who live in the same school district as the site, followed by students in the

⁹Part of the increased full-day enrollment was driven from a shift from half-day to full-day programs. A total of 36,000 children were enrolled in half-day programs in 2013-14 compared to 14,000 in 2014-15. Some is also due to a conversion of 12,700 seats for low-income children, which were previously designed as half-day pre-K coupled with Head Start or child care funding to round out the full day, to full-day pre-K options.

same borough.

The Universal Pre-K expansion came together rapidly and was largely unexpected. Bill de Blasio came from behind to win NYC's mayoral primary in September 2013 on a platform that included the public pre-K expansion. In July of 2013, Bill de Blasio sat in fourth place in the polls at 15.4%, eight percentage points behind the leader, City Council Speaker Christine Quinn. Just two months later, he had surged ahead to first place following a much-lauded campaign ad featuring his son. One of de Blasio's signature proposals was a grand plan to implement free Universal Pre-K for 4-year-olds in NYC with a 0.5 percentage point tax increase on earners making over \$500,000. Any tax increase would need to be approved by the state, and de Blasio's closest competitors, Quinn and Bill Thompson, criticized his plan as infeasible because Governor Cuomo would never approve a tax increase while facing reelection in 2014. But their critiques didn't stop de Blasio from emerging victorious in the September primary, and he went on to win the general election in November in a landslide.

His next challenge was to gain approval for his tax plan from Albany and, in parallel, to prepare logistically to implement UPK in case the funding was approved. De Blasio wanted a secure, earmarked funding stream and insisted that the tax on the wealthy was the only way to accomplish it. Cuomo was equally opposed to raising taxes. A budget deal was reached in late March of 2014 for New York state to contribute \$300 million from the state budget toward the pre-K expansion in NYC, bringing the total spending on UPK in NYC that year to \$575 million.

In addition to funding, de Blasio and his team had to find space for the new UPK classrooms by soliciting proposals from two types of sites, public schools and community-based organizations (CBOs). They began soliciting proposals as early as mid-December 2013 with initial submission deadlines in February 2014, with additional solicitation rounds later in the year. Public school sites were almost exclusively housed in existing elementary schools, and public school principals were invited to submit proposals for new classrooms.¹⁰ Additionally, private schools, day care centers, Head Start programs, and other providers of child care and education in NYC were eligible to apply to host Universal Pre-K classrooms.

The Department of Education had to vet over 3,000 new classrooms for this effort and

¹⁰Although not part of this paper, in 2015, the district opened new "Pre-K Centers", public schools that exclusively housed pre-K classrooms, many of which opened in buildings already owned by the Department of Education, such as closed public schools.

work with the Department of Health and Mental Hygiene to ensure any new sites had the proper license. The first round of new sites was announced in April of 2014, with an additional round of sites announced in June of 2014 and some additional seats continuing to be announced throughout the summer and even into the beginning of the school year.

In order to qualify to host one or more UPK classrooms, CBOs must meet the same minimum criteria as public schools. The only exception is that CBO teachers are not required to have a New York State Birth-Grade 2 certification, though they are required to have a written plan to obtain the certification within five years. Public school UPK teachers are paid higher salaries than typical child care workers. In order to create more equity between public-school-based UPK and CBO-based UPK teachers, de Blasio set aside \$10 million to increase UPK teacher salaries in CBOs.

With the funding and infrastructure in place, NYC opened applications for pre-K students. The first round of applications for students was due in late April with results announced in June, and applications for available seats, including those at new sites, continued to be open until they were filled. For the first year, parents had to use one application for public school sites and apply separately to each CBO site they were interested in. Admission for public school sites was done centrally, with priority given to students with siblings at the school and to students zoned for that school. Most CBO sites admitted students on a first come, first served basis. As a result of the efforts of de Blasio and his staff, there were 25,000 more full-day pre-K seats in 2014 than in 2013.

2.2 Background on the Child Care Market

This paper focuses on the effect of UPK on licensed child care providers who care for children younger than pre-K age. There are two types of licensed child care providers: day care centers and home day cares, which I collectively refer to as child care facilities. Home day cares care for children in residential spaces, while day care centers care for them in non-residential spaces. Children may also be cared for by individuals without a license, such as friends and family or nannies, and I will not be able to speak to this informal segment of the market with this paper.

Home day cares in New York City are licensed, regulated, and inspected by the New York State Office of Children and Family Services (OCFS). Home day cares are required

to obtain a license if they regularly serve three or more children younger than 13 who are unrelated to the provider.¹¹ Home day cares without additional employees, which I will refer to as single home day cares, can care for up to six children, of whom two can be under the age of 2. Home day cares with additional employees, which I will call group home day cares, can care for up to twelve children, with a maximum of two under age 2 for each care provider present (for example, an owner with one additional employee can have four children under 2).¹²¹³¹⁴ OCFS also imposes programming and facility requirements and inspects home day cares for compliance.¹⁵

Day care centers in New York City are licensed, regulated, and inspected by the New York City Department of Health and Mental Hygiene (DOHMH). Day care centers must have a license if they regularly provide care to three or more children under age 6 in a non-residential space.¹⁶ Centers typically provide care in classrooms separated by age with minimum staff-to-child ratios and maximum group sizes determined by the DOHMH. For infants under 12 months of age, there can be at most four children per teacher and eight children per room. Young toddlers ages 12 to 24 months can be cared for in groups of at most ten and must have one teacher for every five children. For 4-year-olds, the maximum is twelve children per teacher and twenty children in a room.¹⁷ The ratios and group sizes for all ages can be found in Table A.1. In general, bachelor's degrees are required for lead teachers and education requirements are lower for aides.¹⁸ DOHMH also imposes programming and

¹¹“Regularly” is defined as three or more hours per day on a regular basis.

¹²Home day cares without additional employees can also have up to eight children if at least two are enrolled in Kindergarten or a later grade, and home day cares with additional employees can have up to sixteen children if at least four are school age. Some home day cares are licensed for slightly lower capacities due to square footage requirements.

¹³When they are present, a provider's own children under age 13 are included when determining maximum capacity unless they are enrolled in kindergarten or a higher grade.

¹⁴Because two providers is the minimum required to use the maximum capacity allowed for a home day care (additional employees do not increase total allowed capacity, only the allowed capacity of infants and young toddlers), throughout the paper, I will assume that group home day cares can care for up to four children under age 2.

¹⁵There are no minimum education requirements for providers, though they must pass a background check and meet continuing education requirements.

¹⁶More specifically, they are required to have a license if care is provided for five or more hours per week for more than thirty days in a twelve-month period.

¹⁷Note that the maximum for UPK classrooms is slightly lower than for child care centers. UPK sites can have at most 18 children if there are two teachers or 20 children if there are three teachers.

¹⁸All programs are required to have an education director with a bachelor's degree and experience working with children. Each classroom must have a lead teacher who must have a bachelor's degree,

facility requirements.

Day care centers undergo unannounced inspections annually. Results of those inspections are used in supplementary analysis in this paper as suggestive evidence on day care safety and quality. A full inspection consists of two separate visits, a facility/environmental initial inspection and a programmatic initial inspection. These can occur anywhere from one day to several months apart. Inspections are also conducted to investigate complaints against the day care or to reinspect to ensure compliance with violations discovered in the initial inspection. Any violations that require reinspection can result in a fine of \$200 to \$2,000. Facilities can also be temporarily shut down until violations are corrected, and repeated or very serious violations can lead to the revocation of their license.

3 Conceptual Framework

Universal Pre-K extends free public schooling and school-day child care to a single age cohort of children. On the private market, centers that care for these 4-year-olds often also care for younger children as well. Therefore, a change in the market for care for pre-kindergarten-aged children could affect care for younger children. Whether this effect is positive, negative, or zero is an empirical question. There are different possible consequences of Universal Pre-K that could affect the supply or demand for private child care in different directions. And if there is little crowd out of 4-year-old care, then there may be little effect on the child care market. The following subsections discuss some reasons why the effect might be positive or negative. When I say it could have a positive effect, I am referring to increasing capacity, reducing prices, and/or increasing quality and vice versa for a negative effect.

3.1 UPK could have a negative impact on child care for younger children

In this section, I will discuss three reasons UPK could have a negative impact on child care for younger children. I describe each briefly in this paragraph and elaborate on these mechanisms in the remainder of the section. First, the viability of child care facilities could be threatened by an inability to attract enough 4-year-olds. For day care centers, 4-year-olds

except for infant/toddler classrooms where the lead teacher can have a high school diploma plus experience or an approved education plan. Assistant teachers must have a high school diploma.

are the least costly to care for, and day cares use them to cross-subsidize younger ages. When these children are crowded out by UPK, day care centers are losing their most profitable children. For home day cares, legally only about one-third of total capacity is available for children under 2. The remaining two-thirds is only available for children ages 2 and older. If home day cares are not able to find enough older children to fill those slots, they may close. Second, UPK increases the demand for early childhood educators, which may make it more difficult for day care centers to attract and retain teachers. Finally, there may be restrictions in the supply of buildings suitable for early childhood education. If more of these classrooms are used for 4-year-olds, there may be less space for younger children.

I will begin by establishing the empirical fact that day cares cross-subsidize younger children with older children. Day care tuition data is not available for New York City, so I will instead look at prices for another large East Coast city, Philadelphia, Pennsylvania.

In Pennsylvania, one teacher can care for at most four infants or at most ten 4-year-olds. With labor accounting for approximately seventy percent of operating costs ([Helburn and Howes, 1996](#)), we would expect tuition to be about twice as high for infants as it is for 4-year-olds.¹⁹

Figure [A.1](#) plots the actual cumulative density functions of the ratio of tuition for younger children of a given age to 4-year-olds in Philadelphia and its surrounding counties as of November 2017.²⁰ The red line represents the expected tuition ratio given the required staff ratios and assuming labor costs are seventy percent of total costs. In subfigure (a), we can see that less than one percent of centers charge at least the expected premium for infant care, and the average center charges only a 22 percent premium. Subfigures (b) and (c) show that ages 1 and 2 are cross-subsidized by 4-year-olds as well. For both, less than one percent of day cares charge at least this expected tuition ratio. For one-year-olds, the average tuition premium is fifteen percent compared to the expected premium of seventy percent, and for

¹⁹Each teacher can care for 2.5 times as many 4-year-olds as infants, so assuming similar wages for teachers in the two classrooms, labor costs are approximately 2.5 times higher for infants. If labor accounts for seventy percent of total costs, then accounting just for increased labor costs, we would expect infant tuition to be $1.5 \times 0.7 = 2.05$ times higher than 4-year-old tuition.

²⁰This data was acquired from the Pennsylvania Department of Human Services through a Right to Know Law request and includes Bucks, Chester, Delaware, Montgomery, and Philadelphia counties. The pricing information includes the prices as the day cares reported it for their licenses. Day cares are supposed to report the tuition that they charge absent any discounts. Similar data is not available for New York as facilities are not required to report their prices to the licensing agencies in New York.

age 2, the mean is 8 percent versus the expected 45 percent, including twenty percent of day cares that charge the same amount for age 2 and age 4. Thus, younger children are substantially cross-subsidized by older children.

The cross-subsidization found in the microdata in Philadelphia is seen in aggregate national data as well. States generally require two to three times as many staff members to care for a group of infants as for a group of 4-year-olds of the same size (Vandell and Wolfe, 2000). Using the approximation that labor costs are about 70 percent of day care facility costs, we would expect infant care to be 70 percent to 140 percent more expensive than care for 4-year-olds. However, according to a national survey, they are on average only 12 percent higher (Schulte and Durana, 2016). Thus, if 4-year-olds are crowded out, day cares are losing their most profitable children, potentially affecting their viability.

The question of why day cares price in this way is addressed with a model in Appendix B. I provide a summary of the intuition here. Demand for infant care and pre-K care is interrelated. Because infant care is more costly to provide, day cares cross-subsidize infants with pre-K students in order to attract more pre-K students. One possible mechanism is switching costs. Parents have to pay a switching cost in order to change child care arrangements, so day care choices are sticky. Day cares charge less profitable infant prices in order to attract infants who will become profitable pre-K children. Another mechanism is sibling effects. Parents may value sending their children to the same facility, so if the day care can attract the infant, they may be able to attract their older sibling as well.

Figure B.1 shows what happens to the private day care market when Universal Pre-K is introduced within a simplified supply and demand framework where the demands for pre-K care and infant care are interrelated. When free Universal Pre-K is introduced, parents decrease their demand for private pre-K care, which reduces the equilibrium price and quantity of private pre-K care. The cross-subsidization of infant care with pre-K care means that the supply curve for infant care depends not only on the price of infant care but also on the equilibrium price and quantity of pre-K care. When the price and quantity of private pre-K care fall, day cares reduce the quantity of infant care they are willing to supply for a given price, shifting their supply curve. This shift reduces the equilibrium quantity of infant care and increases its price.

In this simplified model, the size of the decrease in the equilibrium quantity of infant care depends on the elasticity of demand for infant care. If parents are very inelastic in their

demand for child care centers for infants, then most of the adjustment will come in the form of increased prices, and there will be little adjustment on quantity. On the other hand, if parents are very elastic, most of the effect will be in the form of a decrease in quantity. Demand by parents who can only pay for child care with government-provided vouchers may be almost perfectly elastic so that day cares that take a large number of subsidies may be very responsive on the quantity margin since they cannot respond by increasing prices.

The above model considers market-driven changes to the price and quantity of child care, but child care facilities could also make adjustments to the cost margin by changing their quality. When faced with decreased pre-K enrollment, centers could also respond by lowering their quality in order to lower their costs.

A similar argument holds for home day cares, which are restricted in the number of infants and toddlers they can legally care for. In New York City, a maximum of about one-third of total capacity can be filled by children under age 2.²¹ If a home day care cannot find enough children to fill the “age 2 and up” slots, they may shut down, increase prices, or reduce quality.

Another reason UPK might reduce the equilibrium quantity of care available for younger children is because of constraints in the supply of suitable classrooms for early childhood education. In New York City, sixty percent of UPK enrollment is in community-based organizations, many of which are day care centers. If these centers are using classrooms for 4-year-olds that otherwise would have been used for younger children, it would reduce the quantity of center-based care supplied for younger children. Note that due to facility regulations, it is easier to substitute a 4-year-old classroom for a 3-year-old classroom than for a classroom for even younger children.²² A substitution of 4-year-old classrooms for 3-year-old classrooms is something I will not be able to observe in my data because I have capacity by age separated only into two groups, under age 2 and ages 2 to 5. But any effect from conversions of 4-year-old classrooms to classrooms for children under age 2 would be included in my estimate.

²¹Home day cares in New York City can have at most two children under two for each teacher who is present. The total maximum capacity across all ages is six for a home day care with no additional employees and twelve if they do have additional employees, so slots available to younger children make up approximately one-third of total capacity.

²²For example, NYC regulations require that there be 30 square feet of space in a classroom for each child. A classroom of infants can have at most 8 children, while a classroom of 4-year-olds can have at most 20 children, so infant classrooms tend to be smaller than pre-K classrooms.

Finally, with the abrupt increase in demand for early childhood education teachers, day cares may find it difficult to hire enough staff members to remain open, particularly because salaries in the public school pre-Ks, and to a lesser extent the UPKs in the CBOs, are substantially higher than that of the average day care teacher. Day cares may also find it difficult to retain high-quality teachers. These general equilibrium effects are unlikely to be a large part of the estimated effects in this paper because my estimation strategy estimates the effect of placing one new UPK site in a given area holding fixed the general equilibrium effects of the expansion as a whole. There could certainly be small labor market effects from the addition of one new UPK (i.e. a day care worker leaves to go work at the UPK that opened next door), but overall, labor market effects are likely to be a small part of what I am estimating. It is useful to separate the general equilibrium labor market effects from the partial equilibrium effects because the policy prescriptions are different. A decrease in day care center supply due to labor market effects can be ameliorated by increasing the supply of early childhood educators, but that solution will not be as relevant if the cause of the decline is cross-subsidization across age groups or inelasticity of building supply.

3.2 UPK could have a positive effect on child care for younger children

This section discusses a few reasons why there might be a positive effect of UPK on child care for younger children. First, Universal Pre-K may increase demand for sibling care and/or for before- and after-school care. Parents who would otherwise have stayed home with their 4-year-old and younger child due to high day care expenses for two children may decide to enter the labor force when their 4-year-old becomes eligible for UPK so that they would only need to pay for one child in day care. Similarly, parents who enter the labor force because of UPK may also require before- and after-school care for their child since public pre-K does not last for the entire work day. In both of these cases, the increase in demand is likely to be very geographically close to new UPK sites since parents will want to be able to drop off siblings at nearby locations and before- and after-school care would need to be offered nearby for ease of transportation. Home day cares in particular may be able to provide before- and after-school care because they have more flexibility to be open for only a few hours before and after school than day care centers do.

Another reason UPK could have a positive effect is the flipside of the building constraints story from the previous section. If there is an inelastic supply of buildings suitable for early childhood education and UPK draws students from day care centers to classrooms in places that would not have hosted day care center classrooms otherwise, such as public schools or libraries, it could increase the supply of classrooms available for younger children. Note that it is more difficult to retrofit a 4-year-old classroom to be used for infants than it is to switch it to a 3-year-old classroom, so again converting 4-year-olds classes to be used by younger children may be a less relevant margin when studying care for children under age 2.²³

Day care centers may also benefit from hosting UPK classrooms. Depending on how well they are compensated by the city, the possibility of hosting a UPK classroom could encourage entry of new day cares or prevent closures of existing ones that also serve younger children.

Finally, there could be a “crowd in” effect where the city’s emphasis on quality early childhood education increases demand for quality early childhood education at younger ages, increasing demand for center-based care (or licensed home day care) for younger children. Similarly, through an income effect, UPK could also increase demand for child care at younger ages by decreasing the lifetime costs of sending a child to day care.

4 Data

4.1 Child Care Facility Data

The primary analysis focuses on the number and capacity of child care facilities in New York City using licensure data obtained through Freedom of Information Law requests to New York City’s Department of Health and Mental Hygiene (DOHMH) and New York State’s Office of Children and Family Services (OCFS). The DOHMH data covers years 2008 to 2018 and contains licensure information for day care centers, while the OCFS data covers years 2010 to 2017 and contains licensure information for home day cares. Both data sets contain the name, address, opening date, and capacity of the facilities.²⁴ The OCFS data also

²³Infant and toddler classrooms must have sprinkler systems, which is not required in classrooms for older children. They also must be on the ground floor of buildings to ease evacuation, which at that age is usually done by putting infants in rolling cribs and rolling them outside. Classrooms for older children can be on the second or third floors.

²⁴Per a late 2009 law, home day care providers can opt to remove their street address from public availability. For these observations only zip code is available. Less than 5% of the sample does not have a full address, and they were dropped from the analysis.

contains closing date for home day cares that are no longer in operation. The DOHMH data contains the most recent permit expiration date for the two-year license, and an indicator for whether the facility was closed as of January 2018, when the data was extracted. It also includes an indicator for whether the license is for younger children (ages 0 to 2) or older children (ages 2 to 5). For both data sets, I geocoded the addresses in order to determine the spatial locations of the facilities.

Additionally, data on the results of day care center inspections from the NYC DOHMH are used in complementary analysis as a suggestive indicator of day care safety and quality. These data cover all day care center inspections from 2008 to 2017 with indicators for the date of the visit, the type of visit (routine annual inspection, response to a complaint, or to follow up on violations), and a record of whether any violations were discovered, and if so, what they were.

There are three types of inspection violations: public health hazards, critical, and general. Public health hazards are the most serious. They present an imminent threat to the health and safety of children and must be corrected within one business day. Critical violations are serious violations that do not present imminent threats and must be corrected within two weeks. All other violations are considered general violations and must be corrected within one month. Facilities are reinspected for compliance if they have one or more public health hazard or critical violations or six or more general violations.

The day care center inspections data is also used to estimate a closing date for out-of-business facilities to within approximately one year instead of the two years allowed by the licensure data. Licenses are issued for two years, while inspections are required every year. In the main specification, I consider a provider to be open for their full license term if they had an inspection at any point in the second year of their license. If they did not, then I set their closure date to one year before the expiration date. Results are robust to alternative definitions.

As of January 1, 2014, the last year before the UPK expansion, there were 2,228 licensed day care centers in New York City with capacity for 133,000 children and 7,887 licensed home day cares with capacity for 77,842 children (see Table 2). Of the day care centers, 17 percent were licensed to serve younger children while the remaining were licensed for older children. Facilities licensed for younger children almost always also have a license to care for older children either in the same building or as part of a larger chain.

Figure 2 depicts the trends in number and capacity of day care centers from 2008 to 2018, the period over which I have data.²⁵ The number and capacity of day care centers serving younger children has been steadily increasing over time with a slight slowdown from 2014 to 2017, followed by a jump in 2018, the most recent year. Day care centers serving older children have also increased in size over the sample period with a leveling off in city-wide capacity and slight decline in number of facilities beginning in 2016.

Capacity of home day cares in New York City peaked in 2014 and has dropped by about ten percent since then, from about 78,000 in 2014 to about 70,000 in 2017. There is a steeper drop in number of home day cares than capacity of home day cares since 2014 due to an increase in the fraction of home day cares that are group home day cares, which have at least one additional employee and are therefore licensed for a higher capacity, versus single home day cares, which do not have additional employees.

The raw data point to a slowdown in the growth of the child care market in New York City beginning just after 2014, when UPK was implemented. These, of course, are only correlations. Causality will have to wait until Section 6, where I discuss results.

4.2 Universal Pre-K Data

The locations of new Universal Pre-K sites are used to determine whether and how intensely a given area is “treated” by Universal Pre-K. The locations of the sites are taken primarily from the UPK viewbooks provided to parents to inform them about their options for UPK. I use the viewbooks from 2013-14, the year before the expansion, and 2014-15, the first year of the expansion, both of which I found through the Wayback Machine.²⁶

Viewbooks provide the name, address, and number of available full-day and half-day seats, except for the 2013-14 community-based organizations, for which number of seats in 2013-14 is not available (but an indicator for whether there is full-day UPK does exist). I extracted this data from the PDFs and geocoded the addresses to determine the spatial locations of the UPKs. Part of the UPK expansion included increasing capacity in existing

²⁵Although I have day care center data back to 2008, I only have home day care data going back to 2010. Therefore, most analysis begins with 2010 both for consistency and to restrict to a reasonably short pre-period.

²⁶There were multiple versions of the 2014-15 viewbook due to UPK sites being added. The viewbook I use is the latest I could find and is from June 27, 2014, just over two months before the first day of school.

sites, but because I do not have capacity data for the 2013-14 CBOs, treatment will be based on an indicator for whether a site is offering full-day UPK for the first time in 2014-15.²⁷ Exposure is therefore based on new *sites* not new *capacity*.

The number of sites and total capacity of those sites are provided in Table 1. In 2013-14, there were 514 CBO full-day pre-K sites and 438 public sites, for a total of 952.²⁸ In the expansion, the number of CBO sites in my data increased by 363 and CBO capacity increased by 19,397.²⁹ The number of public school sites increased by 69 and the capacity by 5,590.³⁰ The total combined capacity of the public and CBO sites in the viewbook data is 57,151 in the year of the expansion, which is consistent with enrollment figures of about 53,000 as some seats were not filled.

5 Empirical Strategy

To identify how a new UPK location affects the local day care market, I compare availability of care in areas very close to new UPK sites to areas not as close to new UPK sites before and after the 2014 UPK expansion in NYC. Exposure is measured intensively, so the comparison is between areas with more new UPKs and those with fewer new UPKs. The assumption is that areas that are closer new UPKs will be more affected by the expansion than those that are farther away. Evidence that there is a utility cost of traveling farther to school, particularly at younger ages, supports the claim that the relevant market for child care will be fairly local

²⁷So if a UPK offered some full-day seats in 2013-14 and added more full-day seats in 2014-15, that would not be counted as an expansion since I do not have the data to see that there were fewer seats in 2013. However, if they only offered half-day UPK in 2013-14 and switched to full-day UPK in 2014-15, then that would count as a new site and is treated the same way as a site that did not have any UPK in 2013 that added full-day UPK in 2014.

²⁸Of these CBO sites, only 59 did not have any income requirements in any of its classrooms. Classrooms with income requirements are primarily contracted with the Administration for Children's Services (ACS) and provide half-day pre-K coupled with child care and/or Head Start to make it a full-day option. The Department of Education does not include these classrooms in its count of full-day pre-K seats in the pre-period since they are not exclusively full-day pre-K. However, I do include them since when they were converted to UPK seats in the following year, it was primarily a change in name coupled with a programmatic change, and it provides essentially the same substitute for private child care as it did before.

²⁹The total capacity number is from Shroff *et al.* (2014) since site-level CBO capacity is not available.

³⁰About forty percent of the capacity expansion in public schools is at current sites. Although expansions are not counted as new sites in my analysis, results are robust to including these public school expansions as new sites.

to the new site (Dinerstein and Smith, 2014). The identifying assumption is that changes in the day care market in areas with fewer new UPKs provide a good counterfactual for areas with more new UPKs. This assumption is discussed in more detail in Section 5.2.

The main outcome of interest is the total number of seats in day care center seats for younger children. Note that if I were only interested in closings, day care centers that were open when UPK rolled out would be a natural unit of observation and treatment could be based on some measure of distance to new sites. However, the most important outcome is not the change in the closure rate for currently-open centers but how UPK affected the total stock of day care centers, which incorporates changes in both closings and openings. Stock must be defined over a geographic area, so it is necessary to define these areas.

In addition to defining areas over which to measure the stock of child care facilities, it is necessary to define child care markets. These may be the same or different from the areas used to measure stock. The ideal experiment would be to have a set of markets with uncrossable boundaries where everyone living within a given area worked and attended school and day care within that boundary. Then, if we randomly dropped UPKs in some of them, we could observe what happens to the market. However, in the real world, people do not stick to one clearly-delineated area with uncrossable boundaries, particularly in large and densely-populated areas like NYC, so we must find a way to approximate markets, preferably one that allows markets to overlap.

Defining markets is challenging, especially with little information. The industrial organization literature solves this problem by using structural estimation. The added value of this approach is usually that it can generate counterfactuals, but in this case, where I observe the market before and after a policy change, a simpler reduced-form approach should be sufficient and is more straightforward.

There are two primary reduced-form approaches to this problem: the “ring method” and using administrative boundaries. Neither of these is ideal in this case, and they are discussed in more detail in Appendix C.

Briefly, the “ring method” is not ideal due to the spatial concentration of treatment. This method would involve drawing a ring around each new UPK site to delineate the market affected by that UPK and comparing changes in the stock of child care centers in that ring to changes in the stock of child care centers in a larger “unaffected” ring. The ring method is not ideal here because even if the child care market area affected by a given UPK (and

therefore the relevant ring) is very small, the markets (and rings) would still overlap, and “treated” areas will overlap with areas that were supposed to be “untreated,” biasing the estimates. And the direction of the bias is not clear ex-ante, as explained in Appendix C. Furthermore, areas with more than one UPK would be overweighted in the data.

Using administrative boundaries and measuring changes in the child care market within those boundaries based on their level of exposure to new sites would solve the problem of overlapping areas of observation. However, it creates a new problem because these boundaries are arbitrary and irregularly-shaped, and there is no natural way to test robustness to them. Instead, I divide New York City into a grid of tessellated hexagons. All hexagons are the same size, and because the researcher creates the grid, it is straightforward to test robustness to shifting the grid and changing the size of the hexagons.

5.1 Tessellated Hexagons improve upon alternative methods

Using a grid of tessellated hexagons allows me to address the concerns of using areas of observation that overlap and using arbitrary boundaries. In order to account for potentially-overlapping markets, I allow the child care market in one hexagon to be influenced by the change in the number of UPK sites in that hexagon and the hexagon’s neighbors. Although this is still an approximation of the true markets, it does improve upon existing methods by better handling dense treatment and allowing for robustness checks to the market boundaries. Dividing space into a grid allows exposure to be measured as the number of nearby UPKs and accurately identifies areas with and without new sites. The ring method, on the other hand, does not provide unbiased estimates when treatment is spatially dense. The definition of the grid is also very flexible, which allows me to easily test for robustness to different hexagon grids.

I choose hexagons for the shape of the grid areas for two primary reasons. First, neighboring hexagons are easy to define and all have the same relationship to the central hexagon. There are three options for a grid of regular polygons: triangles, squares, and hexagons. Squares have four neighbors with whom they share a side, and they also have four additional neighbors with whom they share a vertice but not a side. Equilateral triangles similarly have same-side neighbors and vertice-sharing neighbors. For hexagons, on the other hand, all

neighbors are same-side neighbors.³¹ Put another way, the centroids of all of a hexagon’s six neighbors are the same distance from the middle hexagon’s centroid, whereas for a square, the four diagonal neighbors are farther away than the four same-side neighbors. This property of hexagons makes including in the effect of UPKs in neighboring hexagons more natural, and no adjustment is needed to account for different types of neighbors. The inclusion of neighbor effects is necessary to account for overlapping markets. Not accounting for neighbor effects assumes that a UPK in a given hexagon only influences the child care market in that hexagon and not neighboring hexagons.

A second reason for choosing hexagons is that of the three options, hexagons have the lowest area to perimeter ratio, so the average point in a hexagon is closer to its centroid than for a same-area triangle or square, reducing boundary effects. In other words, hexagons are preferred because they are the closest to circles.³² Therefore, I will use a grid of tessellated hexagons.³³³⁴

5.2 Estimation and Identifying Assumptions: Day Care Center Capacity

Now that the geographic boundaries have been chosen, I can implement a difference-in-differences design that leverages the natural experiment of the child care market in some hexagons being more exposed to the 2014 UPK expansion than others.³⁵ The idea is to

³¹See Figure A.2 for a visual.

³²In fact, hexagons are the least-perimeter way to divide a plane into regions of equal area (Hales, 2001).

³³Squares do have the advantage that they are easier to work with using longitude and latitude. See Harari and La Ferrara (Forthcoming) for an application using a grid of squares.

³⁴There are other possible approaches to this problem, primarily from the literature on structural industrial organization. Due to the setting and data here, a structural approach is not necessary. Although it is certainly possible to approach the question in this way, it is not necessary in this case since I observe the outcomes of interest both before and after the policy change. Additionally, a structural approach would require some additional assumptions due to the lack of individual-level data. Another approach is continuous difference-in-difference, as pioneered by Diamond and McQuade (2016). In the application in their paper, there are few enough overlapping markets that they are able to run their estimation assuming zero overlap. In order to apply to this case, the method would have to be adjusted for overlapping markets. Furthermore, the beauty of the hexagon approach vis-a-vis these alternative approaches is in its simplicity, transparency, and minimal required computing power while still providing plausible identification in this setting.

³⁵Because children are more likely to attend school at a location that is closer to them than one that is farther away (Dinerstein and Smith, 2014), differences in the locations of new sites leads

compare changes to the day care market before and after the expansion in hexagons that got more sites to ones that got fewer (or none). Under a set of identifying assumptions discussed below, differential changes in day care availability in areas with more new UPK sites can be attributed to differential exposure to UPK, and the change is the causal effect of a new UPK site.

Level of exposure to the expansion is measured by the number of new full-day UPK sites on net in that hexagon and the hexagon's neighbors. Exposure is defined as the number of full-day UPK sites in 2014 minus the number of full-day UPKs in 2013 and is allowed to be negative.³⁶ I could instead use the number of 2014 sites that did not exist in 2013. I have chosen this definition instead primarily because if a UPK site from 2013 closes but a new one opens next door, the hexagon's exposure to UPK does not change even though there is a new UPK location.³⁷

Because the day care market is likely to take time to adjust to the new sites, I focus on a semi-dynamic specification that allows us to observe how the effect of UPK changes with time since exposure. Day care facilities that will close as a result of UPK probably take at least a year or two to make that decision. There also may be child care facilities that would have opened in the absence of UPK. The decrease in openings may have a compounding effect on stock that takes time to evolve. There are some that would have opened in the first year after UPK, some that would have opened in the second year after UPK, etc.³⁸ If there are dynamic effects, an estimate of the average effect during the post-period would understate the impact of UPK.

Specifically, I estimate the following semi-dynamic equation:

the day care market in some hexagons to be more affected by the expansion than others.

³⁶A better measure of exposure may be the change in the number of full-day pre-K *seats* instead of the change in the number of full-day pre-K *sites*. However, micro-level data on the capacity of community-based organization sites is not available for 2013.

³⁷I also do not have identifiers to easily link 2013 sites to 2014 sites. Linking them without identifiers would introduce noise in the measure.

³⁸A positive effect could also take time to evolve. There are facilities that would eventually close in the absence of UPK who remain open. In this case, the effect on stock would theoretically be seen in the year they would have closed. Similarly for openings, it can take time for new firms to enter the market.

$$Y_{ht} = \beta_0 + \sum_{j=2015}^{2018} [\beta_j * \mathbb{1}_{X_t=j} * New_h + \beta_{nj} * \mathbb{1}_{X_t=j} * NewNB_h] + X_t + \sigma_h + X_t * h + \epsilon_{ht} \quad (1)$$

Observations are at the hexagon-year level. In the main regressions, Y_{ht} is the total capacity across day care centers for younger children ages 0 to 2 years in hexagon h as of $1/1/t$. I also consider the total number of day care centers serving younger children, the capacity of home day cares, and the number of home day cares. $\mathbb{1}_{X_t=k}$ is an indicator for year k , New_h is the net number of new full-day UPK sites in the hexagon in 2014, and $NewNB_h$ is the net number of new sites in hexagon h 's immediate neighboring hexagons. The regression includes year and hexagon fixed effects. I include hexagon-specific linear time trends to account for potential heterogeneity in pre-expansion trends. The data spans years 2010 to 2018. Because the outcome of interest is the stock as of January 1st of each year, the observation from year 2014 is the last one before the expansion. Standard errors are clustered at the hexagon level. The coefficients of interest are the β_j s and β_{nj} s. These estimate the effect of one new Universal Pre-K site in the same hexagon (β_j) or a neighboring hexagon (β_{nj}) on the number of seats for younger children in day care centers in that hexagon in year j .

Identification requires that day care center capacity would have followed a similar trend (net of the hexagon-specific linear time trend) in all hexagons in the absence of the UPK expansion. This assumption would be violated if they are not trending similarly in the pre-period, which I will check for empirically. It can also be violated if there is another shock to the day care market that occurs in 2014 and is correlated with hexagon exposure status. The presence of such a shock is not something I can definitively test for empirically and is an assumption that will be required for causality.

To check for pre-trends, I estimate the fully-dynamic equation:

$$Y_{ht} = \beta_0 + \sum_{k=2011}^{2013} [\beta_k * \mathbb{1}_{t=k} * New_h + \beta_{nk} * \mathbb{1}_{t=k} * NewNB_h] + \sum_{j=2015}^{2018} [\beta_j * \mathbb{1}_{t=j} * New_h + \beta_{nj} * \mathbb{1}_{t=j} * NewNB_h] + X_t + \sigma_h + X_t * h + \epsilon_{ht} \quad (2)$$

This is the same regression as equation (1) except that it also includes interactions of the treatment variable with year indicators for the pre-period.³⁹

5.3 Descriptive Statistics at the Hexagon Level

In the main specification, the hexagons are fairly small (covering approximately 10 city blocks, or about three blocks wide). In two alternative specifications, I consider hexagons that are medium-sized (covering about 20 city blocks) and large-sized (covering about 40 city blocks). In all three cases, I allow the child care market in a given hexagon to be affected by new UPKs in that hexagon and in the hexagon’s neighbors.

See Figure 1 for maps displaying the number of full-day sites in an example hexagon grid in 2014 compared to 2013. In this medium-sized hexagon grid with 998 hexagons, 372 hexagons had more full-day UPK sites in 2014 than they did in 2013. Figure A.3 displays the exposure variable, the number of 2014 UPK sites in the hexagon minus the number of 2013 UPK sites, for the primary hexagon grid. In order to eliminate areas such as parks, water, and airports, hexagons are included in the analysis only if they had a family day care home, day care center, or elementary school at some point in the pre-period.⁴⁰ Of the 2,658 hexagons in this main hexagon grid, 449 had more UPK sites in 2014 than 2013 and 82 had fewer.

Random assignment of UPKs to hexagons is not necessary for identification, and there is no reason to expect exposed and unexposed hexagons to be observably similar. However, comparing these hexagons on observables is still informative in order to better understand the characteristics of the treatment and control groups. Table A.3 compares the number of UPK sites and child care facilities in hexagons that gain one or more UPKs (“expansion” hexagons) versus ones that do not (“non-expansion” hexagons). The number of facilities is measured as of 1/1/2014, just before the UPK expansion. There is no statistical difference between the number of 2013-14 full-day UPK sites between expansion and non-expansion hexagons. The average expansion hexagon had 1.13 more full-day UPK sites in 2014 than it

³⁹These additional interactions with year indicators are necessary to check for pre-trends. It is not the main specification because the semi-dynamic specification is more efficient since it uses the whole pre-period to estimate the hexagon-specific linear time trends, while this specification just uses the trend between the left-out years 2010 and 2014.

⁴⁰I have run robustness checks including the entire sample, and regression results are virtually unchanged by this restriction.

did in 2013. They also had slightly more UPK sites in neighboring hexagons: 0.82 vs. 0.68 for the non-expansion hexagons.

Given the speed of the UPK expansion, UPK sites were primarily located in pre-existing child care and school facilities since there was not time to build or significantly renovate new buildings. So as expected, expansion hexagons have more day care centers serving children ages 2 to 5 since a hexagon is more likely to gain a UPK site if it has more potential sites. And because having more day care center licenses for classrooms for children ages 2 years to 5 years is correlated with having classrooms licensed for ages 0 to 2 years, expansion hexagons also have a higher licensed day care center capacity for children under the age of 2. There are approximately three more seats for children ages 0 to 2 in expansion hexagons than non-expansion hexagons, and the difference is statistically significant. Differences across hexagon types in the number of home day cares are small and not statistically significant.

Table [A.2](#) presents a regression table showing what characteristics of a hexagon predict the number of full-day UPK sites it had in the 2014-15 school year. The characteristics are based on Census block data from the 2010 Census that I mapped to hexagons based on which hexagon the block's centroid fell in. The most important predictors for the number of 2014 sites are the number of 2013 full-day UPK sites and proxies for the number of potential sites, the number of public elementary schools and the number of day care centers serving children ages 2 to 5 as of 1/1/2014. Since public elementary schools and day care centers are the primary providers of UPK, it makes sense that they are strong predictors of where new UPKs are located. New UPKs were also significantly more likely to be in areas with higher fraction of non-white residents. Interestingly, the number of day care centers for children under age 2 and the number of home day cares do not have any additional predictive power on whether a given hexagon has a new UPK. Although these near-zero point estimates and statistical insignificance are not necessary for the identification strategy, this regression provides suggestive evidence that the child care markets for infants and toddlers in hexagons that do not get a UPK provide a good counterfactual for the ones that do.

5.4 Estimation and Identifying Assumptions: Inspections

To complement the results on the change in availability of child care, I estimate the effect of nearby UPK sites on day care center inspection results and complaints, potential indicators of

quality. Day cares have three primary margins on which to adjust when faced with crowd out from UPKs: price, quantity, and quality.⁴¹ One reason changes in day care center capacity are important is because day care centers are on average the highest quality form of non-parental care (Bassok *et al.*, 2016a), so a change in day care center capacity could affect children’s quality of care and therefore future human capital development. But if day care centers reduce their quality, then even children who are able to stay in the generally higher-quality day care market may suffer from reduced quality of care.

To estimate the impact of public pre-K on private day care center quality, I use the same identification strategy as before, but now the unit of observation is at the inspection level for the inspection results and the day care-year level for the complaints outcome. In addition, I dropped the data from the implementation year (2014) since it is not clear exactly what point in the year treatment begins, and that point may even be different for different hexagons depending on how early a given site was announced. Dropping the data from the year 2014 leaves 2013 as the last year before the expansion, so that will be the left out year. Finally, here there is no reason to restrict to inspection results for day care centers for younger children, so I pool inspection results and complaints for licenses for both younger children and older children. I do exclude day care centers that I spatially match to new UPK sites since the treatment of having a new UPK classroom in the day care center is different from having a new UPK site at a different organization nearby.⁴²

For consistency across outcomes, I will use the same measures of exposure as in the analysis of day care center capacity changes, which is the net number of new sites in the hexagon. The regression allows day care centers in a given hexagon to be affected by competition from new UPK sites in that hexagon and neighboring hexagons. Specifically, for complaints, I estimate:

$$Y_{iht} = \beta_0 + \sum_{k \neq 2013} [\beta_k * \mathbb{1}_{t=k} * New_h + \beta_{nk} * \mathbb{1}_{t=k} * NewNB_h] + X_t + \sigma_i + \epsilon_{iht} \quad (3)$$

Where Y_{iht} is an indicator for whether day care i located in hexagon h had any complaints

⁴¹For more details on mechanisms and a model, see Appendix A.

⁴²More specifically, I first spatially matched the 2014 pre-K sites to the 2013 sites to determine which sites were new. I then spatially matched these new sites to the day care center data to determine which day care centers hosted full-day pre-K for the first time in 2014.

during year t , New_h is the net number of new UPKs in hexagon h , and $NewNB_h$ is the net number of new UPKs in hexagon h 's neighbors. I also added a day care center fixed effect, σ_i , so that all estimates are *within-center* changes in quality. Standard errors are clustered at the hexagon level.

I also ran this same regression with an outcome variable of the number of inspections for a day care center per year. The addition of a new UPK site could increase the caseload of the inspectors and lead to a decrease in inspections. A decrease in inspections could affect the violations observed at a given inspection. I do observe a slight decrease in the number of inspections for day care centers that are closer to new UPK sites, so I will control for months since the last inspection in my regressions on inspection outcomes to try to control for any mechanical effect of reduced inspections.

To estimate changes in the number of violations and the probability of requiring a reinspection, I ran regressions at the inspection level with day care center and year fixed effects:

$$Y_{jih}t = \beta_0 + \sum_{k \neq 2013} [\beta_k * \mathbb{1}_{t=k} * New_h + \beta_{nk} * \mathbb{1}_{t=k} * NewNB_h] + Months_t + X_t + \sigma_i + \epsilon_{jih}t \quad (4)$$

Where $Y_{jih}t$ is the outcome of interest from inspection j of day care i located in hexagon h during year t and $Months_t$ is the number of months since the last inspection. This regression is the same as equation (3) except that it is run at the inspection level and includes the control for the number of months since the last inspection. Inspections are included only if they were the initial annual inspection and not if they were due to a complaint or were a reinspection. The outcomes considered are the total number of violations found in the inspection and whether the inspection was poor enough to trigger a reinspection. Facilities are reinspected if they have any of either of the two most serious violations (public health hazard or critical) or if they have six or more violations total.

The identification assumptions for the capacity regressions are also required for these regressions. It must be the case that the day care centers would have followed similar trends in the absence of the UPK expansion. I will again use the fully-dynamic specification to confirm that there are no pre-trends in the data.⁴³

⁴³Another potential problem for identification is differential attrition. If lower quality day cares

6 Results

6.1 UPK reduces availability of infant and toddler care

Figure 3 plots the coefficients from the fully-dynamic specification (equation (2)) of the year indicators interacted with level of exposure, as measured by the net change in the number of UPK sites in the hexagon. With all pre-period coefficients not statistically significant from zero, the parallel trends assumption appears to be reasonably satisfied, and we can see a decrease in the capacity of day care centers for infants and young toddlers as a result of own-hexagon UPKs beginning with the observations from 1/1/2015, the first post-expansion observation. As expected, it takes some time for the child care market to adjust as the effect is growing over time. There appears to be no effect of UPKs in neighboring hexagons, implying that the effect of a new UPK site on child care availability is very local.

Coefficients from the semi-dynamic specification (equation (1)) are in Table 3. By 2018, there was a decline of 1.3 day care center seats for children under the age of 2 years for each new UPK site in the hexagon, a decline of 20% on the pre-expansion base of 6.33 (see Table A.3). With 432 net new UPK sites in 2014, I estimate that the 2014 NYC UPK expansion resulted in 562 fewer day care seats for infants.

The 2014 expansion increased the number of sites by 432 and enrollment by 34,000, but the full UPK program as of 2018 includes four times as many UPK sites. If the results here can be extrapolated to whole program, they imply that there are 2,248 fewer infant and young toddler seats as a result of the UPK program. These seats represent 24% of the total capacity of this sector as of 1/1/2014. To put these numbers in context relative to effects on the child care market found in other papers, Hotz and Xiao (2011) finds that the number of child care centers in an average area falls by about 10% in response to reducing by one the number of infants who can be cared for by one staff member or to increasing the number of years of education required to be a center director by three.

Day care centers are generally regarded as providing the highest quality non-parental care (Bassok *et al.*, 2016a), so these children may be crowded out to lower-quality environments. Furthermore, these regressions measure the effect of adding one new UPK site to a given

are closing (or not opening), it would affect the composition of day cares in the sample over time. Day care center fixed effects alleviate some of this concern, but differential attrition on quality could still impact the year effects. In order to address this concern, I verified that results are qualitatively similar when using a balanced panel. Results available upon request.

area holding everything else fixed. With an expansion as large and fast as this one, there are likely to be general equilibrium effects on the labor market for day care teachers that could make the full effect larger. It is important to separate labor market effects from effects due to the structure of the market because increasing the number of early childhood educators would help with the labor market effects but not the structural ones. Because there seems to be a reduction in day care center seats as a result of the structure of the child care market and not only due to labor market effects, increasing the supply of early childhood workers will not be sufficient to counteract the negative impact of UPK on the day care market.

6.2 Robustness of reduction in infant and young toddler seats

One advantage of using a hexagon grid to define the geography is that it is straightforward to test robustness to the original grid placement. I will consider what happens to the estimates when the grid is shifted and when the hexagons are medium size or large size versus the small size in the main specification.

First, I will look at robustness to shifting the grid. If results are robust and not just an artifact of where the grid happened to be placed, we would expect to see similar results across shifts. I will consider a total of six shifts including the original placement. The shifts are of a “half hexagon” each, and there are three “half hexagons” moving east-west and two “half hexagons” moving north-south, for a total of six shifts. See Figure A.4 for a visual.

Figure A.5 plots the coefficients from the fully-dynamic model and demonstrates that the results are largely robust to shifting the grid. Table A.4 contains the point estimates from the semi-dynamic specification for the interactions of the treatment variable with indicators for the years 2017 and 2018. Note that “East 0, South 0” and column (1) in the table are the original grid. All of the 2018 coefficients indicate a decline of about one to two day care center seats for younger children for each new UPK in the hexagon. Although in one of the shifts the point estimate is not significant at the 95% confidence level, in two others, the effect is over two seats per UPK location by 2018.

The average decline is 1.6 seats per new UPK site as of 2018. Using the average of 1.6 seats, I estimate that the entire UPK program leads to a reduction in the number of infant and young toddler seats of 2,700. With a decline of 1.3 seats per site, the original grid represents a result near the median. It is worth noting, however, that it could have happened that

the “East 2, South 1” shift (column (6)) with the smallest and least statistically significant results could have been the original boundaries. If these had been the given administrative boundaries, it would be easy to think that perhaps UPK did not have a significant effect on nearby day care centers serving younger children when in reality, the remaining boundary shifts demonstrate that is not the case for the average grid position. Being able to test for this robustness is the value of this approach.

Next, I will consider medium-size and large-size hexagons. The medium-size hexagons cover about 20 blocks, while the large ones cover about 40 blocks. In general, as a grid gets larger, we would expect the effect of neighboring hexagons to become less important since they represent UPKs that are farther away from the average point in the hexagon. In this case, the neighbor effects are already zero, so we would expect them to stay zero. When the hexagon size is initially increased, it could be the case that more of the relevant market is being included in the hexagon. In this case, we would expect that the effect of own-hexagon UPKs would be larger as a level since the area being incorporated is also affected by the UPKs in the smaller hexagons, but we might expect the estimate to fall as a percent if the effect dissipates with distance. However, if the area becomes too large and begins to incorporate areas that are not part of the same market, then even the own-hexagon UPKs may not have a large effect on the area as a whole, so the estimate will be attenuated and, for large enough hexagons, eventually go to zero. We might also expect that for larger hexagons, edge effects could be more important, which could make point estimates vary more across shifts.

Figure A.7 plots the coefficients from equation (2) for these two larger grids. The effect is slightly attenuated and has much larger standard errors for the medium-size grid with an area of 20 blocks, and it goes to zero for the large-size grid with an area of 40 blocks. These results indicate that at least for the margin of the opening and shutdown decision, the market that is affected by a given UPK is fairly local. Figure A.8 demonstrates that, as expected, the point estimates for the medium-size grid vary much more across shifts, possibly indicating that edge effects have become more important.

6.3 Heterogeneity: Reduction entirely in high-poverty neighborhoods

The welfare consequences of the reduction in day care center seats depends on the outside option for the marginal families who would have used a day care center in the absence of Universal Pre-K but no longer will due to the spillovers from UPK on the child care market for younger children. Even if the day care center would have been the highest quality option available, in the absence of the center, wealthier families will have more alternative options available to them and are therefore likely to have a higher quality outside option than poorer families. Welfare losses will therefore be larger if the crowd out affects poorer families more than wealthier families.

In order to address this question, I estimate a model to determine whether the effect of Universal Pre-K sites on availability of child care for younger children is stronger in relatively wealthier or poorer areas. I split the hexagons at the median fraction of individuals living below 200% of the federal poverty line. The median hexagon has 30% of individuals with income below this level. I interact indicator variables for the hexagon having above median or below median poverty with the Universal Pre-K exposure variable and year indicators:

$$Y_{ht} = \beta_0 + \sum_{k \neq 2014} [\beta_{ak} * \mathbb{1}_{t=k} * New_h * AboveMed_h + \beta_{bk} * \mathbb{1}_{t=k} * New_h * BelowMed_h] + X_t + \sigma_h + X_t * h + \epsilon_{ht} \quad (5)$$

This is the same regression as equation (2) except for the additional interactions with *AboveMed_h* and *BelowMed_h* and the exclusion of effects from UPKs in neighboring hexagons. The latter are excluded for the sake of brevity since there was little effect from these in the main specification. *AboveMed_h* and *BelowMed_h* are indicators for whether the hexagon has above median or below median fractions of people living below 200% of the federal poverty line. This fraction is determined by applying the American Community Survey 5-year estimates at the block group level to each block within the block group and calculating for each hexagon the average poverty level for the blocks whose centroids are in that hexagon weighting by Census 2010 population at the block level. There are 17 hexagons in the main hexagon grid that have zero population, or 0.6% of the 2,658 hexagons, and these are dropped from the analysis.

Observations are at the hexagon-year level. Y_{ht} is the capacity of day care centers serving younger children in hexagon h as of $1/1/t$, $\mathbb{1}_{X_t=k}$ is an indicator for year k , and New_h is the net new number of full-day UPK sites in the hexagon in 2014. The regression includes year and hexagon fixed effects and hexagon-specific linear time trends. The data spans years 2010-2018; 2014 is the left out year because $1/1/2014$ is the last observation before the expansion, and 2010 is left out in order to identify the linear time trends. Standard errors are clustered at the hexagon level.

As seen in Figure 4, which plots the β_{ak} and β_{bk} coefficients from equation (5), this heterogeneity analysis reveals that the entire decline in seats from the Universal Pre-K introduction occurs in the relatively poorer areas. There is no discernible effect of new UPK sites in the relatively wealthier areas. In poorer areas, by 2.5 years after the expansion, each new UPK site led to just over two fewer seats at day care centers for younger children, a reduction of over one-third.⁴⁴ This result is very robust across the six shifts of the hexagon grid (see Figure A.6). The effect on day care centers in poorer areas is even apparent in the one shift that had less clear results in aggregate (“East 2, South 1”). The addition of Universal Pre-K sites led to a decrease in child care availability for younger children concentrated in the poorest areas.

A likely explanation for the disparate response across areas of different income levels is that families in poorer areas have higher price elasticities, so their demand for child care is more price sensitive. When day care centers in wealthier areas had reduced enrollment of 4-year-olds due to the Universal Pre-K expansion, they may have been able to raise their prices and still attract enough families to stay in business. On the other hand, day care centers in poorer areas may not have been able to attract enough families if they raised prices, so they closed (or did not open) instead. An additional factor in poorer areas is that there are likely more day cares serving children using child care subsidies. Since the subsidy rates are fixed and set by the state, there is no room to adjust prices for these infants and toddlers, leaving only the quantity and quality margins available for adjustment.

⁴⁴This reduction is relative to the 5.8 seats in 2014 in day care centers serving children between the ages of 0 and 2 in hexagons with over 30% of people living below 200% of the federal poverty line that had an increase in the number of Universal Pre-K sites. Hexagons with new UPK sites had higher average day care capacity relative to ones that did not get a new UPK site.

6.4 New UPK Sites increase licensed home day cares very nearby

I now turn to the effect of UPK sites on home day cares. Given that in the absence of UPK, some 4-year-olds might instead go to home day cares for child care, we would expect that some home day cares would close and care might be consolidated. In this scenario, and in the absence of other regulations, even though the total number of home day care seats declines, the only effect on the availability of care for younger children may be the distance parents need to travel to find care. However, New York regulations require that there be only two children under the age of 2 years for each care provider, so a consolidation of seats could lead to fewer slots available for infants and toddlers. As an extreme example, if all home day cares were at capacity for children under the age of 2 years and some of them closed down, it would directly decrease availability and use of care for children under age 2. Because I do not know actual enrollment by age, the results here can only be suggestive indications of possible increases or decreases in availability of care for younger ages.

As seen in Figure 5, which plots the coefficients from equation (2), UPK may increase the total capacity of home day cares very close to new UPKs. Two explanations for an increase in the number of home day cares are that they provide before- and after-school care or sibling care for siblings of 4-year-olds in the UPK. Home day cares may have more flexibility than day care centers to be open for just a few hours a day before and after school or to provide care to just a few children. This result is not especially robust across shifts (see Figure A.9), but at the very least, a new UPK does not appear to decrease the supply of home day cares in its immediate vicinity.

Considering a larger market with a medium-size grid, Figure 7 shows that there may be a decrease in capacity of home day care centers in some areas when a larger market is considered. Part (a) of this figure plots the coefficients on the interaction terms between year and exposure intensity from the main specification, and it appears that UPKs have little effect on the supply of home day cares.

However, the aggregate result hides important heterogeneity. Part (b) of Figure 7 shows that in wealthier areas, there is a statistically-significant increase in home day care capacity, even when considering this larger market. By 2017, there are about 7 more seats in home day cares for each new UPK site in wealthier areas. Based on licensing regulations, approximately one-third of these seats, or just over 2 seats, would be available to children under age 2. On

the other hand, in the poorer areas, if anything, there is a decrease in the capacity of home day care centers close to new Universal Pre-K sites. Therefore, it does not appear that the decrease in day care center capacity is being offset by an increase in licensed home day cares. These results are confirmed by an analysis of robustness to shifting the grid, shown in Figure A.11.

Even if the decrease in day care center capacity in poorer areas were offset by an increase in licensed home day care capacity, we would still be concerned about a decrease in the quality of care as day care centers generally provide the observably highest quality of care. For example, children at licensed home day cares watch on average 1.5 hours of television per day, compared to only a few minutes on average for children in centers (Bassok *et al.*, 2016a). But *without* an increase in licensed home day cares, either younger children are being shifted to informal care arrangements, which are typically even lower quality than home day cares, home day cares are taking more children per home, or their parents are adjusting their labor force participation.

6.5 Evidence on declines in day care center quality

This section considers whether existing day care centers adjusted their quality in response to a new UPK site nearby. To do so, I consider whether there were any changes in safety that can be observed in inspections data. The three outcomes I consider are whether there were any public complaints filed against the day care center in a given year, the number of violations at their initial inspection, and whether reinspection was required.

As seen in Panel (d) of Figure 8, there may be a drop in the number of inspections per year in day cares closer to new UPK sites. Inspectors are assigned to regions, so this finding is consistent with increased case loads due to the new UPK site. I show this result because it could cause some correlation between time between inspections and number of violations. In order to focus on quality decreases due to increased competition from UPK and not quality decreases due to a decrease in inspection frequency, I include the control for time since last inspection in the regressions analyzing inspection results.

Figure 8 plots the coefficients on the interactions between the year indicators and hexagon exposure status. Although none of the coefficients are individually statistically significant at the 5% level, there does appear to be an increase in the total number of violations at each

inspection, in the probability reinspection will be required, and whether there were any complaints in a given year. In all cases, day cares are affected by new UPK sites in their hexagon but not very affected by new sites in neighboring hexagons. The lack of an effect from UPK sites in neighboring hexagons is consistent with the prior results indicating that the relevant markets for day cares are likely relatively small.

Table 5 presents the coefficients from a difference-in-differences regression. These estimates pool the post-period coefficients to test whether they are jointly significant. Since the effects are growing in the post-period with only small effects in the first year after implementation, pooling likely underestimates the total effect. For each new UPK in the hexagon, the number of inspection violations for a day care center increased by 0.28 per inspection, which is an increase of 13% from a mean of 2.2 violations. There was also a modest increase in the probability of requiring reinspection. For each new UPK in its hexagon, a day care center was 2.3 percentage points more likely to require reinspection after an initial inspection, about a 5% increase compared to the mean reinspection rate. Together, these results show that day care center inspection results declined for centers near new UPK sites, consistent with the theory that day care centers may cut corners in order to cut costs when faced with reduced enrollment due to UPK.

The results on complaints can help clarify whether the worse inspection results translated into meaningful decreases in quality. If the public made more complaints, then that may indicate that the day care center has declined in quality enough that parents or other members of the public were concerned. Because parents may be a significant group of complainants, a null result on complaints is difficult to interpret since a decline in enrollment means there are fewer parents to make complaints. The point estimate from Column (3) of Table 5 indicates that the probability of a complaint in a given year increased by 2 percentage points for each new UPK site in the same hexagon as the day care center, which is an increase of 15% relative to the mean, though the estimate is not quite significant at the 10% level ($p=0.11$).

These inspections results provide suggestive evidence that day cares may have responded to crowd out from UPK in other ways in addition to closing or not opening. When faced with lower enrollment, they may have reduced quality in order to lower their marginal costs and stay in operation.

6.6 Number of pre-K switchers to crowd out one infant seat

This section goes through a back-of-the-envelope calculation to estimate the ratio of 4-year-olds who switch from private to public pre-K to the reduction in infant and toddler seats. This parameter could be applied to other pre-K expansions to estimate the number of infant and young toddler seats that will be crowded out based on the expected number of students switching from private to public pre-K.

First, I will estimate the number of students who switch from private pre-K to public pre-K in the presence of Universal Pre-K. There were 15,793 more public pre-K students in NYC in 2015-16 than in 2013-14, and the number enrolled has held pretty steady since 2015-16. This number is smaller than the increase in the number of full-day students because there were more half-day slots in 2013-14 than in 2015-16.

Next, I need to estimate how much of the increased enrollment was from students who would not have enrolled in pre-K otherwise versus how much is from students who would have enrolled in private pre-K. [Fitzpatrick \(2010\)](#) estimates that UPK increases enrollment in pre-K in urban areas by 5 percentage points. There were approximately 100,000 4-year-olds in NYC as of the 2010 Census, so a five percentage point increase translates into approximately 5,000 children who would not have attended pre-K otherwise.

Subtracting the 5,000 children switching from no pre-K to public pre-K from the total increase in public pre-K students leaves 10,793 who switch from private to public. Some of these children would have attended pre-K at private schools, not day care centers. To calculate the desired parameter, I need to isolate those who switched from *day care centers* only. Using enrollment data from the Private School Survey and the capacity data from my data set, I estimate that 81% of the private pre-K students are at day care centers.⁴⁵ Assuming crowd out is proportionally distributed between centers and schools, taking 81% of the 10,793 switchers yields 8,742 who switch from private pre-K in day care centers to

⁴⁵Because capacity data for pre-K in the Private School Survey lumps together private pre-K for 4-year-olds with public pre-K and also private pre-K for age 3, not just age 4, I use kindergarten enrollment as a proxy for age 4 private preschool enrollment at the school. This yields an estimated 12,398 seats. For day care centers, I assume that seats are allocated by age in the same ratio as regulated maximum group size and that every day care includes each age group. Centers can have 12 two year olds in a group, 15 three year olds, or 30 four year olds. Then the fraction of day care center seats available for 4-year-olds is $20/(20+15+12)=43\%$. There were 123,351 day care center seats for children between the ages of 2 and 5 as of 1/1/2014, so I estimate that 52,490 of these were available to 4-year-olds.

public pre-K as a result of the implementation of Universal Pre-K, which is the numerator in the calculation.

Now I need to calculate the number of infant and young toddler seats crowded out. Using the average estimate from the main specification, there were 1.6 fewer day care center seats for children under age 2 for each new full-day UPK site. The number of new full-day sites added by the Universal Pre-K expansion was approximately 776.⁴⁶ Then the decrease in the number of center-based seats for children under age 2 due to the expansion was 1,242. Putting these numbers together yields the result that one under 2 seat was lost for every seven pre-K children who switched from private pre-K to public pre-K.⁴⁷

7 Conclusion

As a result of research demonstrating the importance of early childhood education, there has been a push to expand public schooling to include pre-kindergarten for 4-year-olds. Decreasing the starting age of free public schooling to 4 years old from 5 years old is proposed as a way to reduce inequality, and it is particularly appealing in light of recent research that human capital investments have larger impacts the earlier that they occur. By that same token, investments at ages 0 to 3 may be even more important than investments at age 4, but less attention has been given to this age group.

Extending schooling to 4-year-olds may be more tractable given the existing infrastructure and means of delivering services. However, given the importance of early childhood education, it is important to determine whether this focus on 4-year-olds has had any unintended consequences for even younger children ages 0 to 3.

This paper finds that public pre-K for 4-year-olds had unintended consequences on the child care market for younger children. New UPK sites reduced the number of center-based day care seats for children under age two, with the whole program leading to a decline of about 2,700 seats. This estimate is an underestimate of the overall effect if there are general

⁴⁶The 776 new sites uses the number of full-day sites in 2018 (1,728) and subtracts from it the number of full-day sites in 2013 (952). I use 2018 data because the number of full-day sites was not readily available for years 2015, 2016, or 2017. Since the total number of students has held relatively steady since 2015, the number of sites in 2018 should be a good approximation of the effect of the expansion.

⁴⁷This is from the calculation of 8,742 switchers divided by 1,242 lost infant and young toddler seats.

equilibrium effects on the labor market that caused some day cares to close due to a shortage of early childhood educators. It also only includes children under 2 years. There may also be fewer seats for children ages 2 years and 3 years.

Children displaced from center-based care to home day cares or other types of care are likely to experience a decrease in quality of care that could affect their human capital accumulation. I further find evidence of a decline in quality at day care centers that remain in operation, suggesting that even those children who are still able to attend a day care center may experience lower quality care.

I have shown that public pre-K can decrease the availability of care for children younger than two and suggestive evidence of a decline in day care center quality. Important areas for future research include documenting effects on children ages 2 and 3 and more thoroughly examining other channels through which public pre-K can affect the child care market, such as price and quality of care.

A primary reason for these spillovers may be due to the pricing model in the child care market where older children cross-subsidize more costly-to-care-for younger children. I rationalize this pricing behavior with a model of switching costs. Day cares with local monopoly power have an incentive to attract infants at a lower initial profit because once the infants are at the center, it increases their chance of remaining at that center for pre-kindergarten, where the day care can extract a higher monopolist profit. Note that for this mechanism to be at play, there must be crowd out of 4-year-olds from the private market. Areas with less crowd out will experience less of a decline in care for younger children.

There are some additional explanations for why public pre-K for 4-year-olds could reduce availability of care for younger children. There are likely labor market effects from increased demand for early childhood educators for better-paying jobs in the public sector. Although changes in the child care worker labor market is a likely channel through which public pre-K may affect the child care market, it is likely a small part of the effect I measure as my identification strategy estimates the impact of one new UPK site on the nearby area holding the city-wide labor market fixed. Potentially more relevant for NYC, where space is at a premium, there could be an inelastic supply of buildings and rooms suitable for day care centers, and more of those places may have been shifted to 4-year-old care from care for younger children. Estimating the relative importance of these different channels is an important area for future research.

Although I find negative effects in the context of New York City, there are some forces from public pre-K that could lead to improvements in the child care market for younger children. Perhaps in a different context, these could outweigh the aforementioned forces pushing in the opposite direction. For example, a decrease in lifetime child care expenses induce an income effect that increases demand for child care at younger ages. Also, if public pre-K can be in day care classrooms and reimbursement rates are high enough, it may help day cares stay afloat and continue to offer care for younger children (or begin to offer care for younger children, particularly if there is a sibling-demand effect). Finally, there could be a “crowding in” effect where parents, seeing the emphasis the government is placing on quality early childhood education, may increase their demand for quality early childhood education at younger ages, which could lead to an expansion in the number of seats for younger children. Understanding these effects and what could mitigate the impact of public pre-K for 4-year-olds on the child care market for younger children is becoming increasingly important as more cities and states push to expand the availability of public pre-K to more 4-year-olds and even to 3 year olds. Given the result that the crowd out takes place entirely in areas with higher levels of poverty, policy makers could consider increasing the subsidy rate for infants so it is more in line with the actual cost of providing care, as opposed to the market rate, to reduce market exit and increase entry.

References

- BAKER, M., GRUBER, J. and MILLIGAN, K. (2008). Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy*, **116** (4), 709–745.
- BASSOK, D. (2012). Competition or collaboration?: Head start enrollment during the rapid expansion of state pre-kindergarten. *Educational Policy*, **26** (1), 96–116.
- , FITZPATRICK, M., GREENBERG, E. and LOEB, S. (2016a). Within- and between-sector quality differences in early childhood education and care. *Child Development*, **87**, 1627–1645.
- , — and LOEB, S. (2014). Does state preschool crowd-out private provision? the impact of universal preschool on the childcare sector in oklahoma and georgia. *Journal of Urban Economics*, **83**, 18–33.
- , MILLER, L. C. and GALDO, E. (2016b). 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*, **53**, 87–98.
- BAUM-SNOW, N. and MARION, J. (2009). The effects of low income housing tax credit developments on neighborhoods. *Journal of Public Economics*, **93**, 654–666.
- BERGER, M. C. and BLACK, D. A. (1992). Childcare subsidies, quality of care, and the labor supply of low-income, single mothers. *Review of Economics and Statistics*.
- BERNAL, R. and KEANE, M. P. (2010). Quasi-structural estimation of a model of childcare choices and child cognitive ability production. *Journal of Econometrics*, **156** (1), 164–89.
- BLAU, D. M. (2007). Unintended consequences of child care regulations. *Labour Economics*, **14** (3), 513–38.
- and MOCAN, H. N. (2002). The supply of quality in child care centers. *Review of Economics and Statistics*, **84** (3), 483–496.
- and ROBINS, P. K. (1988). Child-care costs and family labor supply. *Review of Economics and Statistics*, **70** (3), 374–381.

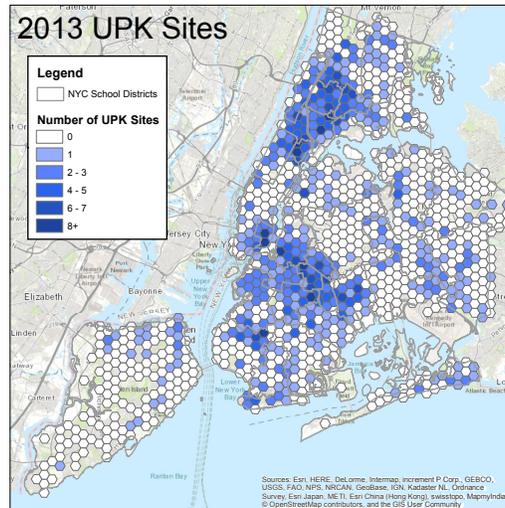
- CASCIO, E. (2009a). Do investments in universal early education pay off? long-term effects of introducing kindergartens into public schools. *NBER Working Paper No. 14951*.
- CASCIO, E. U. (2009b). Maternal labor supply and the introduction of kindergartens into american public schools. *Journal of Human Resources*, **44** (1), 140–170.
- and SCHANZENBACH, D. W. (2013). The impacts of expanding access to high-quality preschool education. *Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution*, **2013**, 127–192.
- CELLINI, S. R. (2009). Crowded colleges and college crowd-out: The impact of public subsidies on the two-year college market. *American Economic Journal: Economic Policy*, **1** (2), 1–30.
- CHIPTY, T. (1995). Economic effects of quality regulations in the day-care industry. *American Economic Review, Papers and Proceedings*, **85** (2), 419–424.
- and WITTE, A. D. (1999). An empirical investigation of firms’ responses to minimum standards regulations. *Children and Youth Services Review*, **21**, 111–146.
- CLEMENS, J. and GOTTLIEB, J. D. (2017). In the shadow of a giant: Medicare’s influence on private payment systems. *Journal of Political Economy*, **125** (1), 1–39.
- CURRIE, J. and WALKER, R. (2011). Traffic congestion and infant health: Evidence from e-zpass. *American Economic Journal: Applied Economics*, **3** (1), 65–90.
- DAVIS, L. W. (2011). The effect of power plants on local housing values and rents. *Review of Economics and Statistics*, **93**, 1391–1402.
- DIAMOND, R. and MCQUADE, T. (2016). Who wants affordable housing in their backyard? an equilibrium analysis of low income property development. *NBER Working Paper No. 22204*.
- DINERSTEIN, M. and SMITH, T. (2014). Quantifying the supply response of private schools to public policies. *SIEPR Working Paper*.

- DOROMAL, J. B., BASSOK, D. and WONG, V. C. (2018). How safe are early childhood education programs? new evidence from the universe of all licensed centers in north carolina. *EdPolicy Works Working Paper Series No. 65*.
- FINKELSTEIN, A. and MCKNIGHT, R. (2005). *What Did Medicare Do (And Was It Worth It)?* Working Paper 11609, National Bureau of Economic Research.
- FITZPATRICK, M. D. (2010). Preschoolers enrolled and mothers at work? the effects of universal prekindergarten. *Journal of Labor Economics*, **28** (1), 51–85.
- FOOTE, A., KUTZBACH, M. J. and VILHUBER, L. (2017). Recalculating... : How uncertainty in local labor market definitions affects empirical findings. *Center for Economic Studies Research Paper*, **17-49**.
- FRIEDMAN-KRAUSS, A. H., BARNETT, W. S., WEISENFELD, G., KASMIN, R., DICRECHIO, N. and HOROWITZ, M. (2017). *The state of preschool 2017: State Preschool Yearbook*. Tech. rep., National Institute for Early Education Research, New Brunswick, New Jersey.
- GARCÍA, J. L., HECKMAN, J. J., LEAF, D. E. and PRADOS, M. J. (2018). Quantifying the life-cycle benefits of a prototypical early childhood program. *NBER Working Paper No. 23479*.
- GELBACH, J. B. (2002). Public schooling for young children and maternal labor supply. *American Economic Review*, **92** (1), 307–322.
- HALES, T. C. (2001). The honeycomb conjecture. *Discrete & Computational Geometry*, **25** (1), 1–22.
- HARARI, M. and LA FERRARA, E. (Forthcoming). Conflict, climate, and cells: A disaggregated analysis. *Review of Economics and Statistics*.
- HELBURN, S. W. and HOWES, C. (1996). Child care cost and quality. *The Future of Children*, **6** (2), 62–82.
- HERBST, C. M. (2018). The rising cost of child care in the united states: A reassessment of the evidence. *Economics of Education Review*, **64**, 13–30.

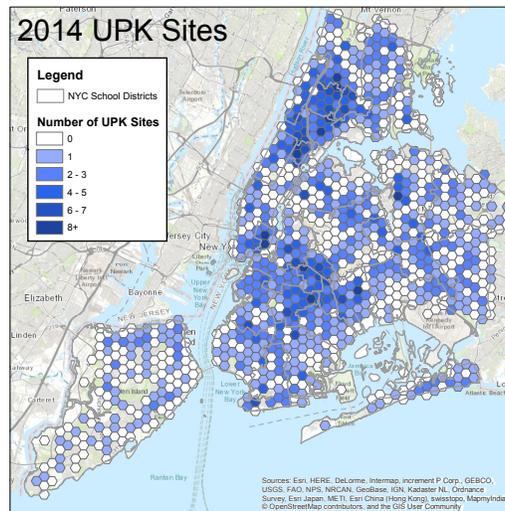
- HOTZ, V. J. and XIAO, M. (2011). The impact of regulations on the supply and quality of care in child care markets. *American Economic Review*, **101**, 1775–1805.
- KLINE, P. and WALTERS, C. (2016). Evaluating public programs with close substitutes: The case of head start. *Quarterly Journal of Economics*, **131** (4), 1795–1848.
- LOEB, S., BRIDGES, M., BASSOK, D., FULLER, B. and RUMBERGER, R. W. (2007). How much is too much? the influence of preschool centers on children’s social and cognitive development. *Economics of Education Review*, **26** (1), 52–66.
- RODGERS, L. P. (2018). Give credit where? the incidence of child care tax credits. *Journal of Urban Economics*, **108**, 51–71.
- SCHULTE, B. and DURANA, A. (2016). The new america care report. *New America Policy Paper*.
- SHROFF, R., DUNKS, R., LIM, J., WANG, H. and CASTRO, M. (2014). Allocation of pre-kindergarten seats in new york city. *The Association for the Advancement of Artificial Intelligence 2014 Workshop on Semantic Cities*.
- SUSIN, S. (2002). Rent vouchers and the price of low-income housing. *Journal of Public Economics*, **83** (1), 109–152.
- VANDELL, D. L. and WOLFE, B. (2000). Child care quality: Does it matter and does it need to be improved? *Institute for Research on Poverty Special Report no. 78*.

Figure 1: Expansion of Full-Day Public Pre-K Sites in New York City in 2014

(a) Number of Full-Day Sites in 2013-14



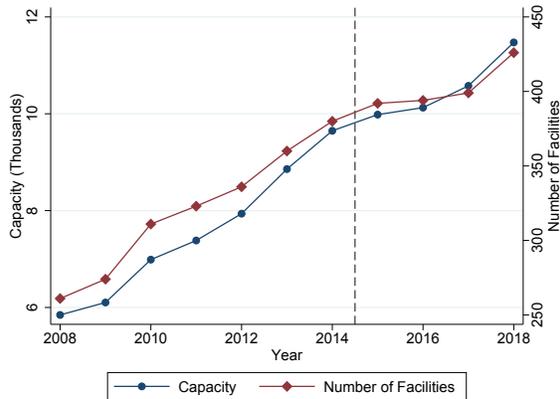
(b) Number of Full-Day Sites in 2014-15



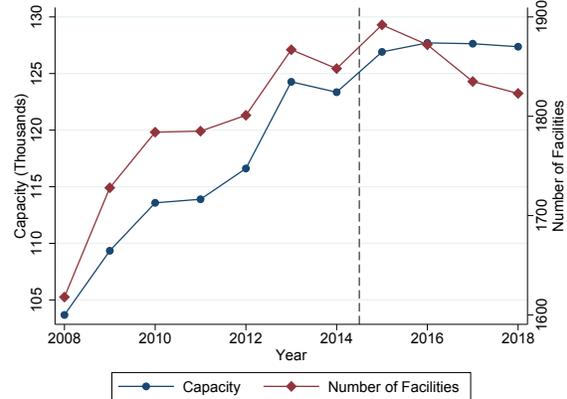
Notes: Maps display the number of full-day Universal Pre-K sites in each hexagonal area in the 2013-14 (a) and 2014-15 (b) school years. New York City expanded public pre-K for 4-year-olds rapidly in the 2014-15 school year, adding 432 new sites and 25,000 seats in one year. Of the 998 hexagons displayed here, 372 had more public pre-K sites in 2014 than they did in 2013. The hexagons in this grid cover approximately 20 city blocks and are the same size as the hexagons in the medium-size grid referred to in the text.

Figure 2: Stock of Day Care Facilities in New York City as of January 1st Each Year

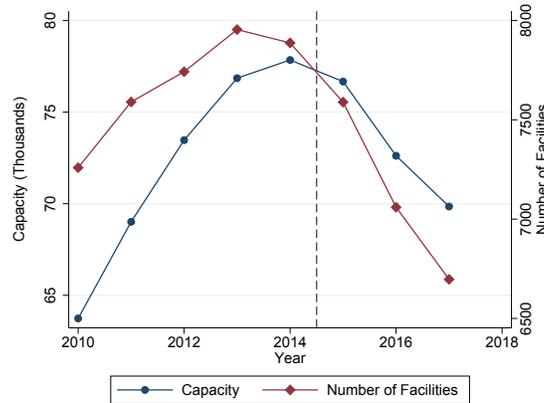
(a) Day Care Centers Serving Younger Children



(b) Day Care Centers Serving Older Children

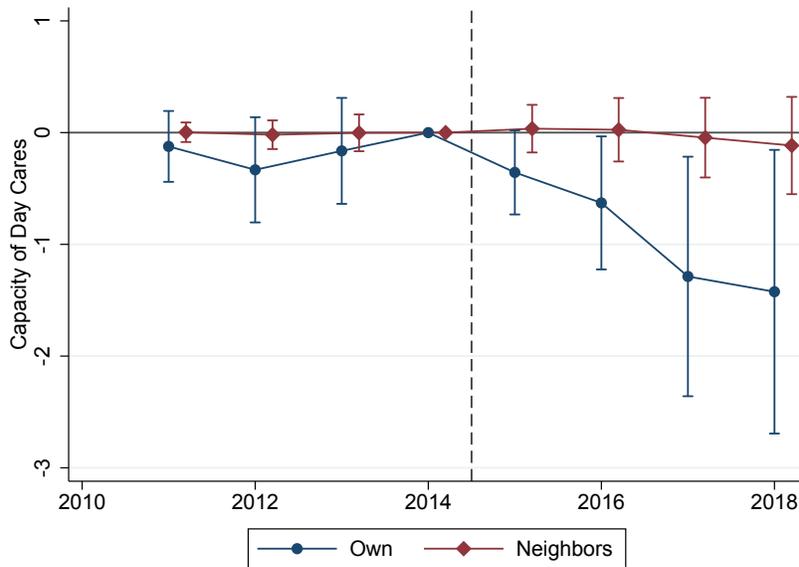


(c) Home Day Cares



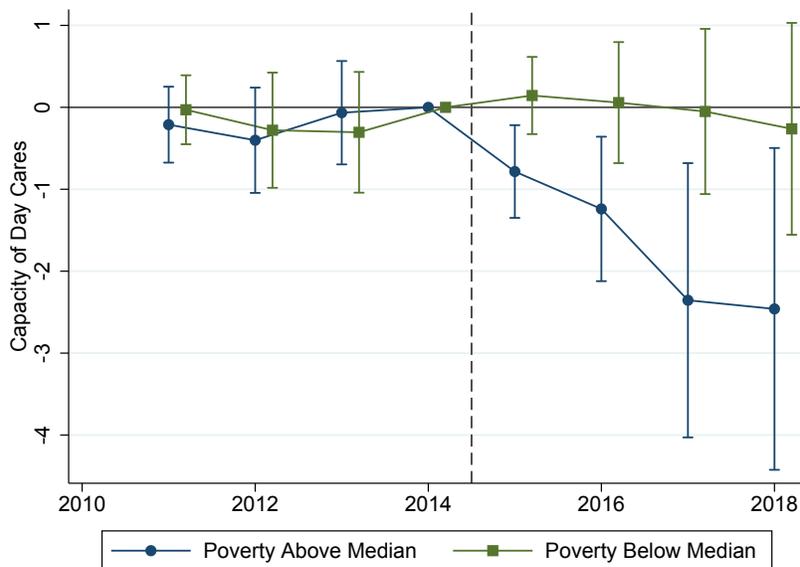
Notes: Figures plot the number and capacity of licensed day care facilities in New York City as of January 1st of each year. Vertical lines indicate the Universal Pre-K expansion in September 2014. Subfigures (a) and (b) are based on day care center licensure data from the NYC Department of Health and Mental Hygiene. Separate licenses are required to care for children under 2 and children ages 2 years through 5 years. Closure dates are estimated based on license expiration date and the date of the last inspection. Refer to Section 4.1 of the text for details. Subfigure (c) is based on home day care licensure data from the New York State Office of Children and Family Services. Approximately one-third of total capacity is available to children under age 2. Refer to Section 2.2 for details.

Figure 3: Effect of a New Pre-K Site on Capacity of Day Care Centers Serving Younger Children



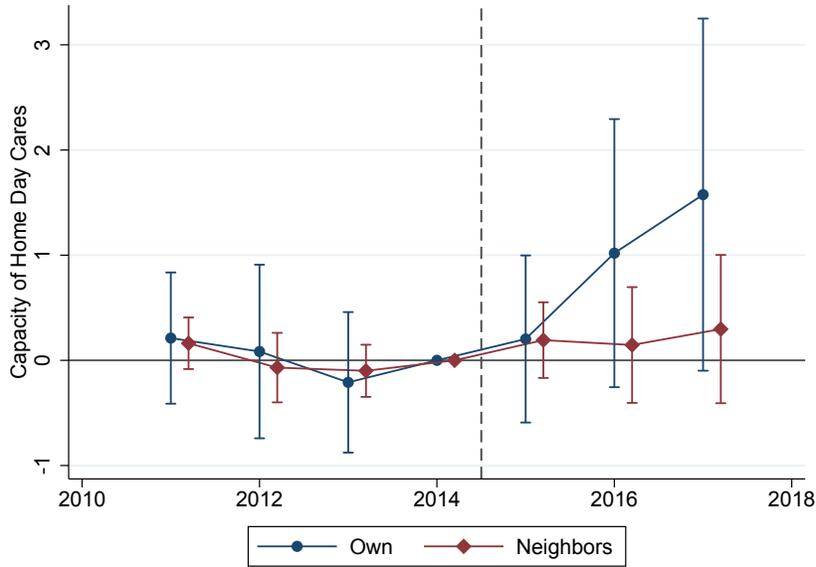
Notes: Figure plots coefficients on interactions between year indicators and the change in the number of Universal Pre-K sites between 2013-14 and 2014-15 in the hexagon (“Own”) and the hexagon’s immediate neighbors (“Neighbors”). Observations are at the hexagon-year level. Hexagons are approximately ten city blocks in area and form a grid that covers New York City. The outcome variable is the total licensed capacity for children between the ages of 0 and 2 in day care centers in the given hexagon. Includes hexagon fixed effects, time fixed effects, and hexagon-specific linear time trends. Bands show the 95% confidence interval with standard errors clustered at the hexagon level.

Figure 4: Effect of a New Pre-K Site on Capacity of Day Care Centers Serving Younger Children: Heterogeneity by Poverty Level



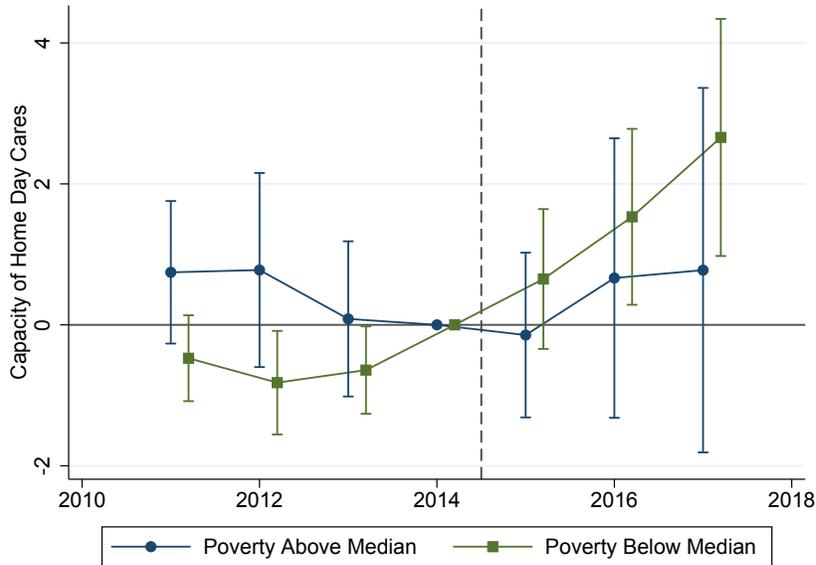
Notes: Figure compares how a new Universal Pre-K site affects the total capacity of day care centers serving ages 0-2 in areas with above-median levels of people living below the 200% poverty line to areas with below-median levels. In the median area, approximately 30% of people live below 200% of the federal poverty line. Figure plots coefficients of interactions of three variables: 1) year indicators, 2) indicators for above-median and below-median fraction of people living below 200% of the federal poverty line as of 2014, and 3) the change in the number of Universal Pre-K sites between 2013-14 and 2014-15 in the hexagon. Observations are at the hexagon-year level. Hexagons are approximately ten city blocks in area and form a grid that covers New York City. The outcome variable is the total licensed capacity for children between the ages of 0 and 2 in day care centers in the given hexagon. Includes hexagon fixed effects, time fixed effects, and hexagon-specific linear time trends. Bands show the 95% confidence interval with standard errors clustered at the hexagon level. Poverty level is determined by applying the American Community Survey 5-year estimates at the block group level to each block within the block group and calculating for each hexagon the average poverty level for the blocks in that hexagon weighting by Census 2010 population at the block level.

Figure 5: Effect of a New Pre-K Site on Capacity of Home Day Cares



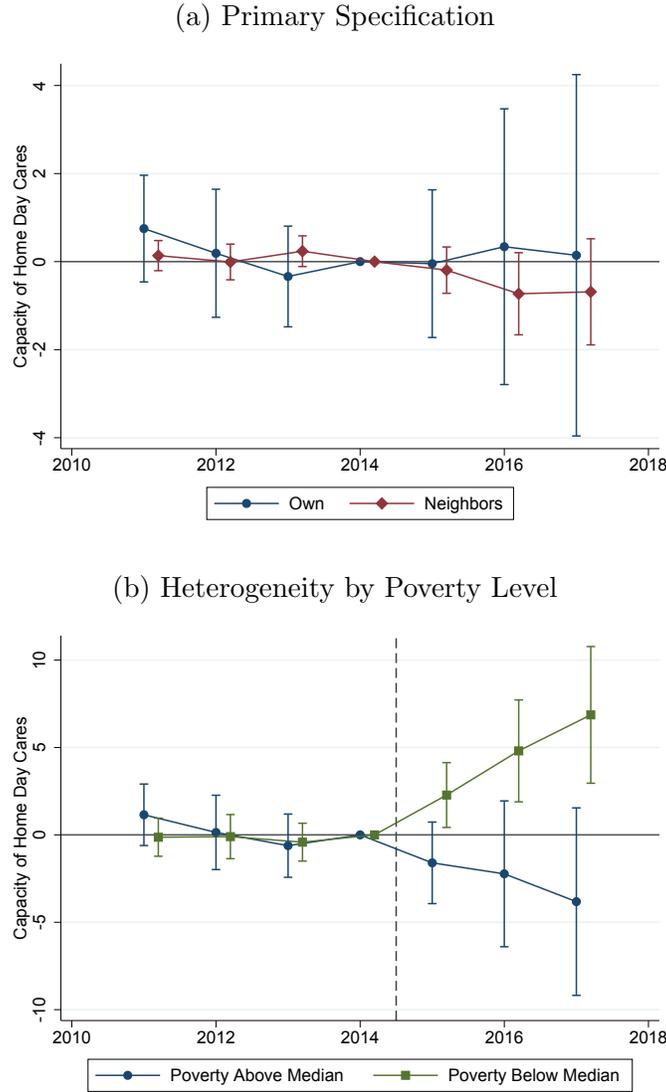
Notes: Figure plots coefficients on interactions between year indicators and the change in the number of Universal Pre-K sites between 2013-14 and 2014-15 in the hexagon (“Own”) and the hexagon’s immediate neighbors (“Neighbors”). Observations are at the hexagon-year level. Hexagons are approximately ten city blocks in area and form a grid that covers New York City. The outcome variable is the total licensed capacity of home day cares in the given hexagon. Note that approximately one-third of licensed capacity is available for children under age 2. Includes hexagon fixed effects, time fixed effects, and hexagon-specific linear time trends. Bands show the 95% confidence interval with standard errors clustered at the hexagon level.

Figure 6: Effect of a New Pre-K Site on Capacity of Home Day Cares: Heterogeneity by Poverty Level



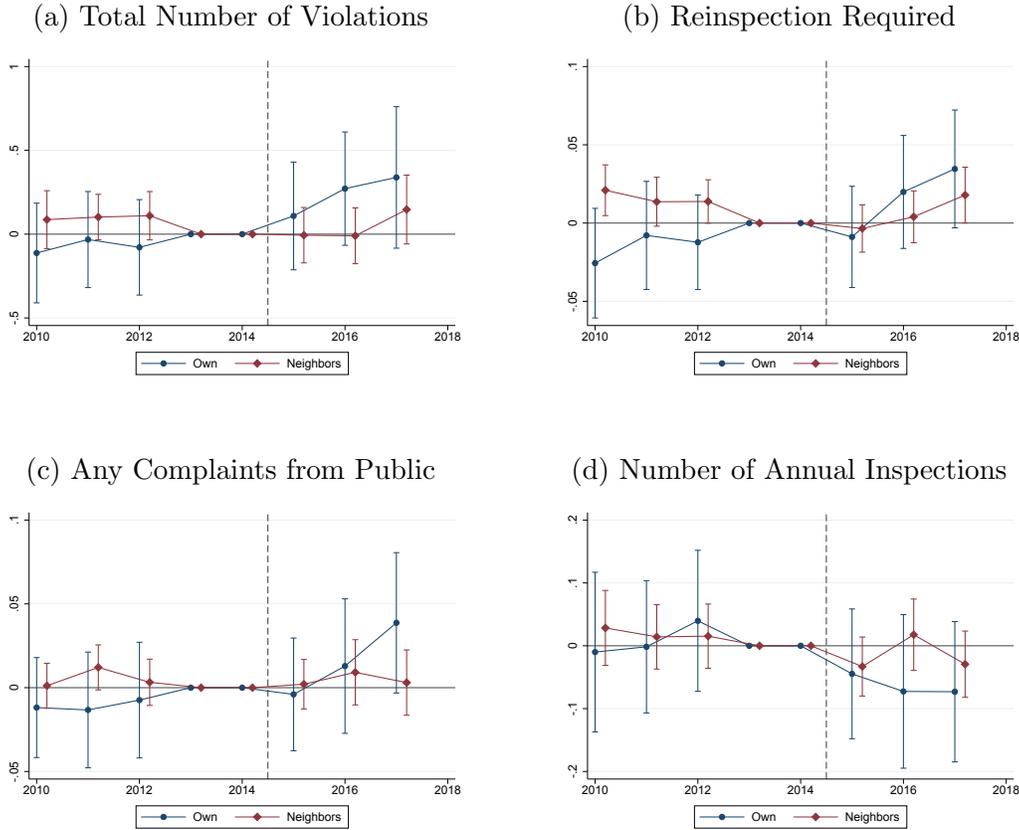
Notes: Figure compares how a new Universal Pre-K site affects the total capacity of home day cares in areas with above-median levels of people living below the 200% poverty line to areas with below-median levels. In the median area, approximately 30% of people live below 200% of the federal poverty line. Figure plots coefficients of interactions of three variables: 1) year indicators, 2) indicators for above-median and below-median fraction of people living below 200% of the federal poverty line as of 2014, and 3) the change in the number of Universal Pre-K sites between 2013-14 and 2014-15 in the hexagon. Observations are at the hexagon-year level. Hexagons are approximately ten city blocks in area and form a grid that covers New York City. The outcome variable is the total licensed capacity of home day cares in the given hexagon. Includes hexagon fixed effects, time fixed effects, and hexagon-specific linear time trends. Bands show the 95% confidence interval with standard errors clustered at the hexagon level. Poverty level is determined by applying the American Community Survey 5-year estimates at the block group level to each block within the block group and calculating for each hexagon the average poverty level for the blocks in that hexagon weighting by Census 2010 population at the block level.

Figure 7: Effect of a New Pre-K Site on Capacity of Home Day Cares: Larger Market



Notes: Figures plot coefficients on interactions between year indicators and a measure of Universal Pre-K exposure in the hexagon (“Own”) and the hexagon’s immediate neighbors (“Neighbors”). In subfigure (a), the measure of Universal Pre-K exposure is the change in the number of Universal Pre-K sites from 2013-14 to 2014-15. In subfigure (b), the measure of Universal Pre-K exposure is, for each hexagon, an indicator for whether there were more Universal Pre-K sites in that hexagon in 2014-15 than in 2013-14. In this specification, the measure of exposure from neighbors runs from 0 to 6 depending on how many neighbors had more sites in 2014. Observations are at the hexagon-year level. Hexagons are approximately *twenty* city blocks in area and form a grid that covers the city. The outcome variable is the total licensed capacity of home day cares in the given hexagon. Note that approximately one-third of licensed capacity is available for children under age 2. Includes hexagon fixed effects, time fixed effects, and hexagon-specific linear time trends. Bands show the 95% confidence interval with standard errors clustered at the hexagon level.

Figure 8: Effect of a New Pre-K Site on Day Care Center Violations



Notes: Figures plot coefficients on interactions between year indicators and the change in the number of Universal Pre-K sites between 2013-14 and 2014-15 in the hexagon that the day care is in. Hexagons are approximately ten city blocks in area and form a grid that covers New York City. Observations in subfigures (a) and (b) are at the inspection level. Observations in subfigures (c) and (d) are at the day care center-year level. Figures (a), (b), and (d) include initial annual inspections only. Figure (c) is an indicator for whether the day care center had any inspections in that year that were prompted by a complaint from the public. Inspections in year 2014 are dropped since based on the Universal Pre-K implementation timeline, it is not clear what part of the year would be in the pre-period and what part would be in the post-period. Day care centers that were themselves new UPK sites in 2014 are dropped. Regressions include day care center fixed effects and year fixed effects. Subfigures (a) and (b) include a control for the number of months since the last inspection. Bands show the 95% confidence interval with standard errors clustered at the hexagon level.

Table 1: Change in Public Pre-K Availability in NYC from 2013-14 to 2014-15

Site Type	2013-14		2014-15		Net Increase	
	Num. Sites	Capacity	Num. Sites	Capacity	Num. Sites	Capacity
CBOs	514 ^a	16,177 ^b	877	35,574	363	19,397
Public	438	15,987	507	21,577	69	5,590
Total	952	32,164	1,384	57,151	432	24,987

^a Includes 455 that have income requirements in one or more of their classrooms. Classes with income requirements are primarily Administration for Children’s Services (ACS) classrooms that combine half-day pre-K with child care and Head Start.

^b Approximately 12,700 of these children are enrolled in ACS programs described in footnote a.

Notes: Table provides the total number and capacity of full-day pre-K sites in NYC in 2013-14, the year before the Universal Pre-K expansion, and 2014-15, the year of the expansion. Note that these are capacities, not enrollments. Because some seats are unused, these do not match enrollment numbers. The difference between these numbers and official Department of Education (DOE) numbers is that DOE excludes ACS programs for low-income children that combine half-day pre-K with child care and Head Start to create a full day. I include these as full-day options in the pre-period because the change to Universal Pre-K in 2014-15 is a programmatic one and does not change the hours that care is available in these locations. The 2013-14 CBO capacity number is only available in aggregate as no microdata on the capacity at each site is available in this year.

Table 2: Child Care Facilities in New York City as of 1/1/2014

Facility Type	Num. Facilities	Total Capacity					Capacity for Younger Children	
		Total	Mean	Min.	Median	Max.	Total	Mean
Day Care Centers Overall	2,228	133,000	59.7	4	45	424	9,649	4.3
Licensed for Younger Children	380	9,649	25.4	4	20	136	9,649	25.4
Licensed for Older Children	1,848	123,351	66.7	8	54	424	0	0
Home Day Cares Overall	7,887	77,842	9.9	3	12	12	31,548	3.3
Single Home Day Cares	2,617	15,310	6.0	5	6	6	5,234	2.0
Group Home Day Cares	5,270	62,532	11.9	7	12	12	26,314	4.0

Notes: Table provides statistics on the licensed child care market in New York City as of 1/1/2014, the last year before the Universal Pre-K expansion. Day care center licenses cover either children under the age of 2 years or ages 2 years through 5 years, and facilities that serve both age groups must have two licenses, one for each age group. All licenses are counted separately in this table. Single home day cares do not have additional employees; care is provided by the owner only. Group home day cares have at least one additional employee. Group home day cares can care for two children under two for each care provider present (up to the total allowed capacity). For these calculations, I have assumed that there is one additional employee since that is the minimum required to care for seven or more children.

Table 3: Effect of New Public Pre-K Site on Stock of Day Care Centers Serving Younger Children

	Capacity		Number of Facilities	
	(1)	(2)	(3)	(4)
$\Delta Sites * (Year = 2017)$	-1.144* (0.588)	-0.967* (0.528)	-0.0553*** (0.0200)	-0.0475*** (0.0181)
$\Delta Sites * (Year = 2018)$	-1.276* (0.693)	-1.060* (0.619)	-0.0609*** (0.0236)	-0.0524** (0.0212)
$\Delta Neighbor Sites * (Year = 2017)$	-0.0388 (0.159)	-0.181 (0.183)	0.00279 (0.183)	0.000461 (0.00574)
$\Delta Neighbor Sites * (Year = 2018)$	-0.109 (0.195)	-0.300 (0.223)	0.00118 (0.00674)	-0.00102 (0.00665)
Dep. Var. Mean	3.408	3.408	0.137	0.137
Dep. Var. Mean (Exposed)	5.690	5.690	0.249	0.249
Incl. Expansions		X		X
N	23922	23922	23922	23922

Note: Observations are at the hexagon-year level. Even columns include public schools that already had full-day pre-k seats in 2013 and increased the number of seats in 2014 as new sites. Includes hexagon and year fixed effects and hexagon-specific linear time trends. Standard errors clustered at the hexagon level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Table provides coefficient estimates on the interaction terms between the 2017 and 2018 indicators and the change in the number of Universal Pre-K sites between 2013-14 and 2014-15 in the hexagon and the hexagon's immediate neighbors from equation (1). Dependent variable is the capacity or number of day care centers serving children between the ages of 0 and 2 in the hexagon. Hexagons are approximately ten blocks in area and cover New York City. Columns (2) and (4) include public schools that had full-day pre-K in 2013-14 and added more seats in 2014-15 as new sites. Universal Pre-K classrooms can also be in community-based organizations, and the analogous information about whether an existing site expanded capacity is not available for those organizations. Includes year and hexagon fixed effects and a hexagon-specific linear time trend. Standard errors are clustered at the hexagon level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Effect of New Public Pre-K Site on Stock of Home Day Cares

	Total Capacity		Infant Capacity		Number of Facilities	
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta\text{Sites}^*(\text{Year}>2014)$	1.084*	0.505	0.339	0.171	0.135**	0.0697
	(0.630)	(0.607)	(0.210)	(0.203)	(0.0649)	(0.0614)
$\Delta\text{Neighbor Sites}^*(\text{Year}>2014)$	0.318	-0.151	-0.112	-0.0443	0.0362	-0.00795
	(0.280)	(0.279)	(0.126)	(0.0940)	(0.0287)	(0.0289)
Dep. Var. Mean	27.19	27.19	9.270	9.270	2.803	2.803
Incl. Expansions		X		X		X
N	21264	21264	21264	21264	21264	21264

Notes: Dependent variable is the capacity or number of home day cares in the hexagon. Hexagons are approximately ten blocks in area and cover New York City. Even columns include public schools that had full-day pre-K in 2013-14 and added more seats in 2014-15 as new sites. Universal Pre-K classrooms can also be in community-based organizations, and the analogous information about whether an existing site expanded capacity is not available for those organizations. Estimates are the average of coefficients from post-treatment years from a model in which each post-period year is interacted with the exposure variable: $Y_{ht} = \beta_0 + \sum_{j=2015}^{2017} [\beta_j * \mathbb{1}_{X_t=j} * New_h + \beta_{nj} * \mathbb{1}_{X_t=j} * NewNB_h] + X_t + \sigma_h + X_t * h + \epsilon_{ht}$. Includes year and hexagon fixed effects and a hexagon-specific linear time trend. Standard errors are clustered at the hexagon level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Effect of New Public Pre-K Site on Day Care Center Violations

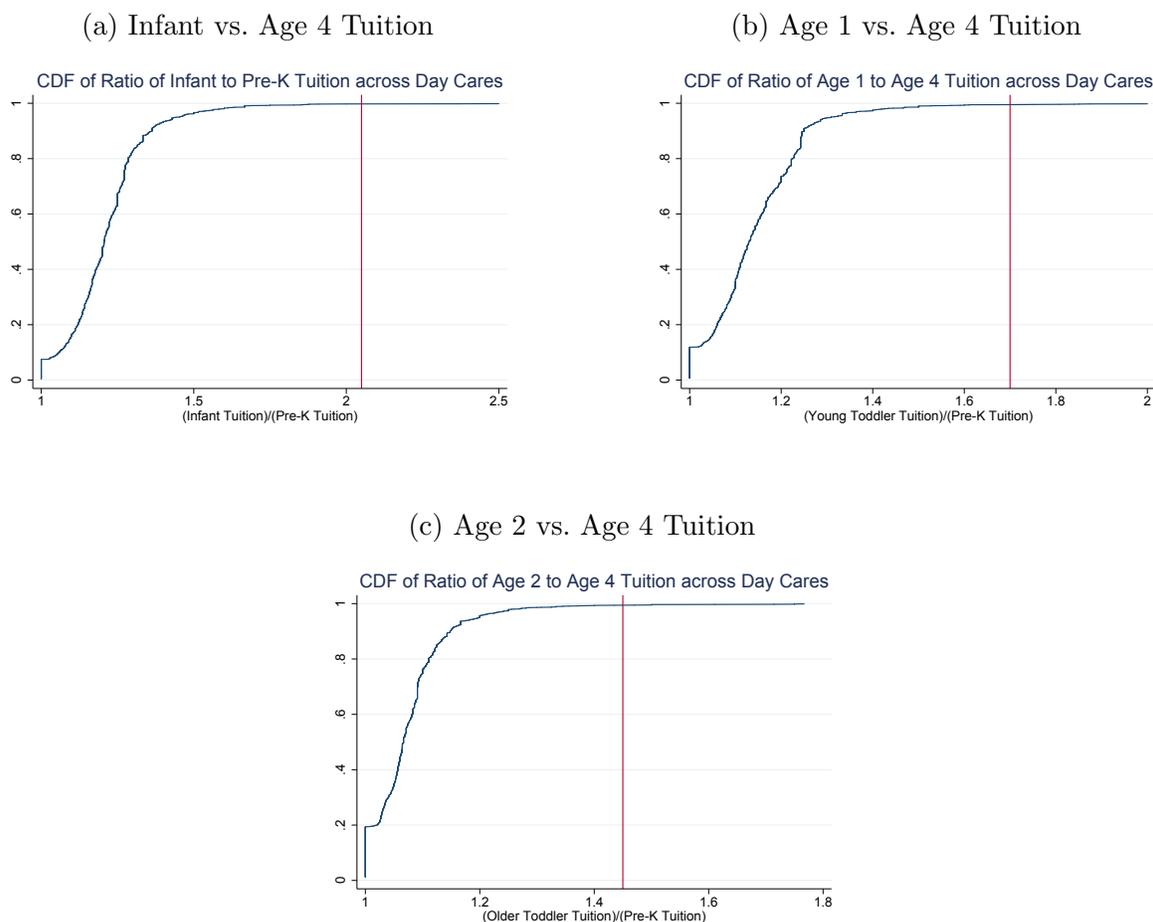
	(1)	(2)	(3)	(4)
	Total Violations	Reinspection Required	Any Complaints	Num. Annual Inspections
$\Delta Sites * (Year > 2014)$	0.278** (0.108)	0.0230** (0.0108)	0.0217 (0.0132)	-0.0692 (0.0428)
$\Delta NeighborSites * (Year > 2014)$	-0.0299 (0.0631)	-0.00564 (0.00517)	0.000891 (0.00659)	-0.0279 (0.0173)
Months Since Last Inspection	-0.000503 (0.00894)	0.00323*** (0.000884)		
Dep. Var. Mean	2.156	0.403	0.140	2.220
N	29437	29437	13260	13260

Notes: This table reports difference-in-differences estimates of the effect of one new full-day public pre-K site in a hexagon on the inspection results of day care centers in that hexagon. Hexagons are approximately ten city blocks in area and cover the city. Observations in columns (a) and (b) are at the inspection level and include only initial annual inspections. The dependent variable in column (a) is the total number of violations recorded in that inspection. The dependent variable in column (b) is an indicator for whether the initial inspection triggered a reinspection. Observations in columns (c) and (d) are at the day care center-year level. The dependent variable in column (c) is an indicator for whether the day care center had any inspections in that year that were triggered by a public complaint against the day care. The dependent variable in column (d) is the number of initial inspections done at that day care in that year. Inspections in year 2014 are dropped since it is not clear what part of the year would be in the pre-period and what part would be in the post-period. All regressions exclude day care centers that were themselves new UPK sites in 2014. Regressions include day care center fixed effects and year fixed effects. Standard errors are clustered at the hexagon level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

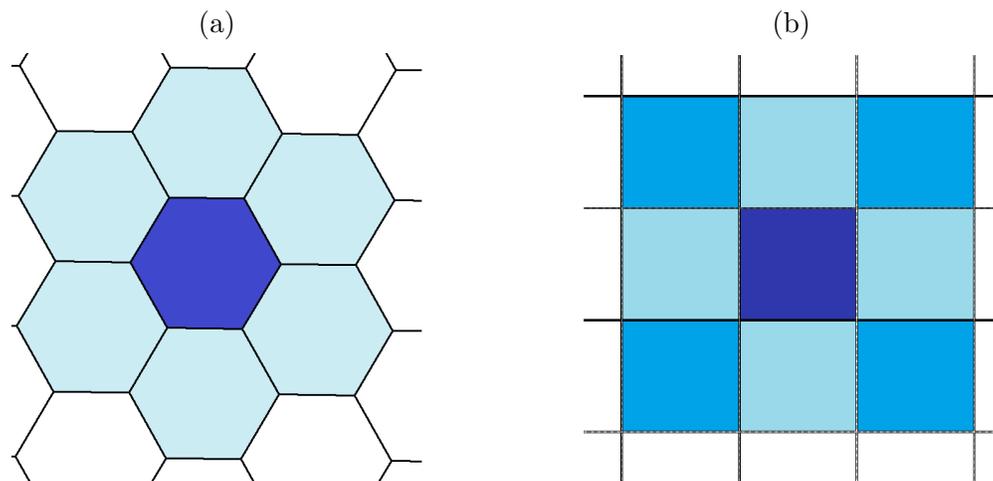
A Additional Figures and Tables

Figure A.1: Ratio of Younger Child Tuition to Age 4 Tuition in the Philadelphia Region



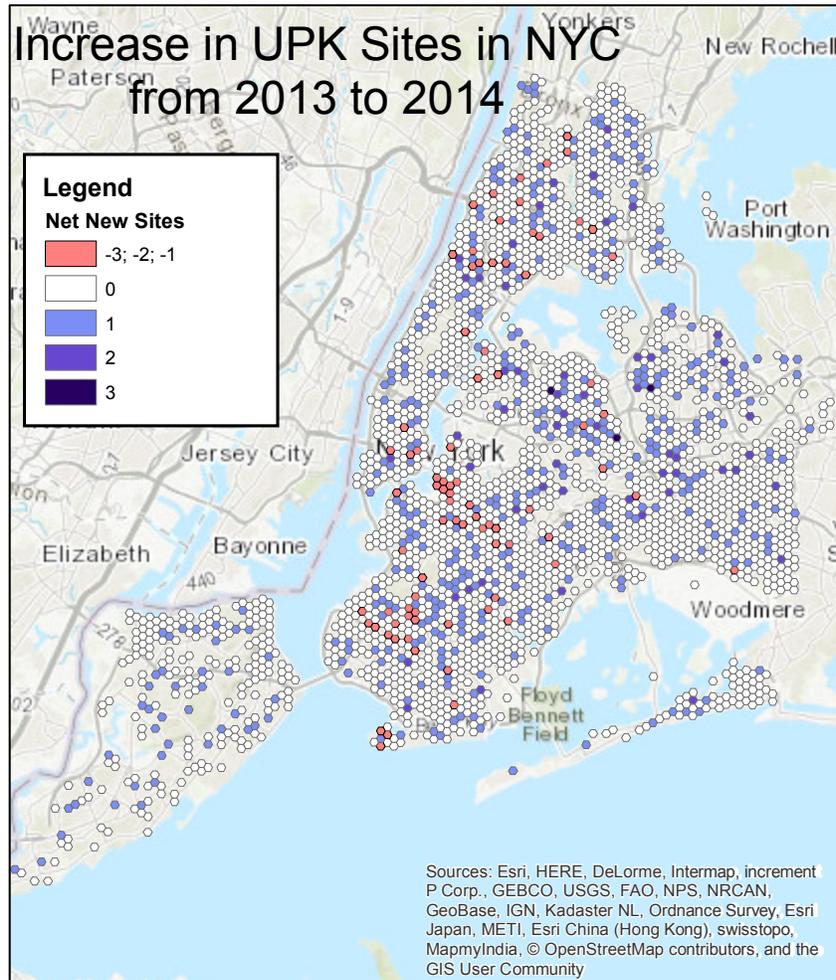
Notes: Using data acquired from the Pennsylvania Department of Human Services (DHS), graphs report the cumulative density function of the ratio of younger child tuition to age 4 tuition as reported by the day cares to DHS, their licensing agency. Data is as of November 2017 and covers Bucks, Chester, Delaware, Montgomery, and Philadelphia counties in Pennsylvania. The vertical red lines indicate the expected ratio based on mandated minimum staff to child ratios and assuming labor accounts for 70% of costs. For example, 1 year olds (shown in (b)) are required to have a 1:5 ratio compared to the 4 year old ratio of 1:10. Therefore, the tuition for 1 year olds should be at least 70% higher, so the tuition ratio should be at least 1.7. The required ratio for infants under 12 months is 1:4 and for 2 year olds it is 1:6.

Figure A.2: Tessellated Hexagons have the same relationship to all of their neighbors



Notes: Figure (a) demonstrates that in a grid of hexagons, all neighbors have the same relationship to the central hexagon (they are same-side neighbors). On the other hand, with a grid of squares, the central square has both same-side (white) neighbors and diagonal (light blue) neighbors, as seen in Figure (b). Because hexagons have the same relationship with all of their neighbors, pulling in neighbor effects is more natural and does not require adjustment for the type of neighbor.

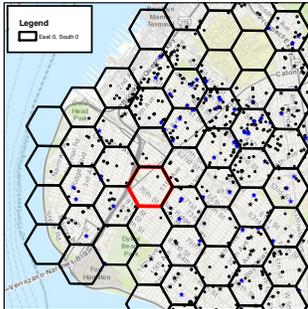
Figure A.3: Change in Number of Universal Pre-K Sites from 2013 to 2014 for Main Hexagon Grid



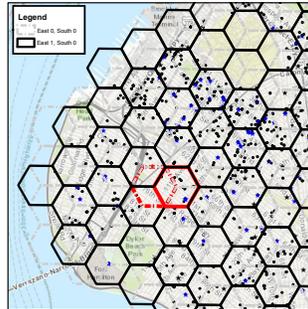
Notes: Map displays the number of full-day UPK sites in each hexagon in 2014-15 minus the number of full-day UPK sites in that hexagon in 2013-14. The change in the number of pre-K sites in the hexagon is the main “exposure” variable and is displayed on the hexagon grid used in the main specification. Hexagons are only included in the regression if they had a family day care home, day care center, or public elementary school at some point from 2008-2013. This restriction is made to make the “control” hexagons more similar to “treated” hexagons by excluding areas that are primarily parks, water, the airport, or other areas that day cares and UPKs cannot operate. Of the 2,658 hexagons shown, 449 had more UPK sites in 2014 than 2013 and 82 had fewer.

Figure A.4: Example of Hexagon Shifting for Robustness

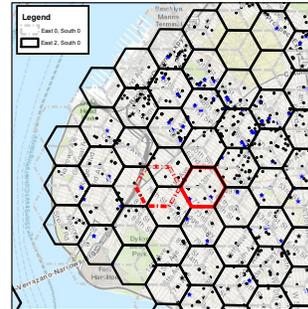
(a) East 0, South 0



(b) East 1, South 0



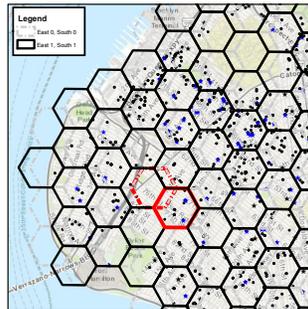
(c) East 2, South 0



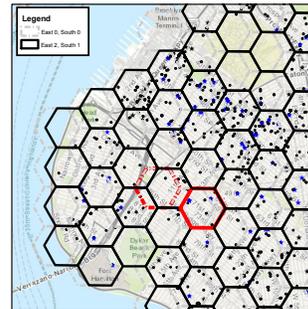
(d) East 0, South 1



(e) East 1, South 1

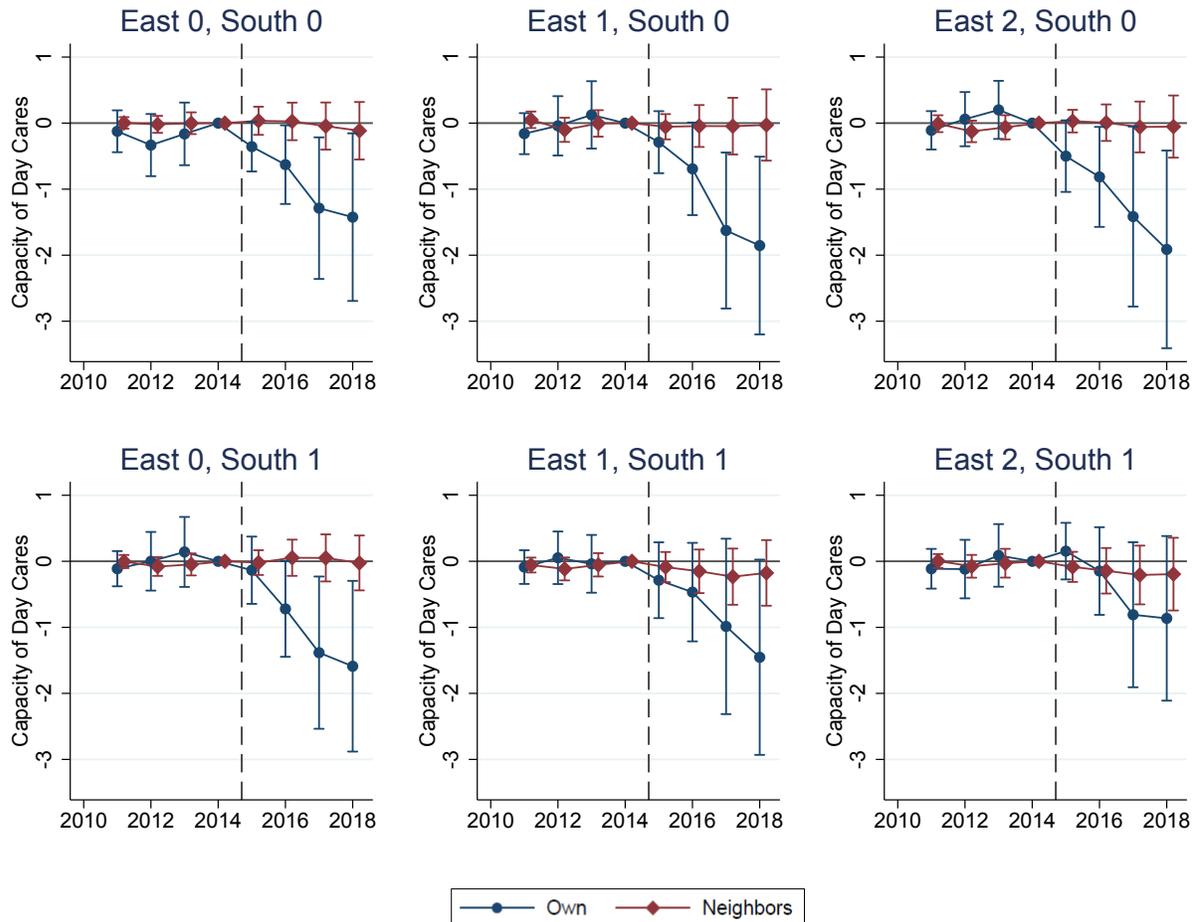


(f) East 2, South 1



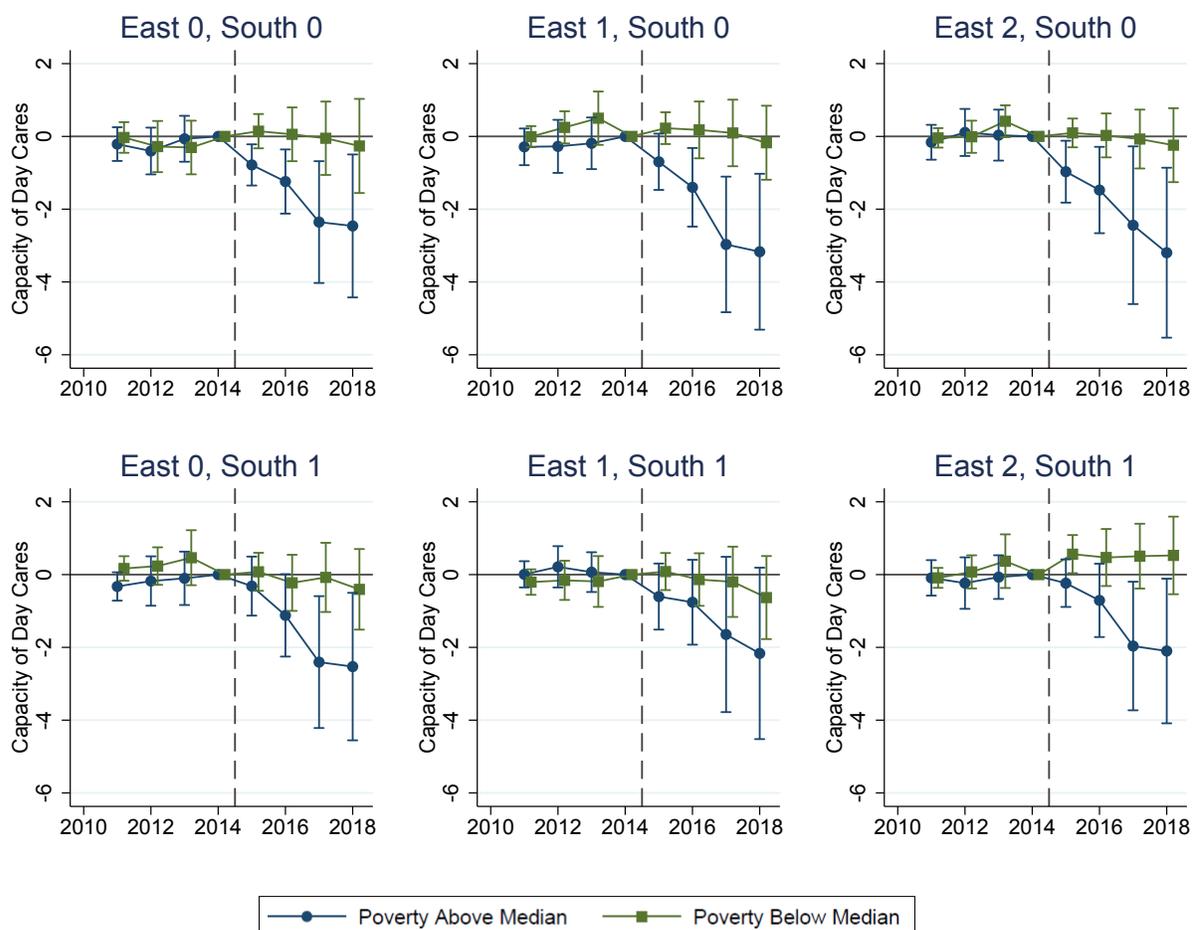
Notes: Figures demonstrate shifting an example hexagon grid. There are three “half hexagon” shifts east-west and two “half hexagon” shifts north-south.

Figure A.5: Effect of New Public Pre-K Site on Capacity of Day Care Centers Serving Younger Children: Robustness to Shifting the Grid



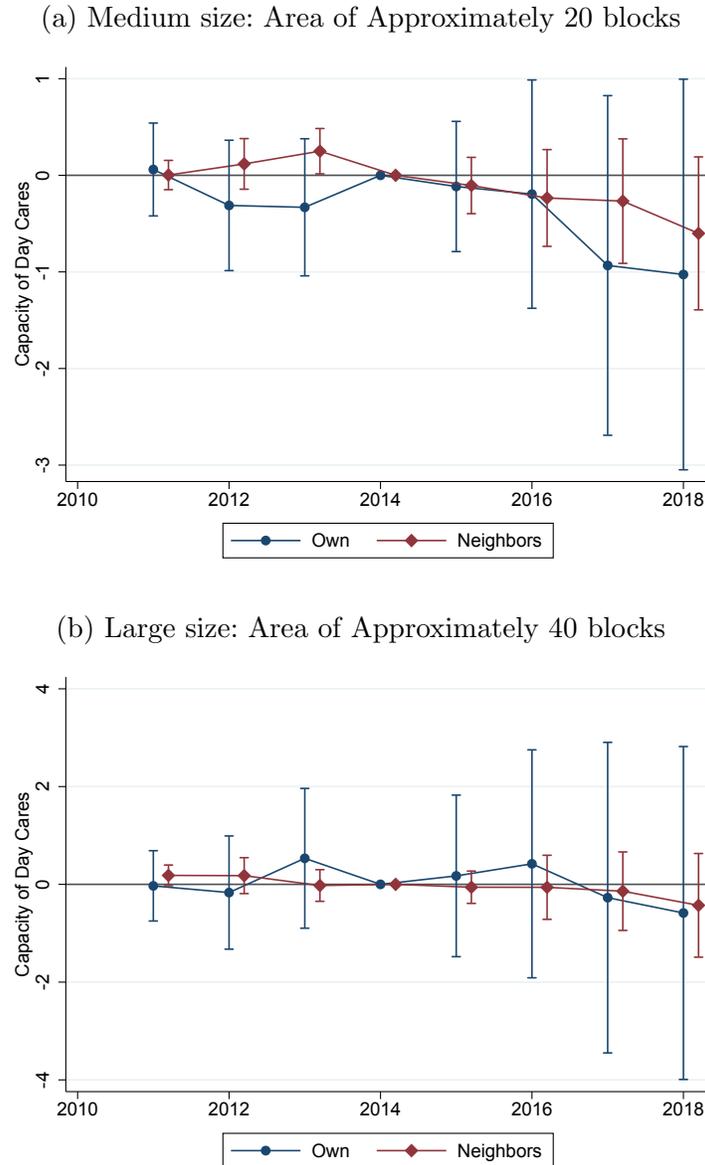
Notes: Figures plot coefficients on interactions between year indicators and the change in the number of Universal Pre-K sites between 2013-14 and 2014-15 in the hexagon (“Own”) and the hexagon’s immediate neighbors (“Neighbors”). Observations are at the hexagon-year level. Hexagons are approximately ten city blocks in area and form a grid that covers the city. Each figure represents a different shift of the hexagon grid. “East 0, South 0” is the original grid, while “East 1, South 0” shifts the whole grid one half hexagon to the east, and so on. There are three half hexagons going east-west and two going north-south, so these represent the complete set of half-hexagon shifts. The outcome variable is the total licensed capacity of day care centers serving children between the ages of 0 and 2. Includes hexagon fixed effects, time fixed effects, and hexagon-specific linear time trends. Bands show the 95% confidence interval with standard errors clustered at the hexagon level.

Figure A.6: Effect of New Public Pre-K Site on Capacity of Day Care Centers Serving Younger Children: Heterogeneity by Poverty Level Robustness to Shifting the Grid



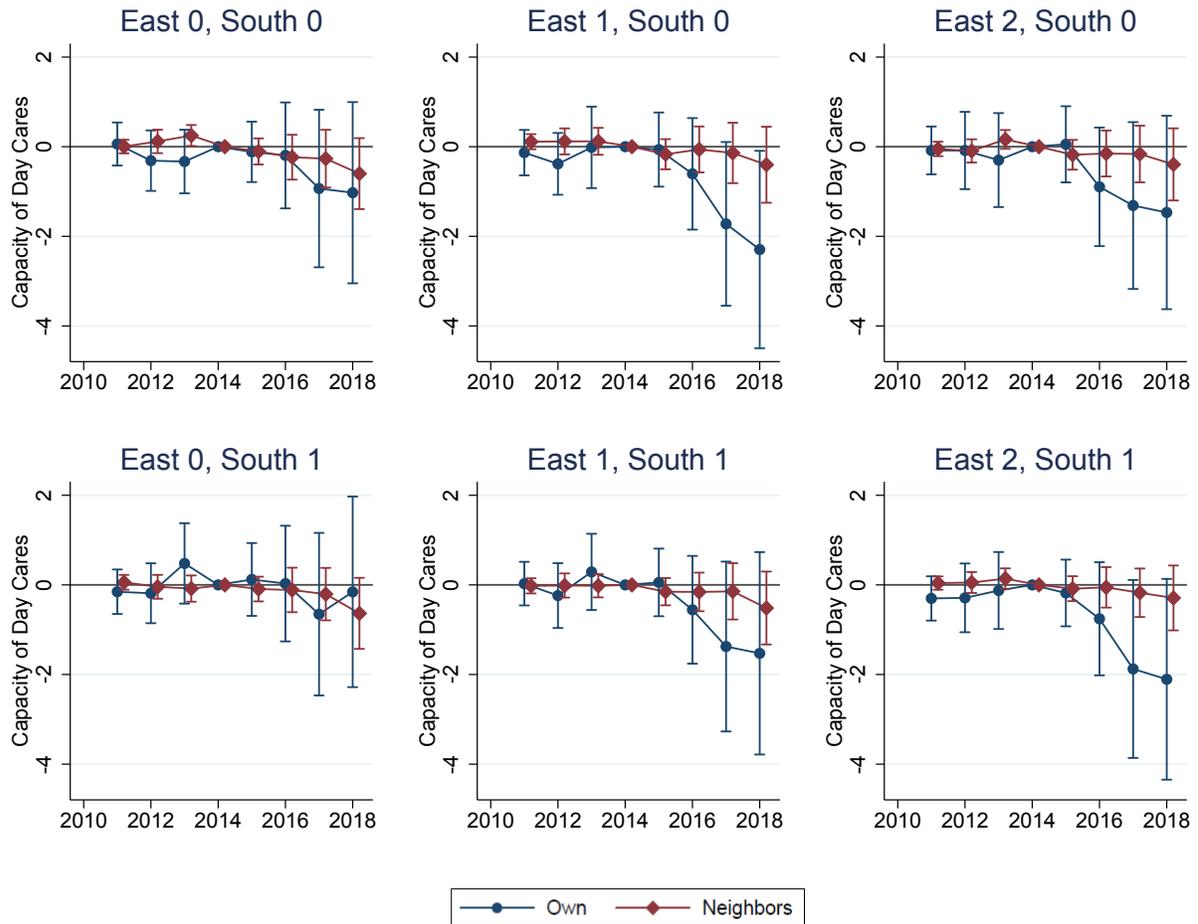
Notes: Figure compares how a new Universal Pre-K site affects the total capacity of day care centers serving children between the ages of 0 and 2 in areas with above-median levels of people living below the 200% poverty line to areas with below-median levels. In the median area, approximately 30% of people live below 200% of the federal poverty line. Figure plots coefficients of interaction of three variables: 1) year indicators, 2) indicators for above-median and below-median fraction of people living below 200% of the federal poverty line as of 2014, and 3) the change in the number of Universal Pre-K sites between 2013-14 and 2014-15 in the hexagon. Observations are at the hexagon-year level. Hexagons are approximately ten city blocks in area. Each figure represents a different shift of the hexagon grid. “East 0, South 0” is the original grid, while “East 1, South 0” shifts the whole grid one half hexagon to the east, and so on. There are three half hexagons going east-west and two going north-south. The outcome variable is the total licensed capacity of home day cares in the given hexagon. Includes hexagon fixed effects, time fixed effects, and hexagon-specific linear time trends. Bands show the 95% confidence interval with standard errors clustered at the hexagon level. Poverty level is determined by applying the American Community Survey 5-year estimates at the block group level to each block within the block group and calculating for each hexagon the average poverty level for the blocks in that hexagon weighting by Census 2010 population at the block level.

Figure A.7: Effect of New Public Pre-K Site on Capacity of Day Care Centers Serving Younger Children: Change in Estimate when Increasing the Size of the Hexagons



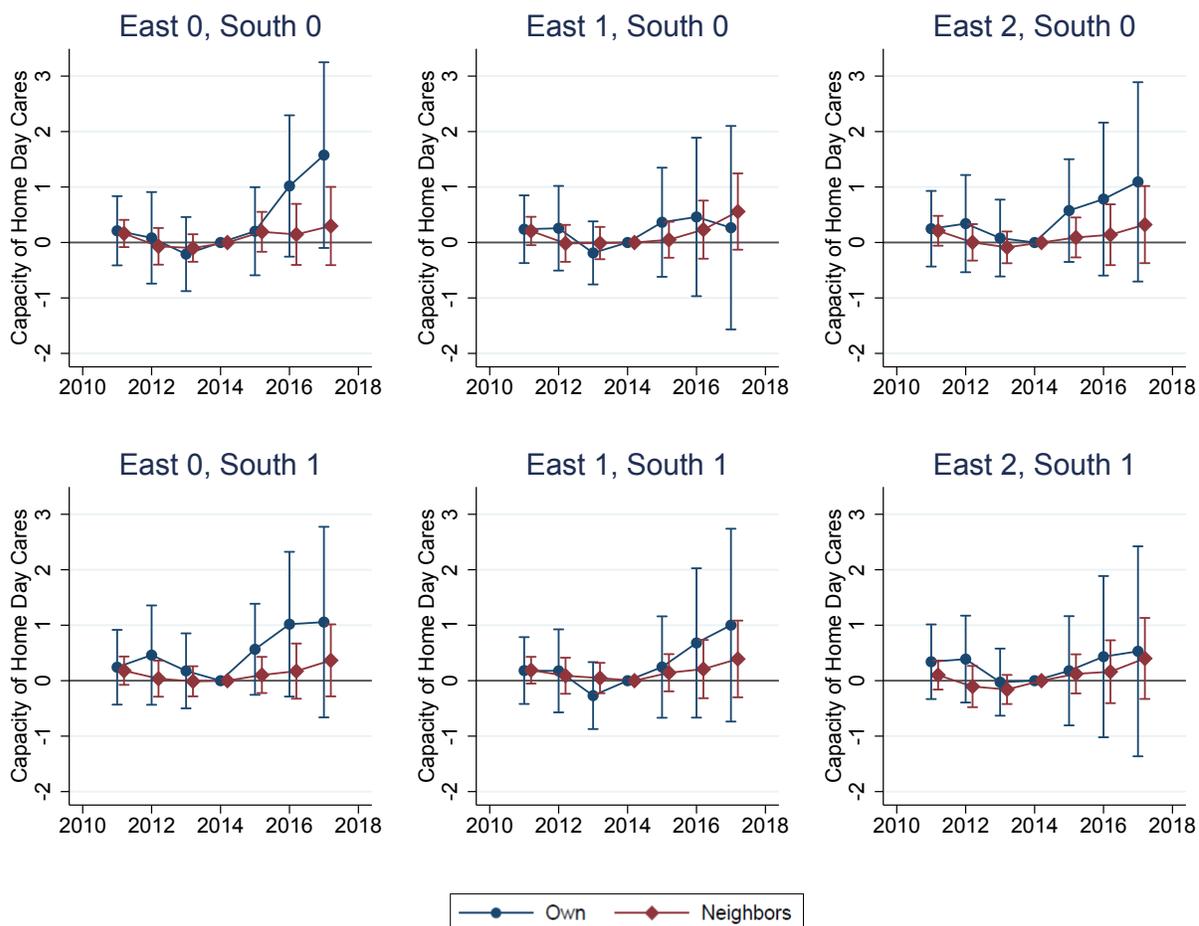
Notes: Plot of regression results from equation (2) with 95% confidence bands. “Own” coefficients estimate the effect of one new UPK site in a hexagon on the total capacity of day care centers for children between the ages of 0 and 2 in that hexagon. “Neighbors” coefficients estimate the effect of a new UPK in a neighboring hexagon on the capacity of day care centers serving children between the ages of 0 and 2 in the given hexagon. These figures show how estimates change for two larger grid sizes: areas of 20 blocks and 40 blocks.

Figure A.8: Effect of New Public Pre-K Site on Capacity of Day Care Centers Serving Younger Children: Medium Grid Robustness to Shifting



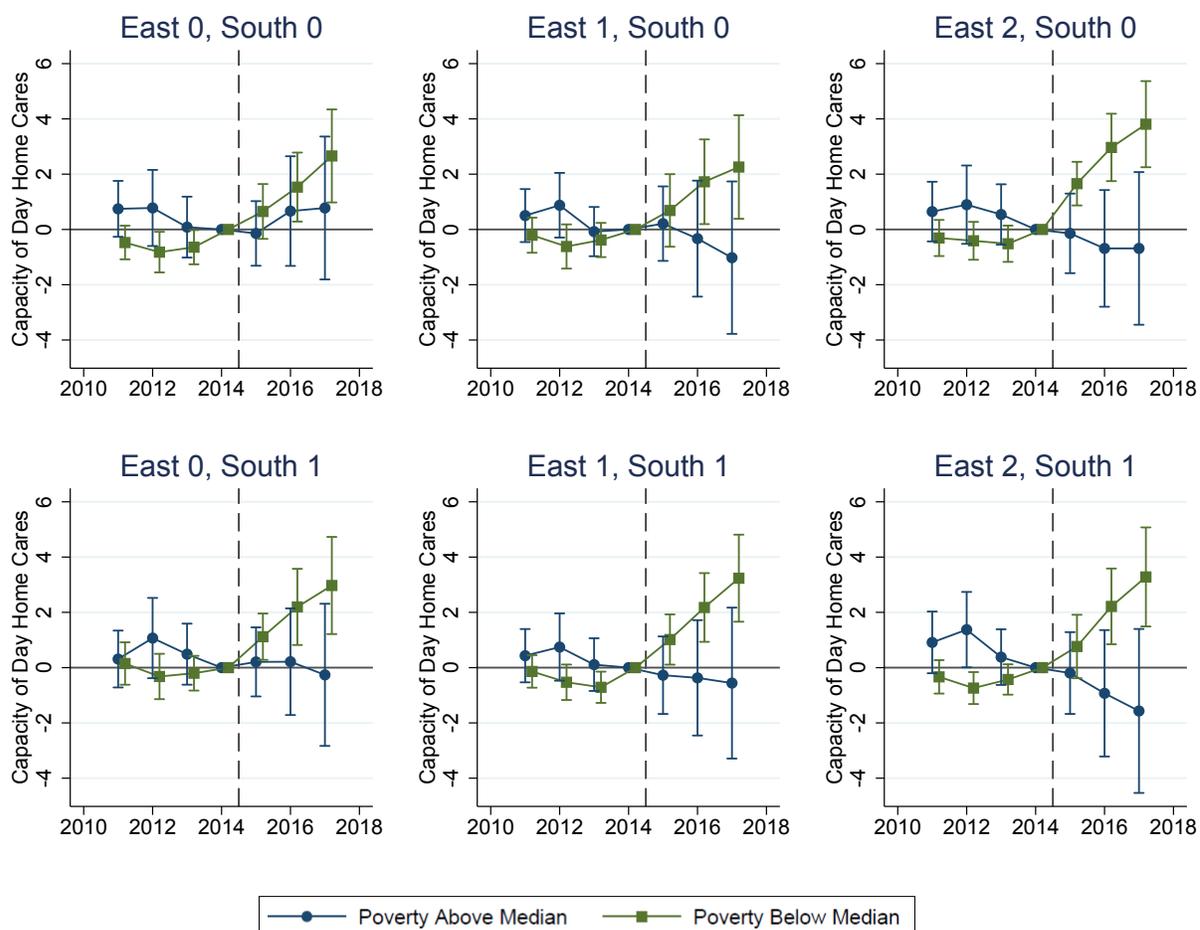
Notes: Figures plot coefficients on interactions between year indicators and the change in the number of Universal Pre-K sites between 2013-14 and 2014-15 in the hexagon (“Own”) and the hexagon’s immediate neighbors (“Neighbors”). Observations are at the hexagon-year level. Hexagons are approximately twenty city blocks in area and form a grid that covers the city. Each figure represents a different shift of the hexagon grid. “East 0, South 0” is the original grid, while “East 1, South 0” shifts the whole grid one half hexagon to the east, and so on. There are three half hexagons going east-west and two going north-south, so these represent the complete set of half-hexagon shifts. The outcome variable is the total licensed capacity of day care centers serving children between the ages of 0 and 2. Includes hexagon fixed effects, time fixed effects, and hexagon-specific linear time trends. Bands show the 95% confidence interval with standard errors clustered at the hexagon level.

Figure A.9: Effect of New Public Pre-K Site on Capacity of Home Day Cares: Robustness to Shifting the Grid



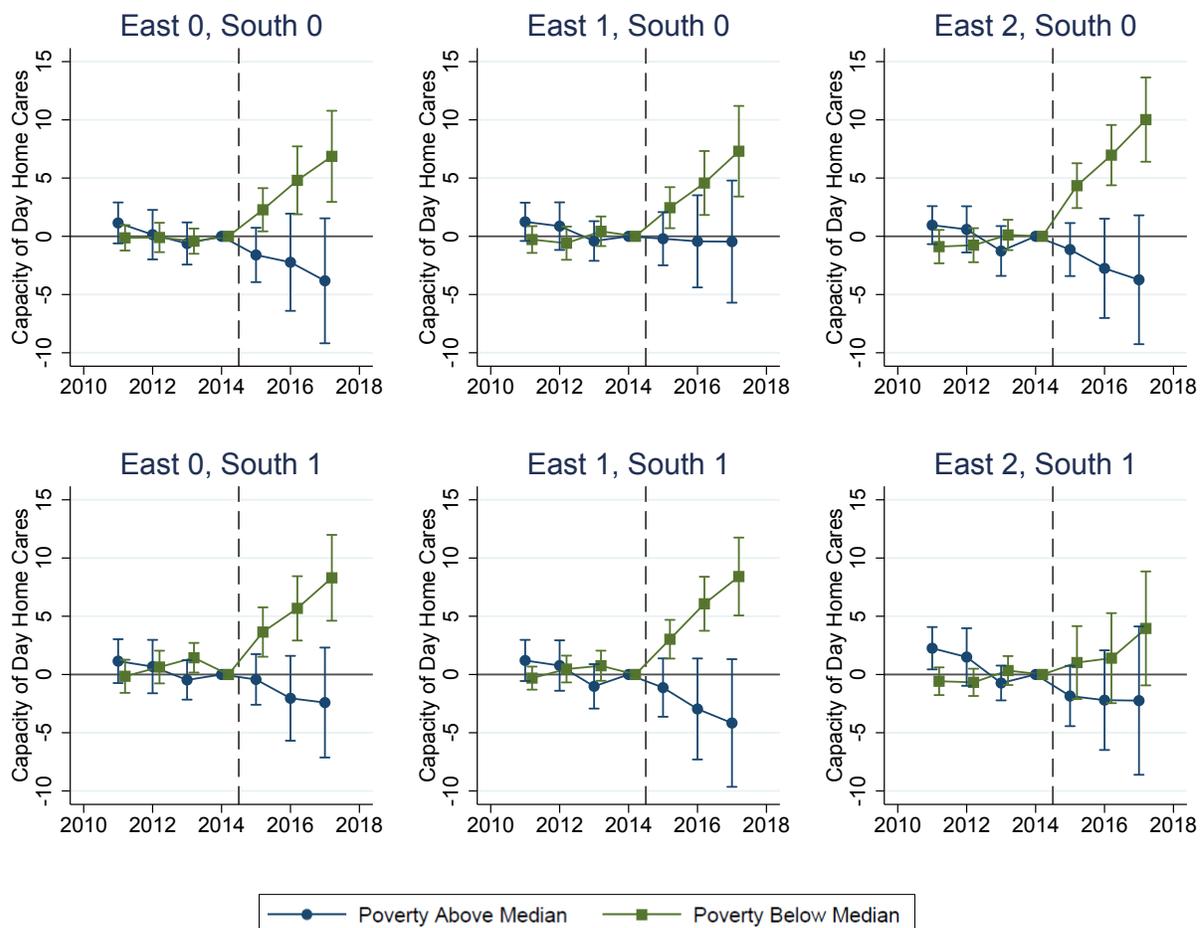
Notes: Figures plot coefficients on interactions between year indicators and the change in the number of Universal Pre-K sites between 2013-14 and 2014-15 in the hexagon (“Own”) and the hexagon’s immediate neighbors (“Neighbors”). Observations are at the hexagon-year level. Hexagons are approximately ten city blocks in area and form a grid that covers the city. Each figure represents a different shift of the hexagon grid. “East 0, South 0” is the original grid, while “East 1, South 0” shifts the whole grid one half hexagon to the east, and so on. There are three half hexagons going east-west and two going north-south, so these represent the complete set of half-hexagon shifts. The outcome variable is the total licensed capacity of home day cares. Note that the maximum capacity for children under age 2 is approximately one-third of total licensed capacity. Includes hexagon fixed effects, time fixed effects, and hexagon-specific linear time trends. Bands show the 95% confidence interval with standard errors clustered at the hexagon level.

Figure A.10: Effect of a New Pre-K Site on Capacity of Home Day Cares: Heterogeneity by Poverty Level Robustness to Shifting Grid



Notes: Figures compare how a new Universal Pre-K site affects the total capacity of home day cares in areas with above-median levels of people living below the 200% poverty line to areas with below-median levels. In the median area, approximately 30% of people live below 200% of the federal poverty line. Figure plots coefficients on interactions between three things: 1) year indicators, 2) indicators for above-median and below-median fraction of people living below 200% of the federal poverty line as of 2014, and 3) the change in the number of Universal Pre-K sites between 2013-14 and 2014-15 in the hexagon. Observations are at the hexagon-year level. Hexagons are approximately ten city blocks in area. Each figure represents a different shift of the hexagon grid. “East 0, South 0” is the original grid, while “East 1, South 0” shifts the whole grid one half hexagon to the east, and so on. There are three half hexagons going east-west and two going north-south. The outcome variable is the total licensed capacity of home day cares in the given hexagon. Includes hexagon fixed effects, time fixed effects, and hexagon-specific linear time trends. Bands show the 95% confidence interval with standard errors clustered at the hexagon level. Poverty level is determined by applying the American Community Survey 5-year estimates at the block group level to each block within the block group and calculating for each hexagon the average poverty level for the blocks in that hexagon weighting by Census 2010 population at the block level.

Figure A.11: Effect of a New Pre-K Site on Capacity of Home Day Care: Medium Grid Heterogeneity by Poverty Level Robustness to Shifting Grid



Notes: Figures compare how a new Universal Pre-K site affects the total capacity of home day cares in areas with above-median levels of people living below the 200% poverty line to areas with below-median levels. In the median area, approximately 30% of people live below 200% of the federal poverty line. Figure plots coefficients on interactions between three things: 1) year indicators, 2) indicators for above-median and below-median fraction of people living below 200% of the federal poverty line as of 2014, and 3) the change in the number of Universal Pre-K sites between 2013-14 and 2014-15 in the hexagon. Observations are at the hexagon-year level. Hexagons are approximately ten city blocks in area. Each figure represents a different shift of the hexagon grid. “East 0, South 0” is the original grid, while “East 1, South 0” shifts the whole grid one half hexagon to the east, and so on. There are three half hexagons going east-west and two going north-south. The outcome variable is the total licensed capacity of home day cares in the given hexagon. Includes hexagon fixed effects, time fixed effects, and hexagon-specific linear time trends. Bands show the 95% confidence interval with standard errors clustered at the hexagon level. Poverty level is determined by applying the American Community Survey 5-year estimates at the block group level to each block within the block group and calculating for each hexagon the average poverty level for the blocks in that hexagon weighting by Census 2010 population at the block level.

Table A.1: New York City Day Care Center Staffing Requirements

Age Group	Min. Staff to Child Ratio	Max. Group Size
Less than 12 months	1:4	8
12 to 24 months	1:5	10
2 years	1:6	12
3 years	1:10	15
4 years	1:12	20
5 years	1:15	25

Notes: Table provides the minimum staff ratio and maximum group sizes allowed by law for children in day care centers in New York City. A ratio of 1:4 means that there can be at most 4 children for each teacher in the classroom. Group size refers to how many children can be cared for together in one room.

Table A.2: Predictors of Having a Full-Day Universal Pre-K Site in 2014-15

	Number of Full-Day Pre-K Sites in 2014-15			
	(1)	(2)	(3)	(4)
Num. 2013 Pre-K Sites	0.865*** (0.0137)	0.862*** (0.0146)	0.661*** (0.0168)	0.657*** (0.0170)
Log Population		0.0146 (0.0124)	-0.0161 (0.0116)	-0.0209* (0.0121)
Frac. HHs with Children		-0.474*** (0.0973)	-0.202** (0.0913)	-0.225** (0.0933)
% Non-White		0.130*** (0.0397)	0.174*** (0.0369)	0.165*** (0.0374)
% Owner-occupied		0.0418 (0.0492)	0.0797* (0.0456)	0.0903* (0.0463)
% Land		0.254 (0.174)	0.137 (0.161)	0.136 (0.161)
Zero Population		-0.00469 (0.166)	-0.207 (0.154)	-0.235 (0.155)
Num. Elem. Schools			0.263*** (0.0190)	0.261*** (0.0191)
Num. 2014 DCC 2-5s			0.188*** (0.0105)	0.188*** (0.0117)
Num. 2014 DCC 0-2s				0.00307 (0.0257)
Num. 2014 HDCs				0.00358 (0.00242)
Dep. Var. Mean	0.512	0.512	0.512	0.512
N	2658	2658	2658	2658

Note: Observations are at the hexagon level. These are the hexagons used in the main specification and each is approximately ten city blocks in area. Demographic characteristics are from block-level data from the 2010 Census. Blocks are mapped to hexagons based on which hexagon their centroid falls in. “Zero population” denotes hexagons with a population of zero; this dummy variable is included so that no hexagons are dropped from the specification due to not having a denominator for fraction of households with children, fraction non-white, or percent owner-occupied. The “% Land” variable is the fraction of the hexagon that is land (as opposed to water). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: Number of Facilities in Non-Expansion v. Expansion Hexagons

Variable	Non-Expansion	Expansion	Diff
Num. 2013 Pre-K Sites	0.36 (0.02)	0.34 (0.03)	-0.026 (0.037)
Num. 2014 Pre-K Sites	0.32 (0.01)	1.47 (0.03)	1.149*** (0.035)
Num. 2014 Pre-K Seats	13.96 (0.67)	56.85 (1.99)	42.891*** (1.734)
Δ Sites per Hexagon	-0.04 (0.01)	1.13 (0.02)	1.176*** (0.014)
Neighbors' Δ Sites	0.68 (0.03)	0.82 (0.06)	0.137** (0.065)
Num. DCC 2-5s as of 1/1/14	0.57 (0.02)	1.31 (0.05)	0.741*** (0.050)
DCC 2-5 Capacity as of 1/1/14	36.88 (1.52)	92.45 (4.45)	55.566*** (3.914)
Num. DCC 0-2s as of 1/1/14	0.12 (0.01)	0.27 (0.02)	0.156*** (0.021)
DCC 0-2 Capacity as of 1/1/14	3.03 (0.25)	6.33 (0.66)	3.302*** (0.624)
Num. HDCs as of 1/1/14	2.99 (0.10)	2.81 (0.22)	-0.186 (0.236)
HDC Total Capacity as of 1/1/14	29.54 (0.92)	27.75 (2.14)	-1.785 (2.252)
HDC Inf/Todd Capacity as of 1/1/14	9.98 (0.31)	9.38 (0.73)	-0.604 (0.764)
N	2209	449	2658

Significance levels: * < 10% ** < 5% *** < 1%

Standard errors in parentheses.

“Expansion” hexagons are ones that had more full-day UPK sites in 2014 than in 2013, and “Non-expansion” hexagons are ones that did not. Statistics are for the hexagon grid in the main specification, which are about 10 city blocks in area.

Table A.4: Robustness: Effect of New Public Pre-K Site on Stock of Day Care Centers Serving Younger Children

	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta Sites * (Year = 2017)$	-1.144* (0.588)	-1.753*** (0.655)	-1.599** (0.703)	-1.517** (0.626)	-0.997 (0.663)	-0.882 (0.630)
$\Delta Sites * (Year = 2018)$	-1.276* (0.693)	-2.009*** (0.737)	-2.128*** (0.769)	-1.748** (0.690)	-1.468** (0.730)	-0.956 (0.711)
$\Delta NeighborSites * (Year = 2017)$	-0.0388 (0.159)	-0.00736 (0.213)	0.00793 (0.190)	0.0991 (0.172)	-0.189 (0.216)	-0.171 (0.203)
$\Delta NeighborSites * (Year = 2018)$	-0.109 (0.195)	0.0157 (0.267)	0.0198 (0.229)	0.0274 (0.199)	-0.133 (0.252)	-0.154 (0.250)
Dep. Var. Mean	3.408	3.385	3.370	3.413	3.411	3.407
Num. South Shifts	0	0	0	1	1	1
Num. East Shifts	0	1	2	0	1	2
N	23922	24120	24156	23967	23994	23922

Notes: Dependent variable is the capacity in or number of day care centers serving children between the ages of 0 and 2 in the hexagon. Hexagons are approximately ten blocks in area and cover New York City. Each model is estimated on a different shift of the original grid. Each shift is half of a hexagon, so that three shifts east brings the grid back to its original position, as does two shifts south. Includes year and hexagon fixed effects and a hexagon-specific linear time trend. Standard errors are clustered at the hexagon level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B Model of Child Care Pricing and Supply

B-1 Model Introduction

The purpose of this model is to provide intuition for why day cares may price such that they make the greatest profit from the oldest children and to use that model to provide a comparative static for the introduction of UPK.

The model assumes that child care providers have some local market power. The market power assumption can be micro-founded with spatial differentiation and parents' disutility of distance travelled to child care. For ease of exposition, I will assume a market structure of monopolistic competition.

Assume that child care facilities provide two products: care for younger children (infants) and care for older children (pre-kindergarteners). The demand for these two products is interrelated because the price of care for younger children at a location may affect the demand for care for older children at that same location, and vice versa. For example, a "low" price for infant care may attract more infants to the day care. If parents face switching costs when changing care arrangements, then they are likely to keep their child at the same day care as they age, so a lower price for infant care in one year increases the demand for toddler care in the next year. Also, parents of more than one child may prefer that they be cared for at the same location (for example, they may pay a utility cost for more than one drop off), so the price for the younger sibling could influence their demand for care for the older sibling at that location and vice versa.

Therefore, the demand for infant care depends on the price of infant care *as well as* the price of pre-K care. Parent demand will also depend on the quality of the facility. Similarly, demand for pre-K depends on the price of infant care, the price of pre-K care, and quality. In addition, parents face a switching cost if they choose a different child care arrangement for their pre-K-aged child than they did in the previous period when that child was an infant. This switching cost is incorporated into the model by a parent's demand for pre-K care at a given child care facility depending on whether the child was at that child care facility when they were an infant in the previous period.

Child care providers face fixed costs and marginal costs. I will assume constant marginal costs for simplicity. The marginal cost for each infant is much larger than the marginal cost for each pre-kindergartener, primarily due to the increased staffing requirements. Providers also choose a level of quality to provide, and their fixed costs and marginal costs are increasing in quality. I assume that there is one level of quality for the center and that the infant care and pre-K care cannot vary in their level of relative quality, though a given quality change could affect the marginal costs of infant and pre-K care differently.

B-2 The Firm's Problem

Algebraically, the provider's problem is:

$$\max_{p_I, p_K} (p_I - \alpha(Q)) * D_I(p_I, p_K, Q + (p_K - \gamma(Q)) * D_K(p_I, p_K, D_I(p_I^{-1}, p_K^{-1}, Q), Q) - \xi(Q) \quad (\text{B-1})$$

where p_I is the price of infant care, p_K is the cost of pre-K care, Q is the quality of the center, and D_I and D_K are demand for infant care and pre-K care, respectively. Marginal costs for infants ($\alpha(Q)$) and pre-K ($\gamma(Q)$) are increasing in quality Q , as are fixed costs ($\xi(Q)$). Note that sibling effects enter this equation through D_K 's dependence on p_I and D_I 's dependence on p_I . Switching costs enter the problem through D_K 's dependence on the previous period's number of infants

$D_I(p_I^{-1}, p_K^{-1}, Q)$.

Now I will make a few simplifying assumptions. First, I will assume that real prices are the same in each period so that $p_I^{-1} = p_I$ and $p_K^{-1} = p_K$. This assumption can be microfounded through relatively sticky day care prices. With this assumption, the equilibrium number of infants in the current period and the previous period is the same: $D_I(p_I, p_K, Q)$.

Next, in order to gain some intuition for why day cares cross-subsidize infants with older children, it is simplest to focus only on switching costs. Therefore, consider the case where sibling effects are not relevant. Then $\frac{\partial K}{\partial p_I} = 0$ because p_I only influences the number of pre-K children through its effect on infant enrollment in the previous period and not through attracting younger siblings. To further simplify, assume that parents of infants are not forward-looking and therefore do not take p_K into account when selecting a day care for their infant. Then $\frac{\partial I}{\partial p_K} = 0$.

Now the provider's problem can be written:

$$\max_{p_I, p_K} (p_I - \alpha(Q)) * D_I(p_I, Q + (p_K - \gamma(Q)) * D_K(p_K, D_I(p_I, Q), Q)) - \xi(Q) \quad (\text{B-2})$$

After taking first order conditions and rearranging, the solution to the firm's problem is:

$$p_K = \gamma(Q) - \frac{D_K}{\frac{\partial D_K}{\partial p_K}} \quad (\text{B-3})$$

$$p_I = \alpha(Q) - \frac{D_I}{\frac{\partial D_I}{\partial p_I}} + \frac{\partial D_K}{\partial D_I} * \frac{D_K}{\frac{\partial D_K}{\partial p_K}} \quad (\text{B-4})$$

$$\frac{\partial \alpha}{\partial Q} D_I + \frac{\partial \gamma}{\partial Q} D_K + \frac{\partial \xi}{\partial Q} = - \frac{D_I}{\frac{\partial D_I}{\partial p_I}} \frac{\partial D_I}{\partial Q} - \frac{D_K}{\frac{\partial D_K}{\partial p_K}} \frac{\partial D_K}{\partial Q} \quad (\text{B-5})$$

The solution for the price for pre-K care is the standard solution in a monopolistic competition model. Without switching costs ($\frac{\partial D_K}{\partial D_I} = 0$), the infant price would be the standard monopoly solution as well. However, with switching costs, $\frac{\partial D_K}{\partial D_I} > 0$ since children who come to the day care as infants are more likely to stay for pre-K. Because $\frac{\partial D_K}{\partial p_K} < 0$, assuming that day cares make positive profit on pre-K children means the whole term added to the monopoly solution is negative, so switching costs lower the price of infant care. Intuitively, the firm takes into account that the marginal revenue from attracting an infant includes not only the marginal revenue from their infant year but also the revenue they would get from their pre-K year times the increased chance they will stay for pre-K given that they came for infant care. Therefore, in the presence of switching costs, day cares charge lower prices to infants because they take into account the additional profits they will make if the child stays for pre-K. In other words, infants are used as loss leaders.

The first order condition for quality is set up so that the left-hand side is the marginal cost of an increase in quality and the right-hand side is the marginal benefit of an increase in quality. The cost is the additional marginal cost times the number of children plus the additional fixed costs, and the benefit is the profit gained from attracting more children due to parent's preferences for higher quality.

B-3 Comparative Static: Universal Pre-K Introduction

Now I will consider how introducing free, public pre-K changes the firm's optimal solution. First, the introduction of UPK reduces demand for private pre-K care. It causes a decrease in the retention of infants as future pre-kindergarteners since public pre-K introduces a new substitute for private

pre-K. In the language of the model, a reduction in the retention of infants is a decrease in $\frac{\partial D_K}{\partial D_I}$ and a reduction in the number of pre-kindergarteners is a decrease in D_K .

A reduction in D_K and/or an increase in $\frac{\partial K}{\partial p_K}$ reduces the profit from pre-K care. Furthermore, the reduction in $\frac{\partial D_K}{\partial D_I}$ means that a child is less likely to stay for pre-K once they come as an infant. So at a given day care facility quality level, p_I increases since attracting an infant is less likely to lead them to stay for the more profitable pre-K age, and even if they do stay, the profit from doing so is lower. The link between starting infant care at a day care and staying for pre-K is weaker, reducing the incentive to charge a “lower” price to attract infants. Note that the amount of change in p_I at a given quality Q increases with the degree of crowd out (the amount that $\frac{\partial D_K}{\partial D_I}$ decreases). The more crowd out, the larger the decrease in the probability of staying for pre-K, and the larger the price increase for infants at a given quality level.

Figure B.1 presents the UPK introduction comparative static in a simple supply and demand framework. Universal Pre-K shifts the demand curve for private pre-K down, lowering the quantity of pre-K care at equilibrium and lowering its price. Due to the lower pre-K price and quantity, day cares shift their supply curves for infant care up, requiring a higher price for a given quantity of infant care supplied. The supply curve shift lowers the quantity of infant care supplied and increases its price. The degree to which quantity and price change depends on how elastic demand is for infant care. If it is relatively inelastic, there will be more adjustment on the price margin and less on the quantity margin relative to if it is relatively elastic. For parents who can only afford to pay for care with child care subsidies, demand for child care centers would be perfectly elastic and all adjustment would be on the quantity margin.

B-4 Quality Enters as a Change in the Marginal Cost

The previous subsection considers changes predicted changes in the price of infant care after the introduction of public pre-K *holding the quality of the day care fixed*. However, day care centers can also choose their quality. Lowering their quality reduces the marginal cost of providing care ($\alpha(Q)$ and $\gamma(Q)$ in the model) as well as the fixed cost ($\xi(Q)$). Therefore, another way child care facilities could react to crowd out of pre-K children would be to reduce their marginal costs by reducing quality of care. For example, they could hire teachers with less experience or education, reduce investments in professional development, lower the quality of provided food, or reduce infrastructure investments or maintenance. These would enter the model as reducing the marginal cost of infants, $\alpha(Q)$, the marginal cost of pre-kindergarteners, $\gamma(Q)$, and/or the fixed costs, $\xi(Q)$.

Turning back to the model, a change in the retention of infants is not predicted to change the equilibrium quality for a given price since $\frac{\partial D_K}{\partial D_I}$ doesn't appear in the equilibrium condition. However, a decrease in the number of pre-kindergarteners (D_K) does affect the equilibrium condition. To determine whether a child care facility lowers quality in response to a decrease in pre-kindergarteners (D_K), we can explore what happens to the equilibrium quality condition when D_K falls. On the left-hand side, a decrease in D_K reduces the cost of an increase in quality since there are now fewer students, but on the right-hand side, the profit from attracting additional children is lower due to the lower price charged for pre-K care. If the decrease in costs ($\frac{\partial \gamma}{\partial Q} \cdot \Delta D_K$) is smaller than the decrease in benefit ($\frac{\frac{D_K}{Q}}{\frac{\partial D_K}{\partial p_K}} \cdot \Delta D_K$), then it will be optimal to lower quality. Rearranging yields the following condition for decreasing quality:

$$\frac{\frac{\partial D_K}{\partial Q}}{\frac{\partial \gamma}{\partial Q}} < -\frac{\partial D_K}{\partial p_K} \quad (\text{B-6})$$

Intuitively, centers will decrease quality if the number of children they would lose from the quality reduction is smaller than the number of children they would lose from increasing prices by an equivalent amount. Therefore, the introduction of public pre-K could also decrease the quality of competing child care facilities.

B-5 Opening/Closing Decision

The model above derives quality of care and prices for infant and pre-K child care given that the firm is open. In each period, the firm also decides whether or not to stay open. This section derives a very simple threshold-crossing model of a firm's decision to open or shut down.

Assume there are some fixed costs to running a day care, which can vary for new firms ($\xi(Q^*)_N$) and incumbent firms ($\xi(Q^*)_I$), where Q^* is the optimal quality level given that the firm chooses to operate. Assume that there is a fixed capacity C per open room of infants. The value of C is generally given by regulations governing the maximum group size for each age group and is also determined by the licensing agency based on square footage of the classroom. Anecdotally, infant and young toddler care often has wait lists, so assume that if the firm is open, infant care is priced such that the infant room will be full ($D_I = C$).

Given these assumptions and the profit maximization problem from equation B-2, profit can be rewritten as a function only of the number of pre-K kids ($\pi(D_K)$). Then the firm is only open in a given period if:

$$\xi(Q^*)_i < \pi(D_K), i \in N, I \tag{B-7}$$

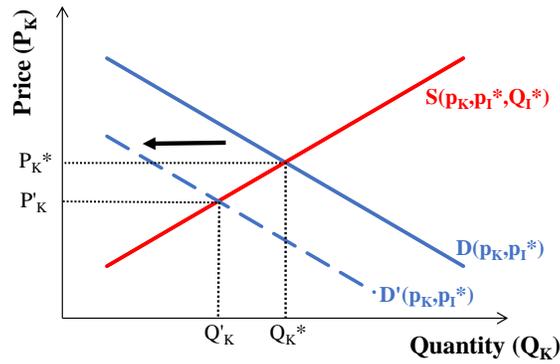
This generates a threshold-crossing rule where the firm will stay open if and only if it can attract a minimum number of students:

$$\underline{D}_{K_i}, i \in N, I \text{ students.} \tag{B-8}$$

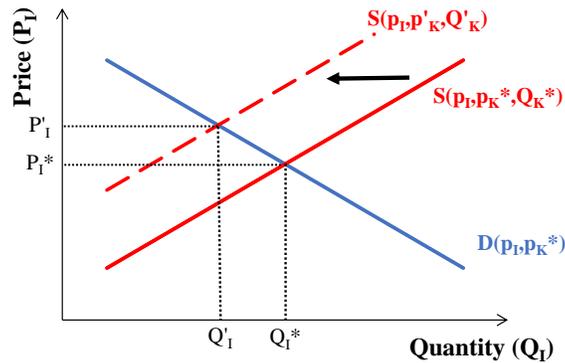
Therefore, if crowd out due to UPK makes it less likely that a firm will have enough pre-K students for each classroom of infants to cover its fixed costs, then currently-open day cares will be more likely to close and new day cares will be less likely to open. This argument also goes through for any home day cares that care for infants - they will be more likely to close and less likely to open in the presence of fixed costs if they cannot attract enough children ages 2 and older to make running the day care worthwhile.

Figure B.1: Effect of Universal Pre-K on Private Child Care Market

(a) Demand for Private pre-K goes down



(b) Supply of Infant Seats goes down due to Lower Private pre-K at Equilibrium



Notes: Suppose private day cares can provide two goods: infant care and pre-K care. The demand for these goods is interrelated, and empirically, day cares cross-subsidize infant care with pre-K care. When Universal Pre-K opens, parents reduce their demand for private pre-K care as in subfigure (a), lowering the equilibrium quantity of private pre-K. The lower equilibrium quantity of private pre-K care shifts the supply curve in the private infant care market to the left, reducing the equilibrium quantity of infant care and increasing its price.

C Pre-existing Reduced Form Identification Methods

This appendix addresses the two current reduced-form approaches that can be used when treatment intensity is based on distance to a point. It describes the methods and how they are implemented and why they are not used in this case.

C-1 Ring Method is biased when rings overlap

There are two ways to implement the ring method, which I will call the “potential sites” design and the “near/far design”. With the former, the first step is to locate all potential sites and identify which became UPKs and which did not. Next, one would compare outcomes in a circle surrounding sites that do become UPK sites to ones that do not. In the “near/far design”, the circles are instead drawn around all new UPK sites. The changes in the number and capacity of day cares in a smaller circle surrounding any new UPK (say, 0.25 miles) is compared to changes in a larger ring around it, say between 0.25 miles and 0.5 miles. Both approaches are common in the literature. See [Dinerstein and Smith \(2014\)](#) or [Davis \(2011\)](#) for “potential sites” and [Currie and Walker \(2011\)](#) for the “near/far” design.

The more feasible design in this case would most likely be the “near/far” design. Implementing the potential sites design would involve identifying all public schools, private schools, and day care centers since these three groups encompass the vast majority of UPK sites. This list still does not give all potential sites, however, because some sites were in facilities that were not operating in the pre-period. The number of potential sites is also very large as there are 1,848 day care centers alone, so even if the rings were very small, there would still be a lot of overlap. Since the near/far design would involve about 400 observed new sites, it would be more feasible.

However, the ring method is likely to be biased when treatment is spatially dense enough that markets (and therefore rings) overlap. And because observations would not be mutually-exclusive since the same area can be part of more than one ring, the standard errors may also be misestimated. With more than 400 sites in an area of about 300 square miles, in this application, the rings would overlap substantially even if they were small. In the “near/far” design, if two sites are far enough apart that their treated areas do not overlap but are close enough that their treated area does overlap with the other’s “control” ring, the estimate would be biased down due to the contamination of the “control” area (see part a of [Figure C.1](#) for a visual). On the other hand, when two sites are close enough together that there is a large overlap in the areas that are “treated” by them, then estimates would be biased up if there is an intensive as well as extensive effect (that is, if there is a larger effect if there are two UPKs in an area as opposed to just one). The observed area is affected by two UPKs, but the regression treats each area as if it is only affected by one UPK, therefore overestimating the effect by attributing all of it to one UPK instead of two (see part (b) of [Figure C.1](#) for a visual). Similar arguments go through for the alternative “treated/untreated” design. Therefore, when treatment is spatially dense, estimates based on the ring method could be biased and in an unknown direction.

C-2 Testing Robustness to Administrative Boundaries is Challenging

Sometimes data restrictions or the identification strategy require the use administrative boundaries as the geographic units of observation. Sometimes the relevant market is defined as the unit of observation itself without allowing for overlapping markets ([Cellini, 2009](#)). For example, I could

allow the child care market in a given Census tract to be affected by all UPKs in that Census tract but not any that are outside that Census tract. Other times, the market is defined by a ring around the centroid of the administrative unit (Baum-Snow and Marion, 2009). For example, I could allow the child care market in a given Census tract to be affected by all UPKs in Census tracts whose centroids are within a certain radius of that tract's centroid.⁴⁸

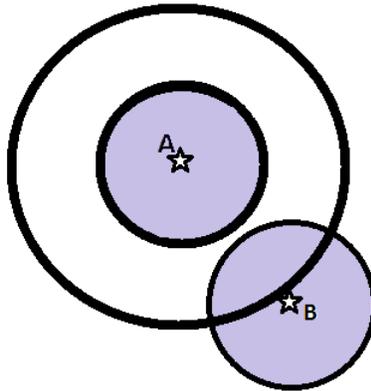
However, administrative boundaries are arbitrary, irregularly-shaped, and there is no natural way to test robustness to them, so they may not be the best option when they are not required by data or other constraints. Foote *et al.* (2017) demonstrate that uncertainty in the definition of commuting zones is empirically relevant for analysis of local labor markets. Therefore, testing robustness to boundaries can be important. Although using different boundaries could change results, this issue is often brushed under the rug. And for good reason since there is no natural way to shift the boundaries or make them larger or smaller, so there is often little that a researcher can do.

One reason that these boundaries may be preferred even when they are not required by the data is that, although they are arbitrary, they are not defined by the researcher, so they are less subject to manipulation. I will address this concern by demonstrating how results change when boundaries are shifted or enlarged in a systematic way.

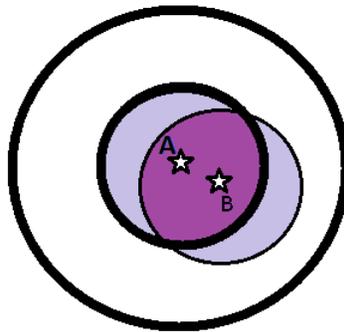
⁴⁸Alternatively, one could allow the child care market in a Census tract to be affected by all UPKs within a given radius of the tract's centroid regardless of which tract the UPK is located in. In practice, it is not usually implemented this way due to the data restrictions that required the analysis to be done at the level of the administrative unit in the first place.

Figure C.1: Potential Bias with Ring method

(a) “Treated” area overlaps “un-treated” area: biased down



(b) Two “treated” areas overlap: unknown bias



Notes: Assume the ring method model is correctly specified: the effect of the treatment dissipates with distance and the inner ring covers the area affected by the treatment. Then, if in (a), two sites are close enough together that the treated ring overlaps the other’s control ring but far enough apart that the treated areas do not overlap, then the estimate of the effect would be biased down in absolute value due to the contamination of the control ring with the other treated ring. If, on the other hand, they are close enough that their treated areas overlap, the effect could be biased up or down. Because all of the effect in the inner ring is attributed to one site in the regression even though in the dark purple shaded area, it is affected by two, it could bias the result away from zero. But if the inner rings overlap then there must also be the issue of the treated ring overlapping the control ring, which biases the effect toward zero. Which effect dominates depends on how close the sites are and how rapidly the effect dissipates with distance.