Racial Disparities in Pre-K Quality: Evidence from New York City's Universal Pre-K Program

Scott Latham
Princeton University

Sean P. Corcoran
Vanderbilt University

Carolyn Sattin-Bajaj
University of California, Santa Barbara

Jennifer L. Jennings
Princeton University

New York City's universal pre-kindergarten program, which increased full-day enrollment from 19,000 to almost 70,000 children, is ambitious in both scale and implementation speed. We provide new evidence on the distribution of pre-K quality in NYC by student race/ethnicity, and investigate the extent to which observed differences are associated with the spatial distribution of higher-quality providers. Relative to other jurisdictions, we find the average quality of public pre-K providers is high. However, we identify large disparities in the average quality of providers experienced by black and white students, which is partially explained by differential proximity to higher-quality providers. Taken together, current racial disparities in the quality of pre-K providers may limit the program's ability to reduce racial achievement gaps.

VERSION: July 2020

Racial Disparities in Pre-K Quality:
Evidence from New York City's Universal Pre-K Program

Scott Latham
Princeton School of Public and International Affairs
288 Wallace Hall, Princeton, NJ, 08544
slatham@princeton.edu
(732) 513-0964

Sean P. Corcoran
Peabody School of Education
Vanderbilt University
205C Payne Hall, Nashville, TN, 37203
sean.corcoran@vanderbilt.edu
(615) 322-8021

Carolyn Sattin-Bajaj
Gevirtz Graduate School of Education
University of California, Santa Barbara
ED 3117, Santa Barbara, CA, 93106
carolynsattin-bajaj@ucsb.edu

Jennifer L. Jennings (Corresponding author)
Princeton School of Public and International Affairs
159 Wallace Hall, Princeton, NJ, 08544
jlj@princeton.edu
(609) 258-4422

Racial Disparities in Pre-K Quality:

Evidence from New York City's Universal Pre-K Program

Abstract

New York City's universal pre-kindergarten program, which increased full-day enrollment from 19,000 to almost 70,000 children, is ambitious in both scale and implementation speed. We provide new evidence on the distribution of pre-K quality in NYC by student race/ethnicity, and investigate the extent to which observed differences are associated with the spatial distribution of higher-quality providers. Relative to other jurisdictions, we find the average quality of public pre-K providers is high. However, we identify large disparities in the average quality of providers experienced by black and white students, which is partially explained by differential proximity to higher-quality providers. Taken together, current racial disparities in the quality of pre-K providers may limit the program's ability to reduce racial achievement gaps.

In the last two decades, states and districts have expanded public pre-kindergarten (pre-K), often with the stated goal of reducing socioeconomic and racial achievement gaps (Friedman-Krauss et al., 2018). Studies documenting the short and long-term effects of targeted, highly-resourced preschool programs provide an evidence base for these polices (Campbell et al., 2012; Heckman et al, 2010), as do studies demonstrating long-term effects of larger programs like Head Start (Deming, 2009). However, public pre-K programs have thus far struggled to replicate earlier programs' effects on early educational outcomes (Yoshikawa et al., 2013; Lipsey, Farran, & Hofer, 2015; Phillips et al., 2017), leading scholars and policymakers to emphasize the importance of program quality.

Previous research has focused on "structural" and "process" conceptualizations of pre-K program quality. Structural quality refers to classroom features such as the physical space and its layout, as well as teacher qualifications. Existing research suggests that these aspects of the classroom may be necessary but not sufficient for improving early math and language skills (Early et al., 2006; Phillips et al., 2017). In contrast, measures of process quality evaluate student-teacher and student-student classroom interactions. Trained observers use established rubrics, such as the Early Childhood Education Rating Scale - Revised Edition (ECERS-R, henceforth ECERS; Harms, Clifford, & Cryer, 2004), and the Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008), to measure process quality.

Research suggests that higher scores on both the ECERS and CLASS scales are associated with student learning in preschool, including math and language skills (Mashburn et al., 2008). Higher scores on the CLASS, in particular, have been consistently related to better student outcomes (Araujo et al., 2016; Sabol et al., 2013). By contrast, several recent studies have reported

weak or no association between ECERS ratings and later child outcomes (Brunsek et al, 2017; Gordon et al., 2015; Hong et al., 2019).

One potential explanation for the inconsistent association between classroom quality and student outcomes is that the relationship is nonlinear: classroom quality appears to be more strongly associated with outcomes in the higher end of the quality distribution. For instance, one recent examination of the association between ECERS and student outcomes found that these associations were stronger above a certain baseline level of quality (Le, Schaack, & Setodji, 2015). Similarly, the association between CLASS scores and student outcomes is stronger among higher quality providers (Hatfield et al., 2016; Weiland et al, 2013). For this reason, disparities at the higher end of the quality distribution may be more consequential for student outcomes than disparities at the lower end of the distribution. It follows that understanding quality disparities between groups, even when quality is generally high, is important if a goal is to reduce socioeconomic and racial disparities in educational outcomes.

Racial disparities in the quality of early childhood experiences have been documented in other contexts. The NCEDL-SWEEP study, which included 11 states that served over 80% of the U.S. preschool population in 2001-02, found that black and Latino children attended programs of lower average quality than did white children (Early et al., 2010; Valentino, 2018). More recently, similar disparities were documented in California and Georgia (Karoly et al., 2008; Bassok & Galdo, 2016). Racial and income segregation, which is more pronounced for pre-K than other grade levels, is associated with higher levels of racial disparities in quality (Frankenberg, 2016; Valentino, 2018), suggesting that geographic variation in the supply of high-quality programs may play a role in driving these disparities. For instance, a recent analysis of 26 state-funded pre-K programs concluded that black and Latino children have disproportionately low access to high-

quality providers (Education Trust, 2019). However, to our knowledge, no studies to date have explicitly examined the spatial distribution of pre-K quality within individual markets.

In this paper, we examine the case of New York City's (NYC) full-day Universal Pre-K (UPK) program, which grew from 19,000 to almost 70,000 children over two years, making it the largest pre-K program in the country, with only one state (Georgia) serving more children in full-day pre-kindergarten (Potter, 2015). New York City provides a compelling context to examine disparities in a universal pre-K program due to its ambitious scope, racial and ethnic diversity, and high degree of residential segregation. NYC is not unique in this regard; other large urban districts are likely to encounter many of the same issues when adopting universal programs of their own. We examine the following research questions:

1. What is the distribution of UPK program quality in NYC? To what extent does this vary by students' race and ethnicity?

2. To what extent are racial/ethnic differences in program quality driven by spatial differences in local program supply?

NYC's UPK program has the potential to reduce racial and socioeconomic gaps in educational outcomes through differential increases in pre-K attendance for disadvantaged groups, differential access to high-quality programs, or differential benefits from programs of the same quality. However, NYC is the most segregated school district in the nation (Kucsera, 2014), and Potter (2016) found that UPK classrooms were *more* segregated than NYC's kindergarten classrooms. To the extent that advantaged groups attend higher quality programs, UPK also has the potential to exacerbate, or leave unchanged, existing disparities.

Expansion of Universal Pre-K in New York City

New York City's pre-K program was initiated in 1998 with the goal of providing all 4-year-olds with free public preschool, but lacked funding to achieve this goal. During his 2013 mayoral campaign, Bill de Blasio promised to expand UPK to provide free high-quality, full-day pre-K to all 4-year-olds in the city. After his election, pre-K expansion unfolded over two years; by fall 2015, over 70,000 4-year-olds (about 65% of 4-year-olds in NYC) were enrolled in free full-day pre-K across the city, a dramatic increase from the 19,000 in full-day programs and 39,000 in part-day programs served previously (Potter, 2016).

Given space limitations at UPK providers run by the NYC Department of Education, the DOE expanded supply through contracts with community-based organizations (CBOs). Although the DOE previously contracted with CBOs to offer free preschool, most slots at these sites were half-day. Under the expanded UPK program, the city required CBOs to offer full-day slots, and to meet the same program standards as DOE-run programs. To signal that CBOs were now an integral part of NYC's pre-K system, these programs were renamed "New York City Early Education Centers" (NYCEECs, pronounced "nigh-seeks") (Potter, 2015). Importantly, NYCEECs remain independently operated, and may be administered by non-profits, nursery schools, day care centers, or faith-based organizations.

The provision of high-quality pre-K was an explicit goal of UPK expansion, and the DOE selected only 60% of applicants to operate as NYCEECs (Goldstein, 2016). All providers are required to meet a comprehensive set of program standards that emphasize rigorous instruction, a supportive classroom environment, teacher collaboration, effective school leadership, and strong family-community ties (NYC Department of Education, 2016). These include specific provisions geared toward promoting both structural and process quality. In terms of structural quality, full-

day programs must offer a minimum of 6 hours, 20 minutes of instruction, and include time for whole-class activities, student-directed free play, gross motor function (such as outdoor play), meals/snacks, and naps or rest time. All teachers at DOE-run programs were already required to hold a New York State teaching certificate, which requires a BA, and the DOE required that NYCEEC teachers earn one within a 3-year period (Reid et al., 2019). To promote process quality, the city collects and monitors scores across the aforementioned ECERS (Harms, Clifford, & Cryer, 2005) and CLASS rating scales (Pianta, La Paro, & Hamre, 2008). Teachers in all classrooms are offered the same opportunities for professional development and onsite coaching.

In order to enroll in UPK, parents file an application online or at a district-run family welcome center. Families can choose up to 12 programs and rank them in their order of preference. Each program establishes its own admissions priorities, which may give priority to current students (if the program also serves 3-year-olds) or to siblings. In addition, all families have priority to attend programs in the community school district in which they live (there are 32 such districts in NYC). For programs located in traditional elementary schools, families zoned for the school have priority. There is one enrollment deadline (March 30th), and all families that apply receive a pre-K offer in May. Families can then choose to accept their offer by registering in person at the program, and can remain on waitlists for programs which they ranked higher but to which they were not admitted.

The proportion of four-year-olds enrolled in UPK varies significantly by race and ethnicity. In the 2017-18 school year, for instance, about 62%, 72%, and 88% of black, Hispanic, and Asian four-year-olds in the city were enrolled, compared to half (52%) of white four-year-olds.[1]

Data and Methods

Data

We used publicly available data from New York City's OpenData portal[2] to create a dataset that includes all preschool providers in the city that served four-year-olds during the 2017-18 school year. For all UPK providers in New York City, these data include program characteristics for the 2017-18 school year (e.g., full-day classes, early drop off/late pick up, number of meals served, program capacity), administrative information such as street addresses, and auspice (whether a program is district-run or a NYCEEC). These data also contain multiple measures of program quality collected between the 2014-15 and 2016-17 school years, including the two most widely used measures of early childhood classroom quality (ECERS and CLASS). Finally, they contain provider-level enrollment for the 2017-18 year, both overall and disaggregated by race/ethnicity and by gender. We supplement the provider-level data with census tract-level data from the 2013-2017 American Community Survey, which includes information about racial/ethnic and socioeconomic composition, as well as population density and total land area, for each census tract.

We conduct analyses at both the provider and the census tract-levels. Our provider-level sample contains 1,869 UPK providers that reported enrollment in the 2017-18 school year. Our census tract-level sample includes the 2,128 census tracts that have non-zero population in New York City.

Measures

UPK quality. We use two measures of quality: the ECERS – Revised Edition and the CLASS. The ECERS-Revised Edition is a 43-item scale that accounts for a wide variety of interactions, activities, and materials in pre-K. Assessors spend a minimum of 3 hours in each

classroom and 30-45 minutes interviewing the lead classroom teacher. This information is used to rate providers across six dimensions of quality: space and furnishings, personal care routines, language and literacy, learning activities, interaction, and program structure.

Assessors rate multiple items within each dimension of quality. Items are scored on a scale from 1-7, with 1, 3, 5, and 7 indicating "inadequate," "minimal," "good," or "excellent" quality, respectively.[3] Programs are assigned a subscale score for each of the six dimensions of quality listed above. They are also assigned an overall ECERS score, which is the simple mean of the 43-item scores (i.e., not the mean of the subscale scores). The result is a continuous measure of program quality on a 1-7 scale (Harms, Clifford, & Cryer, 2004). While we report mean subscale scores by subgroup, we note that the ECERS subscales have not been individually psychometrically validated.

UPK providers were typically notified at least two weeks in advance of an ECERS assessment, to ensure that the assessment occurred on a "typical day" (i.e., no field trips, assemblies, or other special events). A single classroom was rated at each provider, and this classroom was selected randomly on the morning of assessment (NYC DOE, 2016).

The CLASS is an observational measure of quality that focuses on the interactions between teachers and children in a pre-K classroom. Classrooms are observed during multiple 20-minute segments, and are rated on three dimensions of quality (emotional support, classroom organization, and instructional support). Each dimension is scored on a 1-7 scale. For all three dimensions, a score of 5 is considered the threshold for a "high quality" program; in prior research, instructional support scores are generally lower than scores for the other two dimensions (Pianta, La Paro, & Hamre, 2008). The subscale scores are commonly considered separately, but we also construct a measure of overall classroom quality by taking the average score across the three dimensions.

In contrast to the ECERS, most or all classrooms at each UPK provider were assessed using the CLASS. Specifically, all classrooms were assessed for providers with fewer than seven classrooms. For providers with seven or more classrooms, at least 75% of classrooms were randomly selected to be assessed. Each classroom was observed for a minimum of two 20-minute segments (NYC DOE, 2016).

Notably, because the DOE observers inspect pre-K programs on a rolling basis, some providers in our sample are missing these quality ratings. Specifically, ECERS and CLASS ratings are missing for about 4.5% and 17% of the 1,869 UPK providers in our sample, respectively. Appendix Table A1 shows OLS estimates that examine how programs with missing quality ratings differed from those with non-missing ratings. As expected, programs that were new entrants into the UPK program were more likely to be missing both types of ratings. While we find no evidence of differences across programs that were missing ECERS ratings, we find that programs missing CLASS ratings were more likely to be DOE-sponsored, and served fewer black students. We keep this caveat in mind throughout, particularly when conducting spatial analyses.

UPK demographics. We classify providers based on the majority race/ethnicity of enrolled students. Specifically, we create indicators for programs that were majority white, black, Hispanic, and Asian, or had no majority (i.e., no group over 50%). Appendix Table A2 provides further details about how the demographic composition of these programs varies. Specifically, it shows the percent of white, black, Hispanic, and Asian students enrolled at the 25th, 50th, and 75th percentiles, by majority race/ethnicity.

UPK availability. To evaluate the availability of UPK providers in different areas of the city, we calculate the straight-line distance between the centroid of each census tract and each UPK provider in the city. We then calculate whether each tract is located within .25 miles or .5 miles of

a high-quality UPK provider, defined as a provider in the top 50%, 25%, or 10% of *either* the ECERS or CLASS ratings (high-quality ECERS and high-quality CLASS providers are defined separately). We chose .25 and .5 miles to represent programs that were located within a short walking distance. In a separate analysis of student-level enrollment data in NYC, we find that about 34 and 61% of UPK students attended a program within .25 and .5 miles of their residence, respectively (Authors, 2019).

Census tract demographics. We construct indicators for census tracts that are majority white, black, Hispanic, Asian, or no majority (i.e., no group over 50%), based on total population counts from the 2013-2017 American Community Survey (ACS). We also construct a composite measure of census tract-level socioeconomic status (SES) from multiple measures contained in the ACS. Specifically, we conduct a factor analysis of unemployment and poverty rates, as well as the percentage of the population that uses public assistance, and the percentage of housing units that are vacant, and use the primary factor in our models.

Analytic strategy

We first describe characteristics of the entire UPK sample. Next, we examine the average demographic composition of UPK providers across all children in UPK (i.e. weighting provider observations by enrollment). We also examine the average characteristics of programs for white, black, Hispanic, and Asian children (weighting by enrollment in these respective categories). We then follow the same procedure to examine mean program quality as measured by the ECERS and CLASS scales, for all students and separately by race and ethnicity.

Next, we turn to OLS regression to examine differences in UPK quality based on the racial/ethnic composition of enrolled students. We estimate models of the form:

$$Y_{pt} = \beta_0 + \beta_1 RACE'_{pt} + \beta_2 DOE_{pt} + \beta_3 INC\_REQ_{pt} + \Gamma_t + \varepsilon_{pt} \qquad (1)$$

Here, Y_{pt} is a measure of quality (overall ECERS or CLASS rating) for provider p in census tract t. $RACE'_{pt}$ is a vector that includes indicators for providers that are majority black, majority Hispanic, majority Asian, and no majority (omitting majority white). DOE_{pt} is an indicator for providers that are district-operated, and INC_REQ_{pt} is an indicator for providers that had an income requirement for at least some of their pre-K slots. We alternately include district and tract-level fixed effects (Γ_t) to account for spatial differences in program quality. Both the DOE and income requirement indicators as well as the fixed effects are used to assess whether any racial/ethnic disparities we observe are associated with systematic differences in provider type or location. The β_1s are the coefficients of interest and represent estimates of how quality differs between providers that are majority white and majority black/Hispanic/Asian.

Next, we shift our level of analysis to the census tract-level, and estimate how availability of UPK differs depending on the demographics of the local population. Here, we estimate a series of logistic regression models predicting whether a given census tract was located within .25 or .5 miles of a *high-quality* UPK provider (defined as a provider in the top 50%, 25%, or 10% of ECERS ratings). These models take the following form:

$$\ln\left(\frac{Y_t}{1-Y_t}\right) = \beta_0 + \beta_1 RACE'_t + \beta_2 SES_t + \beta_3 POPDEN_t + \beta_4 LAND\_AREA_t + \Gamma_i + \varepsilon_t \quad (2)$$

Where Y_t is a tract-level measure of nearby UPK availability (e.g., any nearby high-quality providers) and $RACE'_t$ is a vector that includes indicators for tracts that are majority black, majority Hispanic, majority Asian, or no majority (omitting majority white). These models also control for a composite measure of SES as well as the population density and total land area of each census tract. Γ_i is either a borough or district fixed effect. The β_1s, which provide estimates

of how availability of UPK supply varies by the racial/ethnic majority of census tracts, are the coefficients of interest.

Results

UPK provider and Census tract characteristics

Table 1, Panel A shows average UPK program characteristics across all providers in NYC. It shows that in 2017, 34 percent of providers were DOE programs operating in K-12 schools, 4 percent were DOE-sponsored independent pre-k centers, 61 percent were NYCEECs, and just 1 percent were charters. Table 1 also highlights student composition across UPK providers; 17 percent were majority white, 20 percent were majority black, 30 percent were majority Hispanic, 10 percent were majority Asian, and 23 percent had no racial/ethnic majority. The average ECERS rating was 4.24 and the average CLASS rating was 5.25.

Panel B shows the demographic composition of Census tracts across NYC. Thirty-four percent of tracts were majority white (i.e., more than 50% white residents), 21% were majority black, 19 percent were majority Hispanic, and 5% were majority Asian. In 21% of tracts, no racial/ethnic subgroup constituted more than 50% of residents.

Demographic composition of UPK providers

Table 2 shows the demographic composition of UPK providers overall, as well as the average demographic composition experienced by white, black, Hispanic, and Asian students. Across the entire system, 20.4 percent of students were white, 22 percent were black, 37.7 percent were Hispanic, 16.8 percent were Asian, and 3.2 percent were some other race/ethnicity.

Similar to Potter (2016), we find substantial segregation by race/ethnicity. For instance, in 2017-18 white children on average attended a program with about 61% white children and fewer

than 6% black children. For black children, those percentages are nearly reversed (57% black children, 5% white children). We also document considerable segregation of Hispanic and Asian students. For instance, on average, Hispanic children attended providers that enrolled about 59% Hispanic children, despite making up only about 38% of children in the city. Similarly, Asian children attended providers with about 50% Asian students on average, despite making up 17% of the population. In Appendix Table A3, we show analogous results for kindergarten and first grade, using publicly available data from the NYC DOE. Consistent with Potter (2016), we find that UPK is more racially segregated than either kindergarten or first grade.

Segregation indices offer another approach to describing segregation in UPK. We employ the multigroup version of Theil's entropy index (H), which provides an estimate of how the diversity of individual providers compares to the diversity of the UPK system as a whole (Reardon & Firebaugh, 2002). We find that UPK has an entropy index value of .44, which means that individual providers are 44% less diverse, on average, than the UPK system overall. We also calculated a series of two-group entropy indices to examine the segregation of individual groups. Here we find that white children are the most heavily segregated (H=.50), followed by black children (H=.43), Asian children (H=.41), and Hispanic children (H=.29).

Program quality

To provide context for how quality in NYC compares to other contexts, Appendix Tables A4 and A5 compare NYC pre-K's average ECERS and CLASS ratings, respectively, with available estimates of other state and district pre-K programs. The average ECERS rating in NYC was 4.24, and the average scores for CLASS emotional support, classroom organization, and instructional support were 6.38, 6.20, and 3.13, respectively. As other sources have noted

(CityHealth & NIEER, 2019), overall quality in NYC is quite high, outpacing the quality ratings of other longstanding pre-K programs, particularly in terms of CLASS emotional support and classroom organization subscales.

Given the uneven distribution of students by children's race/ethnicity, we next examine whether and how classroom quality differs along the same dimensions. We note the correlations between ECERS and CLASS ratings in NYC in Appendix Table A6. In general, we find that the overall correlation between these measures is quite low (.17), as are the correlations between more closely related subscales. For comparison, Mashburn et al. (2008) considered the association between ECERS ratings and CLASS subscales, and reported correlations of .54 and .43 for CLASS emotional support and CLASS instructional support, respectively. Although this difference is notable, the correlations in Mashburn et al. were based on ratings collected in the same classroom. By contrast, in NYC these assessments are frequently conducted in different classrooms at the same UPK site.[4] Observed differences may also reflect that UPK teacher quality patterns parallel those from K-12, where scholars have determined that most of the variation in teacher quality is within, rather than between, schools (Jackson, Rockoff, & Staiger, 2014).

With this caveat in mind, Table 3 shows ECERS and CLASS ratings for UPK providers throughout the city, and examines how average quality ratings differ for white, black, Hispanic, and Asian children. On average, children in each group attended a provider that scored below 5.0 (i.e. "good quality") on the ECERS scale (ranging from 4.06 to 4.44). These ratings are in line with those reported in other contexts (see Table A4). By contrast, CLASS ratings were quite high across all groups of children. For instance, the average scores for the emotional support and classroom organization subscales were well above the "high quality" benchmark of 5 for all groups of children (ranging from 6.11 to 6.52), and were considerably higher than those reported for other

prominent pre-K programs including Tulsa (Phillips, Gormley, & Lowenstein, 2009), Boston (Weiland et al., 2013), and Georgia (Bassok & Galdo, 2016). Instructional support scores were lower (ranging from 3.06 to 3.28), but comparable to that found in in other contexts (see Table A5).

Next, we document large race- and ethnicity-based gaps on both measures. For instance, white children, on average, attended programs that were approximately .51 standard deviations (SD) higher on ECERS and .48 SD higher on CLASS than did black children. We also document significant (and generally large) black-white gaps across all subscales for each measure (Table 3, Panel B). We note smaller, but still substantial, Hispanic-white and Asian-white gaps in provider quality. For instance, white children, on average, attended providers that scored about .23 SD higher on the ECERS and .34 SD higher on the CLASS than did Hispanic children. The ECERS and CLASS gaps between white and Asian children were .09 and .17 SD, respectively. In Appendix Table A7, we also examine how a host of other program characteristics vary by race and ethnicity.

Moving to a regression framework, in Table 4 we examine how ECERS and CLASS ratings differed based on the demographic composition of UPK programs. Panel A shows estimates of disparities in ECERS ratings between providers that served majority white students and those that served majority black, Hispanic, and Asian students. We find that providers serving majority black children scored on average .6 SD lower on the ECERS than did providers serving majority white children (column 1); the parallel differences for majority-Hispanic and majority-white providers is a gap of .26 SD. Coefficient estimates from the models with census tract fixed effects (column 3) reflect similarly large differences in quality among providers in the same census tracts. These

estimates are not always statistically significant, however, which may reflect the reduction in variation in providers' majority racial group within individual tracts.

We also examine the extent to which these disparities change when controlling for a limited set of program characteristics (columns 4 and 5 of Table 4). We find that ECERS disparities are partly explained by the higher proportion of black children who attend DOE-sponsored programs as well as programs that are means-tested and available only to low-income children (column 5 compared with column 3). However, even with these controls, there is still a large gap in the average quality of programs attended by black and white children, though this difference is not statistically significant.

Panel B shows disparities in quality for the CLASS scale. Similar to the ECERS, we document large gaps in quality for providers that served predominantly white children and those that served predominantly black or Hispanic children. For instance, majority-black providers scored .5 SD lower, on average, than did majority white providers (Panel B, column 1). The analogous gap for majority-white and majority-Hispanic providers was .39 SD. Unlike the ECERS results, we find that majority-Asian providers scored significantly lower on the CLASS than did majority-white providers (β = .25 SD). We find that the magnitude of disparities is generally consistent when examining these patterns within the same school district (column 2) or the same census tract (column 3), though again the standard errors are considerably larger at the census tract level. Importantly, these patterns are unchanged if we limit the sample to providers who had both ECERS and CLASS ratings (results available upon request).

Appendix Figures A1 and A2 show the distributions of ECERS and CLASS scores for majority white, black, Hispanic, and Asian providers. These figures show that the disparities we

document in Table 4 occurred throughout the quality distribution, and were not driven by a few extremely low-quality programs.

Geographic distribution of UPK supply

We next examine how proximity to high-quality UPK providers (as measured by ECERS ratings) varied with the demographic composition of the local population; these models are estimated at the census tract level. Table 5 shows how the likelihood of being located near a high-quality UPK provider differed depending on the majority racial/ethnic group in each census tract (again these regression models control for SES, tract size and population density). We include analogous results based on CLASS ratings in Appendix Table A8, though we caution that these are likely to suffer from selection bias due to the missing data issues described above as well as compression in the CLASS ratings distribution (i.e., the standard error for these ratings is .47, compared with .73 for ECERS ratings – Table 3, Panel A).

Table 5 (Panel A) shows that tracts with a majority black population were significantly less likely to be located within .25 miles of a provider that scored high on the ECERS, whether we consider the top 50%, the top quartile, or the top decile of providers. For instance, compared to majority white census tracts, the average tract with more than 50% black residents was 17 percentage points less likely to be located within .25 miles of a program in the top half of the ECERS distribution (column 1). When considering tracts within the same school district (columns 4-6), the magnitude of these associations was smaller, but we still find a negative association between tracts with predominantly black residents and proximity to high-quality UPK providers. Specifically, majority black tracts were 8 and 11 percentage points less likely, respectively, to be located within .25 miles of a provider in the top half or top quartile of the ECERS distribution.

By contrast, we find no consistent evidence that proximity to high-quality UPK providers differs for residents of majority Hispanic or Asian tracts. Panel B shows analogous estimates for providers located within .5 miles. We find no consistent evidence that tract demographics were associated with the presence of high-quality providers at this distance.

We demonstrate these findings visually in Appendix Figures A3-A7, with a series of maps that show the locations of higher and lower quality UPK providers in each borough, superimposed on a map depicting the percentage of black residents in each census tract. Consistent with Table 5, availability of high-quality providers varies systematically with the racial composition of tracts. Specifically, providers in the bottom quartile of ECERS ratings were considerably more likely to be located in census tracts with a higher percentage of black residents, and those in the top quartile were more likely to be located in areas with a lower percentage of black residents.

Discussion

In this paper, we contribute to the body of literature assessing racial disparities in access to high-quality pre-K programs. To do so, we examine the case of NYC's Universal Pre-K program, which underwent an ambitious and rapid expansion with a strong focus on quality. Today, NYC serves nearly 5% of all 4-year old US children attending a public pre-K program. To its credit, the district's average pre-K quality measures are comparable to or higher than those in other highly regarded programs.

At the same time, we document meaningful variation in pre-K classroom quality across programs attended by students of different racial/ethnic backgrounds. These disparities are largest between white and black children, but we also document substantial quality gaps between programs attended by Hispanic and white children. These disparities are apparent across multiple measures of program quality, and within each of the city's five boroughs.

To benchmark the magnitude of a .5 SD disparity in CLASS and ECERS ratings between the average programs attended by white and black students, consider that a recent meta-analysis found that a one SD increase in ECERS was associated with between a .01 and .02 SD increase in language and pre-literacy skills (Hong et al., 2019). The same meta-analysis found that a one SD increase in CLASS was associated with between .03 and .06 SD in the same domains. This suggests that the quality disparities we document are likely to contribute to small but meaningful disparities in student outcomes. Importantly, a number of studies that examined the associations between ECERS and CLASS ratings and student outcomes found no significant associations (Brunsek et al, 2017; Hong et al., 2019). We note, however, that classroom environment disparities by race/ethnicity, which include measures of learning activities, social interactions, and organization, are of policy interest, even if they do not directly impact future outcomes.

Exploring the mechanisms that may produce these patterns, we provide evidence that white-black disparities in provider quality are partially due to differences in local supply. Specifically, we find that areas of the city with predominantly black residents are less likely to be located near high-quality UPK providers, even after accounting for residents' socioeconomic status. Further, we find that disparities in the spatial distribution of quality occur in a highly localized context, apparent within a .25 but not a .5 mile radius of children's homes. Notably, many students attend pre-k within close proximity of their home, and in NYC, black students are more likely to attend UPK within .25 miles of their home than are white students (34% versus 25% in 2017-18) (Authors, 2019). These findings provide guidance to policymakers who are considering choosing sites for new programs, or allocating program improvement support.

Given the high degree of residential segregation in NYC, our estimates may be conservative in portraying differences in program quality across tracts. For instance, in our data

majority black districts had an average of 72% black residents, compared to just 3% black residents in majority white districts. By contrast, majority white districts averaged 71% white residents and 5% black residents. For this reason, the differences between majority white and majority black tracts may underestimate the extent of disparities.

An important factor we were not able to examine in our provider-level analysis is the extent to which race/ethnicity-based gaps in quality also reflect income-based quality disparities. Research suggests that programs that are racially segregated tend to be economically segregated as well (Orfield & Frankenberg, 2014). Studies from K-12 find that schools that serve lower-income students are harder to staff, attracting teachers and principals with less experience and poorer credentials. Although not well-studied, the same dynamic may exist within the pre-K sector.

A related possibility is that differences in quality may result not from characteristics of teachers or providers, but directly due to differences in peer composition. Peer composition plays an important role in student learning during the preschool year, as lower-performing children learn more when they interact with higher-performing peers (Reid & Ready, 2013; Weiland & Yoshikawa, 2014). This problem may be exacerbated when providers are required to admit students based on an income requirement as a function of their funding stream, effectively concentrating poor children in higher-poverty classrooms.

Another potential explanation for documented disparities is differences in the extent to which parents are constrained in their ability to learn and use information about program quality. Though parents' quality evaluations are largely unrelated to objective measures of structural and process quality (Bassok, Markowitz, Player, & Zagardo, 2018), higher-income parents have more time and resources they can dedicate toward finding programs, more information available through social networks, and greater familiarity with institutional navigation (Authors). For these reasons,

differences in pre-K quality are likely to reflect differences in parent constraints rather than differences in parent preferences. For instance, although most parents report a strong preference for high-quality preschool, lower-income parents are more likely to be limited by cost and proximity (Meyers & Jordan, 2006). This is important when thinking about the types of supports necessary to improve access to high-quality pre-K for low-income, black, and Hispanic children.

Despite the unique size and scope of New York City's universal pre-K program, the results we document reflect challenges that other communities are likely to face when adopting universal public pre-K programs. Although the stated goal of these programs is often to close race and income-based achievement gaps, this can be quite difficult to accomplish in a context with a high degree of residential segregation and structural inequality. Universal programs generally enjoy more public and financial support, but high levels of segregation present additional barriers to providing uniform levels of quality to all groups (Gomez-Velez, 2015; Valentino, 2018).

There are a few limitations worth noting. First, this research is descriptive. While we identify the spatial distribution of supply as one potential driver of black-white racial disparities, we cannot make definitive claims about the causes of these disparities. Second, we lack information about the socioeconomic status of children enrolled in UPK programs, and thus are unable to examine the extent to which quality differs by income. Third, our analyses of the spatial supply rely on the fact that 61% of students enroll in a program within .5 miles of their home address (Authors); future analyses should use group-specific distance enrollments to consider how local supply affects racial disparities in the program quality students experience. Finally, white New York 4-year olds are disproportionately enrolled in private pre-K programs for which we lack data on program quality. Whether private programs offer higher-quality environments than UPK

providers is an open empirical question, and the findings presented here reflect within-UPK racial

disparity estimates.

Notes:

[1] Authors' estimates, using publicly-available UPK enrollment data through NYC's OpenData portal in conjunction with American Communities Survey data from 2017.

[2] We use information from the following four datasets on the NYC OpenData portal: the 2018 "Universal Pre-K School Locations" file, the "2018 Pre-K School Directory", the "2016-2017 School Quality Reports PK" file, and the "2015-2018 Demographic Snapshot Pre-K For All."

[3] The 43 items in the ECERS are each scored based on multiple indicators of quality, each of which is rated on a scale from 1-7. The assigned score for each item is the minimum score across relevant indicators (i.e. the *maximum* score at which all indicators of quality are met).

[4] We also examined correlations between ECERS and CLASS ratings for providers with a small number of students, reasoning that these providers are likely to have only have a single classroom. We found that these correlations are a bit larger, but still quite similar to those that use the entire sample. For instance, the correlations between CLASS & ECERS ratings for providers with 25 or fewer students is .197 (N=520), compared with .171 for the full sample (N=1633). Analogous correlations for providers with 20 or fewer (N=448), 18 or fewer (N=374), and 15 or fewer (N=183) students are .196, .204, and .188, respectively.

References

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2016). Teacher quality and learning outcomes in kindergarten. *The Quarterly Journal of Economics*, *131*(3), 1415–1453. https://doi.org/10.1093/qje/qjw016

Bassok, D., Fitzpatrick, M., & Loeb, S. (2011). Disparities in child care availability across communities: Differential reflection of targeted interventions and local demand. *CEPA Working Paper*.

Bassok, D., & Galdo, E. (2016). Inequality in preschool quality? Community-level disparities in access to high-quality learning environments. *Early Education and Development*, *27*(1), 128–144. https://doi.org/10.1080/10409289.2015.1057463

Bassok, D., Markowitz, A. J., Player, D., & Zagardo, M. (2018). Are parents' ratings and satisfaction with preschools related to program features? *AERA Open*, *4*(1), 2332858418759954. https://doi.org/10.1177/2332858418759954

Brunsek, A., Perlman, M., Falenchuk, O., McMullen, E., Fletcher, B., & Shah, P. S. (2017). The relationship between the Early Childhood Environment Rating Scale and its revised form and child outcomes: A systematic review and meta-analysis. *PLOS ONE*, *12*(6), e0178512. https://doi.org/10.1371/journal.pone.0178512

Burchinal, M., Nelson, L., Carlson, M., & Brooks-Gunn, J. (2008). Neighborhood characteristics, and child care type and quality. *Early Education and Development*, *19*(5), 702–725. https://doi.org/10.1080/10409280802375273

Campbell, F. A., Pungello, E. P., Burchinal, M., Kainz, K., Pan, Y., Wasik, B. H., Barbarin, O. A., Sparling, J. J., & Ramey, C. T. (2012). Adult outcomes as a function of an early childhood educational program: An Abecedarian Project follow-up. *Developmental Psychology*, *48*(4), 1033–1043. https://doi.org/10.1037/a0026644

Campbell, S. L., & Ronfeldt, M. (2018). Observational Evaluation of Teachers: Measuring More Than We Bargained for? *American Educational Research Journal*, *55*(6), 1233–1267. https://doi.org/10.3102/0002831218776216

CityHealth, & NIEER. (2019). *Pre-K in American cities*. https://static1.squarespace.com/static/5ad9018bf93fd4ad7295ba8f/t/5c4883ef7924e8130fafee7d/1548256243676/Pre-K+in+American+Cities+1-23.pdf

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics*, *1*(3), 111–134. https://doi.org/10.1257/app.1.3.111

Early, D. M., Bryant, D. M., Pianta, R. C., Clifford, R. M., Burchinal, M. R., Ritchie, S., Howes, C., & Barbarin, O. (2006). Are teachers' education, major, and credentials related to classroom quality and children's academic gains in pre-kindergarten? *Early Childhood Research Quarterly*, *21*(2), 174–195. https://doi.org/10.1016/j.ecresq.2006.04.004

Early, D. M., Iruka, I. U., Ritchie, S., Barbarin, O. A., Winn, D.-M. C., Crawford, G. M., Frome, P. M., Clifford, R. M., Burchinal, M., Howes, C., Bryant, D. M., & Pianta, R. C. (2010). How do pre-kindergarteners spend their time? Gender, ethnicity, and income as predictors of experiences in pre-kindergarten classrooms. *Early Childhood Research Quarterly*, *25*(2), 177–193. https://doi.org/10.1016/j.ecresq.2009.10.003

Frankenberg, E. (2016). *Segregation at an early age*. Center for Education and Civil Rights.

Friedman-Krauss, A., Barnett, W. S., Weisenfeld, G. G., Kasmin, R., DiCrecchio, N., & Horowitz, M. (2018). The state of preschool 2017: State preschool yearbook. *National Institute for Early Education Research*.

Goldstein, D. (2016, September 7). Bill de Blasio's Pre-K crusade. *The Atlantic*. https://www.theatlantic.com/education/archive/2016/09/bill-de-blasios-prek-crusade/498830/

Gomez-Velez, N. (2014). Can universal pre-K overcome extreme race and income segregation to reach New York's neediest children—The importance of legal infrastructure and limits of the law. *Cleveland State Law Review*, *63*(2), 319–354.

Gordon, R. A., Hofer, K. G., Fujimoto, K. A., Risk, N., Kaestner, R., & Korenman, S. (2015). Identifying high-quality preschool programs: New evidence on the validity of the Early Childhood Environment Rating Scale–Revised (ECERS-R) in relation to school readiness goals. *Early Education and Development*, *26*(8), 1086–1110. https://doi.org/10.1080/10409289.2015.1036348

Harms, T., Clifford, R., & Cryer, D. (2005). *Early Childhood Environment Rating Scale: Revised edition*. Teacher's College Press.

Hatfield, B. E., Burchinal, M. R., Pianta, R. C., & Sideris, J. (2016). Thresholds in the association between quality of teacher–child interactions and preschool children's school readiness skills. *Early Childhood Research Quarterly*, *36*, 561–571. https://doi.org/10.1016/j.ecresq.2015.09.005

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. (2010). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics*, *94*(1), 114–128. https://doi.org/10.1016/j.jpubeco.2009.11.001

Hong, S. L., Sabol, T. J., Burchinal, M. R., Tarullo, L., Zaslow, M., & Peisner-Feinberg, E. S. (2019). ECE quality indicators and child outcomes: Analyses of six large child care studies. *Early Childhood Research Quarterly*, *49*, 202–217. https://doi.org/10.1016/j.ecresq.2019.06.009

Jackson, C. K., Rockoff, J. E., & Staiger, D. O. (2014). Teacher effects and teacher-related policies. *Annual Review of Economics*, *6*(1), 801-825.

Kaplan, T., & Hernández, J. (2014, March 29). State Budget Deal Reached; $300 Million for New York City Pre-K. *The New York Times*. https://www.nytimes.com/2014/03/30/nyregion/cuomo-new-york-state-budget.html?module=inline

Karoly, L. A., Ghosh-Dastidar, B., Zellman, G. L., Perlman, M., & Fernyhough, L. (2008). *Prepared to learn: The nature and quality of early care and education for preschool-age children in California* [Product Page]. https://www.rand.org/pubs/technical_reports/TR539.html

Kucsera, J. (2014). *New York State's extreme school segregation: Inequality, inaction and a damaged future*. Civil Rights Project / Proyecto Derechos Civiles. https://eric.ed.gov/?id=ED558739

Le, V.-N., Schaack, D. D., & Setodji, C. M. (2015). Identifying baseline and ceiling thresholds within the Qualistar Early Learning Quality Rating and Improvement System. *Early Childhood Research Quarterly*, *30*, 215–226. https://doi.org/10.1016/j.ecresq.2014.03.003

Lipsey, M. W., Farran, D. C., & Hofer, K. G. (2015). *A randomized control trial of a statewide voluntary prekindergarten program on children's skills and behaviors through third grade*. Peabody Research Institute. https://eric.ed.gov/?id=ED566664

Loeb, S., Kalogrides, D., & Horng, E. L. (2010). Principal preferences and the uneven distribution of principals across schools. *Educational Evaluation and Policy Analysis*, *32*(2), 205–229. https://doi.org/10.3102/0162373710369833

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., Burchinal, M., Early, D. M., & Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, *79*(3), 732–749. https://doi.org/10.1111/j.1467-8624.2008.01154.x

Meyers, M. K., & Jordan, L. P. (2006). Choice and accommodation in parental child care decisions. *Community Development*, *37*(2), 53–70. https://doi.org/10.1080/15575330609490207

NYC Department of Education. (2016). *Pre-K for All Handbook: 2016-17*.

NYC Department of Education. (2018). *Pre-K program assessments: Classroom Assessment Scoring System (CLASS) and Early Childhood Environmental Rating Scale-Revised (ECERS-R) release*.

Orfield, G., & Frankenberg, E. (2014). Increasingly segregated and unequal schools as courts reverse policy. *Educational Administration Quarterly*, *50*(5), 718–734. https://doi.org/10.1177/0013161X14548942

Phillips, D. A., Gormley, W. T., & Lowenstein, A. E. (2009). Inside the pre-kindergarten door: Classroom climate and instructional time allocation in Tulsa's pre-K programs. *Early Childhood Research Quarterly*, *24*(3), 213–228. https://doi.org/10.1016/j.ecresq.2009.05.002

Phillips, D., Lipsey, M. W., Dodge, K., Haskins, R., Bassok, D., Burchinal, M. R., Duncan, G. J., Dynarski, M., Magnuson, K. A., & Weiland, C. (2017). *Puzzling it out: The current state of scientific knowledge on pre-kindergarten effects*. Brookings.

Pianta, R., La Paro, K., & Hamre, B. (2008). *Classroom Assessment Scoring System Manual: Pre-K*. Paul H. Brookes Publishing Co.

Potter, H. (2015). *Lessons from New York City's Universal Pre-K expansion: How a focus on diversity could make it even better*. The Century Foundation.

Potter, H. (2016). *Diversity in New York City's Universal pre-K classrooms*. The Century Foundation.

Reardon, S. F., & Firebaugh, G. (2002). Measures of Multigroup Segregation. *Sociological Methodology*, *32*(1), 33–67. https://doi.org/10.1111/1467-9531.00110

Reid, J. L., Melvin, S. A., Kagan, S. L., & Brooks-Gunn, J. (2019). Building a unified system for universal Pre-K: The case of New York City. *Children and Youth Services Review*, *100*, 191–205. https://doi.org/10.1016/j.childyouth.2019.02.030

Reid, J. L., & Ready, D. D. (2013). High-quality preschool: The socioeconomic composition of preschool classrooms and children's learning. *Early Education and Development*, *24*(8), 1082–1111. https://doi.org/10.1080/10409289.2012.757519

Sabol, T. J., Hong, S. L. S., Pianta, R. C., & Burchinal, M. R. (2013). Can rating pre-K programs predict children's learning? *Science*, *341*(6148), 845–846. https://doi.org/10.1126/science.1233517

Shapiro, A., Martin, E., Weiland, C., & Unterman, R. (2019). If you offer it, will they come? Patterns of application and enrollment behavior in a universal prekindergarten context. *AERA Open*, *5*(2), 2332858419848442. https://doi.org/10.1177/2332858419848442

Simon, N. S., & Johnson, S. M. (2015). Teacher turnover in high-poverty schools: What we know and can do. *Teachers College Record*, *117*(3).

The Education Trust. (2019). *Young learners, missed opportunities: Ensuring that black and Latino children have access to high-quality state-funded preschool*. https://s3-us-east-2.amazonaws.com/edtrustmain/wp-content/uploads/2014/09/05162154/Young-Learners-Missed-Opportunities.pdf

Torquati, J. C., Raikes, H. H., Huddleston-Casas, C. A., Bovaird, J. A., & Harris, B. A. (2011). Family income, parent education, and perceived constraints as predictors of observed program quality and parent rated program quality. *Early Childhood Research Quarterly*, *26*(4), 453–464. https://doi.org/10.1016/j.ecresq.2011.03.004

Valentino, R. (2018). Will public pre-K really close achievement gaps? Gaps in prekindergarten quality between students and across states. *American Educational Research Journal*, *55*(1), 79–116. https://doi.org/10.3102/0002831217732000

Weiland, C., Ulvestad, K., Sachs, J., & Yoshikawa, H. (2013). Associations between classroom quality and children's vocabulary and executive function skills in an urban public prekindergarten program. *Early Childhood Research Quarterly*, *28*(2), 199–209. https://doi.org/10.1016/j.ecresq.2012.12.002

Weiland, C., & Yoshikawa, H. (2014). Does higher peer socio-economic status predict children's language and executive function skills gains in prekindergarten? *Journal of Applied Developmental Psychology*, *35*(5), 422–432. https://doi.org/10.1016/j.appdev.2014.07.001

Wells, A. S., Fox, L., & Cordova-Cobo, D. (2016, September). How racially diverse schools and classrooms can benefit all students. *The Education Digest; Ann Arbor*, *82*(1), 17–24.

Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., Gormley, W. T., Ludwig, J., Magnuson, K. A., Phillips, D., & Zaslow, M. J. (2013). *Investing in our future: The evidence base on preschool education*. Society for Research in Child Development. http://repositorio.minedu.gob.pe/handle/123456789/4015

Table 1. *Sample characteristics*

Panel A. UPK providers

	M	SD
DOE K12	0.34	
DOE pre-K	0.04	
NYCEEC	0.61	
Charter	0.01	
Total enrollment	38.09	25.89
Majority white	0.17	
Majority black	0.20	
Majority Hispanic	0.30	
Majority Asian	0.10	
No majority	0.23	
ECERS rating	4.24	0.73
CLASS rating	5.25	0.53
N	1869	

Panel B. Census tracts

	M	SD
Majority white	0.34	
Majority black	0.21	
Majority Hispanic	0.19	
Majority Asian	0.05	
No majority	0.21	
N	2128	

Note. Authors' calculations using publicly-available data from NYC's OpenData portal.
DOE K12 - Pre-k program in K-12 school run by NYC Department of Education; DOE pre-k - Independent pre-k center run by NYC Department of Education; NYCEEC - New York City Early Education Center

Table 2. *Average demographic composition of UPK classrooms, by student race/ethnicity*

	All students		White		Black		Hispanic		Asian	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
% white	20.4	(28.7)	60.7	(29.4)	5.1	(11.9)	10.3	(17.0)	14.1	(18.8)
% black	22.0	(27.8)	5.5	(10.9)	57.2	(28.7)	17.3	(19.7)	7.0	(13.6)
% Hispanic	37.7	(28.2)	19.1	(18.5)	29.7	(23.7)	58.8	(25.2)	24.9	(19.9)
% Asian	16.8	(23.8)	11.6	(14.7)	5.3	(10.4)	11.1	(16.5)	50.5	(29.3)
% other race/ethnicity	3.2	(5.4)	3.3	(4.7)	2.7	(4.5)	2.5	(4.2)	3.6	(5.4)

Note. Authors' calculations using publicly-available data from NYC's OpenData portal. All estimates are weighted to be representative of the average child in each group in the 2017-2018 school year.

Table 3. *Pre-k program quality by student race/ethnicity*

Panel A. Program quality

	All students		White		Black		Hispanic		Asian	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Overall ECERS score (1-7)	4.28	(0.73)	4.44	(0.72)	4.06	(0.73)	4.27	(0.73)	4.38	(0.69)
Language reasoning	5.32	(1.01)	5.46	(0.97)	5.08	(1.08)	5.33	(0.99)	5.41	(0.94)
Interaction	5.27	(1.23)	5.47	(1.05)	4.94	(1.38)	5.31	(1.23)	5.38	(1.16)
Activities	4.63	(0.96)	4.81	(0.97)	4.37	(0.95)	4.63	(0.97)	4.77	(0.87)
Personal care routines	2.76	(0.86)	2.78	(0.87)	2.71	(0.82)	2.77	(0.85)	2.75	(0.94)
Space and furnishings	3.96	(0.77)	4.06	(0.79)	3.82	(0.74)	3.95	(0.76)	4.04	(0.77)
Program structure	4.00	(1.42)	4.42	(1.38)	3.64	(1.42)	3.87	(1.40)	4.21	(1.35)
Overall CLASS score (1-7)	5.27	(0.47)	5.39	(0.43)	5.17	(0.53)	5.23	(0.45)	5.31	(0.43)
Emotional support	6.43	(0.45)	6.52	(0.38)	6.32	(0.53)	6.41	(0.45)	6.49	(0.42)
Classroom organization	6.26	(0.54)	6.38	(0.45)	6.11	(0.63)	6.24	(0.53)	6.34	(0.48)
Instructional support	3.12	(0.80)	3.28	(0.84)	3.08	(0.84)	3.06	(0.75)	3.12	(0.76)

Note. Authors' calculations using publicly-available data from NYC's OpenData portal. All estimates are weighted to be representative of the average child in each group in the 2017-2018 school year. Gaps in Panel B are reported in standard deviation units, and may not match reported differences in Panel A exactly due to rounding. ECERS - Early Childhood Environment Rating Scale; CLASS - Classroom Assessment Scoring System
* $p < .05$.

Table 3 (continued). *Pre-k program quality by student race/ethnicity*

Panel B. Gaps in program quality

	black-white gap	Hispanic-white gap	Asian-white gap
Overall ECERS score (1-7)	-0.51 *	-0.23 *	-0.09 *
Language reasoning	-0.38 *	-0.13 *	-0.06 *
Interaction	-0.42 *	-0.12 *	-0.07 *
Activities	-0.46 *	-0.19 *	-0.05 *
Personal care routines	-0.09 *	-0.02	-0.04 *
Space and furnishings	-0.31 *	-0.14 *	-0.02
Program structure	-0.55 *	-0.39 *	-0.15 *
Overall CLASS score (1-7)	-0.48 *	-0.34 *	-0.17 *
Emotional support	-0.45 *	-0.25 *	-0.08 *
Classroom organization	-0.50 *	-0.27 *	-0.07 *
Instructional support	-0.25 *	-0.28 *	-0.21 *

Note. Authors' calculations using publicly-available data from NYC's OpenData portal. All estimates are weighted to be representative of the average child in each group in the 2017-2018 school year. Gaps in Panel B are reported in standard deviation units, and may not match reported differences in Panel A exactly due to rounding. ECERS - Early Childhood Environment Rating Scale; CLASS - Classroom Assessment Scoring System
* $p < .05$.

Table 4. *Relationship between demographic composition of UPK providers and standardized provider quality ratings*

Panel A. ECERS ratings

	(1)	(2)	(3)	(4)	(5)
Majority black	-0.60 ***	-0.49 ***	-0.59 **	-0.57 **	-0.41
	(0.09)	(0.11)	(0.21)	(0.22)	(0.21)
Majority Hispanic	-0.26 **	-0.13	-0.27	-0.26	-0.16
	(0.08)	(0.10)	(0.18)	(0.19)	(0.17)
Majority Asian	-0.11	-0.30 **	-0.04	-0.04	-0.07
	(0.10)	(0.09)	(0.26)	(0.26)	(0.25)
No majority	-0.28 **	-0.30 **	-0.17	-0.17	-0.06
	(0.09)	(0.09)	(0.19)	(0.19)	(0.18)
N	1782	1782	1782	1782	1782

Panel B. CLASS ratings

	(1)	(2)	(3)	(4)	(5)
Majority black	-0.50 ***	-0.43 ***	-0.55 *	-0.49 *	-0.47 *
	(0.08)	(0.09)	(0.21)	(0.22)	(0.22)
Majority Hispanic	-0.39 ***	-0.33 ***	-0.41 *	-0.37	-0.37
	(0.07)	(0.07)	(0.19)	(0.19)	(0.20)
Majority Asian	-0.25 **	-0.28 **	-0.13	-0.14	-0.21
	(0.09)	(0.08)	(0.29)	(0.29)	(0.27)
No majority	-0.19 **	-0.17 *	-0.24	-0.25	-0.25
	(0.07)	(0.07)	(0.17)	(0.18)	(0.17)
N	1545	1545	1545	1545	1545

School district fixed effects		X			
Census tract fixed effects			X	X	X
Income requirement				X	X
DOE program					X
DOE pre-k program					X

Note. Authors' calculations using publicly-available data from NYC's OpenData portal. Outcomes are reported in standard deviation units. The omitted category is majority white. Estimates are weighted by 2017 school enrollment. ECERS - Early Childhood Environment Rating Scale; CLASS - Classroom Assessment Scoring System. * p <.05 ** p <.01 *** p <.001.

Table 5. *Census tract level estimates of the association between tract demographics and presence of at least 1 nearby UPK provider who scored high on the ECERS*

Panel A. Providers within .25 miles

	Any providers with ECERS rating >=					
	(1)	(2)	(3)	(4)	(5)	(6)
Tract level variable	50th pctile	75th pctile	90th pctile	50th pctile	75th pctile	90th pctile
Majority black	-0.17 **	-0.18 ***	-0.09 ***	-0.08	-0.11 *	-0.04
	(0.06)	(0.05)	(0.01)	(0.05)	(0.05)	(0.04)
Majority Hispanic	-0.03	-0.04	0.01	-0.04	-0.05	0.04
	(0.07)	(0.06)	(0.02)	(0.05)	(0.05)	(0.04)
Majority Asian	-0.02	-0.04	-0.07 ***	0.04	0.01	-0.03
	(0.09)	(0.08)	(0.02)	(0.07)	(0.06)	(0.04)
No majority	-0.08 *	-0.05	-0.01	-0.07 **	-0.04	0.02
	(0.03)	(0.03)	(0.01)	(0.02)	(0.03)	(0.02)
N	2128	2128	2128	2128	2128	2128
Borough FE	X	X	X			
District FE				X	X	X

Panel B. Providers within .5 miles

	Any providers with ECERS rating >=					
	(1)	(2)	(3)	(4)	(5)	(6)
Tract level variable	50th pctile	75th pctile	90th pctile	50th pctile	75th pctile	90th pctile
Majority black	-0.03	-0.06	-0.12 **	-0.03	0.02	0.01
	(0.02)	(0.06)	(0.04)	(0.06)	(0.06)	(0.07)
Majority Hispanic	-0.02	-0.05	0.04	-0.03	0.01	0.14
	(0.03)	(0.07)	(0.05)	(0.05)	(0.07)	(0.08)
Majority Asian	0.01	0.12 *	-0.04	0.03	0.17 ***	0.06
	(0.01)	(0.06)	(0.09)	(0.05)	(0.04)	(0.08)
No majority	0.00	0.02	0.00	0.00	0.03	0.05
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
N	2128	2128	2128	2128	2128	2128
Borough FE	X	X	X			
District FE				X	X	X

Note. Authors' calculations using publicly-available data from NYC's OpenData portal. Estimates are presented as average marginal effects, and can be interpreted as the average percentage point difference between majority white tracts and each respective group. All estimates control for the total land area and population density of each census tract, as well as a composite measure of SES. ECERS - Early Childhood Environment Rating Scale * $p < .05$ ** $p < .01$ *** $p < .001$.

Appendix

Table A1. *Associations between missing ECERS & CLASS scores and provider-level characteristics*

	Missing ECERS				Missing CLASS			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Prop. Black	-0.04	0.01	0.01	0.01	-0.15 ***	-0.04	-0.07 *	-0.06 *
	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)	(0.03)	(0.03)	(0.03)
Prop. Hispanic	-0.01	0.01	0.00	0.00	0.03	0.07 *	0.03	0.03
	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)	(0.03)	(0.03)	(0.03)
Prop. Asian	0.01	0.01	0.01	0.01	-0.04	-0.04	-0.05	-0.05
	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.03)	(0.03)
Prop. Other	0.06	0.06	0.05	0.06	0.01	0.00	-0.06	-0.07
	(0.08)	(0.07)	(0.07)	(0.07)	(0.14)	(0.11)	(0.10)	(0.10)
Entry (enrollment in 2017, not 2015)		0.42 ***	0.42 ***	0.35 ***		0.87 ***	0.92 ***	0.85 ***
		(0.02)	(0.02)	(0.02)		(0.03)	(0.02)	(0.03)
DOE K12			0.03 **	0.01			0.24 ***	0.24 ***
			(0.01)	(0.01)			(0.01)	(0.01)
DOE prek			0.02	0.02			0.00	-0.02
			(0.02)	(0.02)			(0.03)	(0.03)
Charter			0.01	0.01			0.01	0.00
			(0.05)	(0.05)			(0.07)	(0.07)
Average ECERS				-				0.02
								(0.01)
Missing ECERS				-				0.26 ***
								(0.05)
Average CLASS				0.00				-
				(0.01)				
Missing CLASS				0.06				-
				(0.05)				
N	1869	1869	1869	1869	1869	1869	1869	1869

Note. ECERS - Early Childhood Environment Rating Scale; CLASS - Classroom Assessment Scoring System. * p<.05 ** p<.01 *** p<.001.

Table A2. *Demographic composition of providers by majority race/ethnicity*

	Racial/ethnic majority (>=50%)				
	Majority white	Majority black	Majority Hispanic	Majority Asian	No majority
% white					
25th pctile	61.2	0.0	0.0	0.0	5.7
50th pctile	73.4	0.0	1.0	2.1	23.1
75th pctile	96.7	2.0	6.6	8.3	38.9
% black					
25th pctile	0.0	64.7	3.4	0.0	2.8
50th pctile	0.0	76.2	13.8	0.0	11.1
75th pctile	2.9	87.5	25.0	4.0	31.3
% Hispanic					
25th pctile	0.0	6.8	62.1	5.6	24.1
50th pctile	9.2	15.4	72.2	13.9	35.1
75th pctile	17.1	27.0	82.4	21.8	43.8
% Asian					
25th pctile	0.0	0.0	0.0	60.7	6.3
50th pctile	4.4	0.0	1.3	72.2	17.6
75th pctile	11.4	4.0	6.7	88.9	33.3
N	310	379	562	192	426

Table A3. *Demographic composition of kindergarten and first grade classrooms in NYC, by student race/ethnicity*

Panel A. Kindergarten

	All students		White		Black		Hispanic		Asian	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
% white	16.5	(22.2)	46.3	(23.6)	5.6	(11.4)	10.1	(16.0)	17.7	(18.6)
% black	24.4	(27.5)	8.2	(13.5)	55.3	(27.5)	19.1	(20.0)	7.7	(13.9)
% Hispanic	40.3	(26.4)	24.7	(17.3)	31.6	(23.1)	57.5	(24.3)	28.1	(19.2)
% Asian	15.9	(21.0)	17.0	(16.1)	5.0	(9.6)	11.1	(15.9)	43.6	(25.5)
% other non-white	2.8	(4.5)	3.8	(3.8)	2.5	(3.7)	2.1	(3.0)	2.9	(4.1)

Panel B. First grade

	All students		White		Black		Hispanic		Asian	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
% white	15.9	(21.7)	45.4	(23.8)	5.6	(11.1)	9.9	(15.6)	17.4	(18.5)
% black	24.6	(27.2)	8.6	(13.9)	54.7	(26.9)	19.0	(20.0)	8.5	(14.7)
% Hispanic	41.7	(26.7)	25.9	(18.1)	32.2	(22.9)	58.8	(24.5)	28.9	(19.5)
% Asian	15.0	(20.1)	16.3	(15.7)	5.2	(9.7)	10.4	(15.3)	42.0	(25.1)
% other non-white	2.8	(5.6)	3.8	(4.8)	2.3	(4.0)	1.9	(3.3)	3.1	(4.6)

Note. Authors' calculations using student-level enrollment data from the Research Alliance for New York City Schools. All estimates are weighted to be representative of the average child in each group in the 2017-2018 school year. * $p < .05$.

Table A4. *Comparison of reported Early Childhood Environment Rating Scale (ECERS) scores across study populations*

Location	N	Year	Total score M	SD	Source
North Carolina	2447	2014-15	5.52	0.50	Authors' calculations
Tennessee	1755	2015-16	4.88	0.91	Authors' calculations
Boston pre-K	83	Spr. 2010	4.47	0.50	Weiland, Ulvestad, Sachs, & Yoshikawa (2013)
Tennessee pre-K	114	n.r.	4.40	0.82	Denny, Hallam, & Homer (2012)
Head Start FACES*	370	Spr. 2010	4.30	n.r.	Moiduddin et al. (2012)
New York City pre-k	**1724**	**2016-17**	**4.24**	**0.73**	**Authors' calculations**
A Midwestern state	189	Fall 2013	4.12	0.84	Lin & Magnuson (2018)
California	615	Spr. 2007	4.10	1.09	Karoly et al. (2008)
NCEDL-SWEEP* (11 states)	671	2001-04	3.85	0.82	Mashburn et al. (2008)

Note. n.r. - not reported. *NCEDL: National Center for Early Development & Learning's multi-state study of pre-kindergarten; SWEEP: State-Wide Early Education Programs study; Head Start FACES: Head Start Family and Child Experiences Survey.

Table A5. *Comparison of reported Classroom Assessment Scoring System (CLASS) scores across study populations*

Location	N	Year	Emotional		Classroom		Instructional		Source
			M	SD	M	SD	M	SD	
New York City pre-k	**1493**	**2016-17**	**6.38**	**0.51**	**6.20**	**0.61**	**3.13**	**0.87**	**Authors' calculations**
San Mateo County, CA	8	2006-07	6.20	n.r.	5.10	n.r.	3.80	n.r.	American Institutes for Research (2007)
Head Start	319	2015-16	6.00	0.29	5.73	0.35	2.83	0.52	Office of Head Start (2017)
San Francisco County, CA	32	2006-07	6.00	n.r.	5.20	n.r.	3.70	n.r.	American Institutes for Research (2007)
Louisiana	636	2014-15	5.76	0.55	5.47	0.68	2.92	0.79	Bassok, Markowitz, Player, & Zagardo (2018)
Boston pre-K	83	Spr. 2010	5.63	0.60	5.10	0.68	4.30	0.84	Weiland, Ulvestad, Sachs, & Yoshikawa (2013)
NCEDL-SWEEP* (11 states)	671	2001-04	5.57	0.68	n.r.	n.r.	2.08	0.83	Mashburn et al. (2008)
Georgia pre-k	3883	2010-11	5.53	0.76	5.17	0.89	2.06	0.72	Bassok & Galdo (2016)
California	615	Spr. 2007	5.50	0.88	4.90	1.06	2.60	1.05	Karoly et al. (2008)
Eastern NC & Central PA	807	2003-04	5.35	0.69	4.82	0.83	2.59	0.95	Broekhuizen et al. (2016)
Tennessee pre-K	114	n.r.	5.30	0.87	4.70	0.93	2.50	0.86	Denny, Hallam, & Homer (2012)
Head Start FACES*	370	Spr. 2010	5.30	n.r.	4.70	n.r.	2.30	n.r.	Moiduddin et al. (2012)
Tulsa pre-k	77	Spr. 2006	5.23	0.57	4.96	0.69	3.21	0.93	Phillips, Gormley, & Lowenstein (2009)
10 cities across the US	314	2008-10	5.11	0.87	5.05	0.76	2.36	0.86	Hamre, Hatfield, Pianta, & Jamil (2013)

Note. n.r. - not reported. *NCEDL: National Center for Early Development & Learning's multi-state study of pre-kindergarten; SWEEP: State-Wide Early Education Programs study; Head Start FACES: Head Start Family and Child Experiences Survey.

Table A6. *Correlations between ECERS and CLASS scales*

	Overall CLASS score (1-7)	Emotional support	Classroom organization	Instructional support
Overall ECERS score (1-7)	0.17	0.15	0.15	0.12
Language reasoning	0.12	0.08	0.07	0.13
Interaction	0.12	0.10	0.11	0.08
Activities	0.16	0.12	0.13	0.13
Personal care routines	0.05	0.05	0.04	0.04
Space and furnishings	0.12	0.14	0.13	0.05
Program structure	0.16	0.17	0.15	0.10

Note. ECERS - Early Childhood Environment Rating Scale; CLASS - Classroom Assessment Scoring System.

Table A7. *Pre-k program characteristics and gaps by student race/ethnicity*

Panel A. Program characteristics

	All students		White		Black		Hispanic		Asian	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
DOE program	0.42	(0.49)	0.38	(0.49)	0.40	(0.49)	0.45	(0.50)	0.44	(0.50)
DOE program in K-12 school	0.35	(0.48)	0.30	(0.46)	0.38	(0.48)	0.37	(0.48)	0.32	(0.47)
DOE pre-k center	0.07	(0.26)	0.08	(0.27)	0.03	(0.16)	0.08	(0.27)	0.11	(0.31)
NYCEEC program	0.57	(0.50)	0.62	(0.49)	0.58	(0.49)	0.54	(0.50)	0.56	(0.50)
Provider offers full day seats	0.97	(0.17)	0.91	(0.29)	0.99	(0.12)	0.99	(0.12)	0.99	(0.10)
Provider offers half day seats	0.07	(0.25)	0.12	(0.32)	0.07	(0.25)	0.06	(0.23)	0.03	(0.16)
# of pre-k seats (2017)	64.46	(52.29)	66.81	(68.92)	55.44	(37.58)	65.94	(46.63)	71.17	(57.88)
Dual language program	0.04	(0.20)	0.02	(0.15)	0.03	(0.18)	0.06	(0.25)	0.03	(0.16)
Enhanced language program	0.10	(0.30)	0.12	(0.32)	0.06	(0.23)	0.12	(0.32)	0.10	(0.30)
Income requirement	0.19	(0.39)	0.08	(0.27)	0.29	(0.45)	0.23	(0.42)	0.09	(0.29)
# of meals/day (incl. snacks)	2.43	(0.51)	2.32	(0.54)	2.55	(0.50)	2.46	(0.50)	2.36	(0.49)
Offers indoor play area	0.73	(0.44)	0.76	(0.43)	0.73	(0.44)	0.75	(0.44)	0.68	(0.47)
Offers outdoor play area	0.96	(0.19)	0.96	(0.20)	0.98	(0.15)	0.96	(0.20)	0.95	(0.22)
Offers early dropoff	0.41	(0.49)	0.48	(0.50)	0.41	(0.49)	0.35	(0.48)	0.43	(0.50)
Offers late pickup	0.45	(0.50)	0.54	(0.50)	0.45	(0.50)	0.39	(0.49)	0.48	(0.50)

Note. Authors' calculations using publicly-available data from NYC's OpenData portal. All estimates are weighted to be representative of the average child in each group in the 2017-2018 school year. Gaps in Panel B are reported as raw (unstandardized) differences between groups, and may not match reported differences in Panel A exactly due to rounding. NYCEEC - New York City Early Education Center. * $p < .05$.

Table A7 (continued). *Pre-k program characteristics and gaps by student race/ethnicity*

Panel B. Gaps in pre-k program characteristics

	black-white gap	Hispanic-white gap	Asian-white gap
DOE program	0.02 *	0.07 *	0.05 *
DOE program in K-12 school	0.08 *	0.07 *	0.02 *
DOE pre-k center	-0.05 *	0.00	0.03 *
NYCEEC program	-0.04 *	-0.07 *	-0.05 *
Provider offers full day seats	0.08 *	0.08 *	0.08 *
Provider offers half day seats	-0.05 *	-0.06 *	-0.09 *
# of pre-k seats (2017)	-11.37 *	-0.87	4.36 *
Dual language program	0.01 *	0.04 *	0.00
Enhanced language program	-0.06 *	0.00	-0.02 *
Income requirement	0.21 *	0.15 *	0.01 *
# of meals/day (incl. snacks)	0.24 *	0.14 *	0.05 *
Offers indoor play area	-0.03 *	-0.02 *	-0.08 *
Offers outdoor play area	0.02 *	0.00	-0.01 *
Offers early dropoff	-0.07 *	-0.13 *	-0.05 *
Offers late pickup	-0.09 *	-0.14 *	-0.06 *

Note. Authors' calculations using publicly-available data from NYC's OpenData portal. All estimates are weighted to be representative of the average child in each group in the 2017-2018 school year. Gaps in Panel B are reported as raw (unstandardized) differences between groups, and may not match reported differences in Panel A exactly due to rounding. NYCEEC - New York City Early Education Center. * $p < .05$.

Table A8. *Census tract level estimates of the association between tract demographics and presence of at least 1 nearby UPK provider who scored high on the CLASS*

Panel A. Providers within .25 miles

Tract level variable	Any providers with CLASS rating >=					
	50th pctile	75th pctile	90th pctile	50th pctile	75th pctile	90th pctile
Majority black	-0.05	-0.06	-0.04	0.02	-0.01	-0.05
	(0.06)	(0.06)	(0.03)	(0.07)	(0.04)	(0.03)
Majority Hispanic	-0.03	-0.01	-0.04	-0.03	-0.02	-0.05
	(0.06)	(0.06)	(0.03)	(0.06)	(0.05)	(0.04)
Majority Asian	0.02	0.02	-0.02	0.06	0.03	-0.04
	(0.07)	(0.07)	(0.09)	(0.05)	(0.07)	(0.07)
No majority	-0.03	-0.05	0.00	-0.02	-0.05	-0.01
	(0.06)	(0.06)	(0.02)	(0.04)	(0.04)	(0.04)
N	2128	2128	2128	2128	2128	2128
Borough FE	X	X	X			
District FE				X	X	X

Panel B. Providers within .5 miles

Tract level variable	Any providers with CLASS rating >=					
	50th pctile	75th pctile	90th pctile	50th pctile	75th pctile	90th pctile
Majority black	-0.02	-0.06	-0.08	0.08	0.14 *	-0.03
	(0.03)	(0.05)	(0.04)	(0.04)	(0.06)	(0.06)
Majority Hispanic	-0.03	-0.04	-0.08	0.01	0.01	-0.06
	(0.06)	(0.11)	(0.07)	(0.03)	(0.05)	(0.06)
Majority Asian	0.04	0.04	0.03	0.07	0.09	0.02
	(0.03)	(0.07)	(0.10)	(0.07)	(0.10)	(0.10)
No majority	0.01	-0.04	0.02	0.03	-0.02	0.01
	(0.02)	(0.05)	(0.05)	(0.02)	(0.03)	(0.05)
N	2128	2128	2128	2128	2128	2128
Borough FE	X	X	X			
District FE				X	X	X

Note. Authors' calculations using publicly-available data from NYC's OpenData portal. Estimates are presented as average marginal effects, and can be interpreted as the average percentage point difference between majority white tracts and each respective group. All estimates control for the total land area and population density of each census tract, as well as a composite measure of SES. CLASS - Classroom Assessment Scoring System. * $p<.05$ ** $p<.01$ *** $p<.001$.

Figure A1. *Distribution of provider ECERS ratings, by majority race/ethnicity of students*
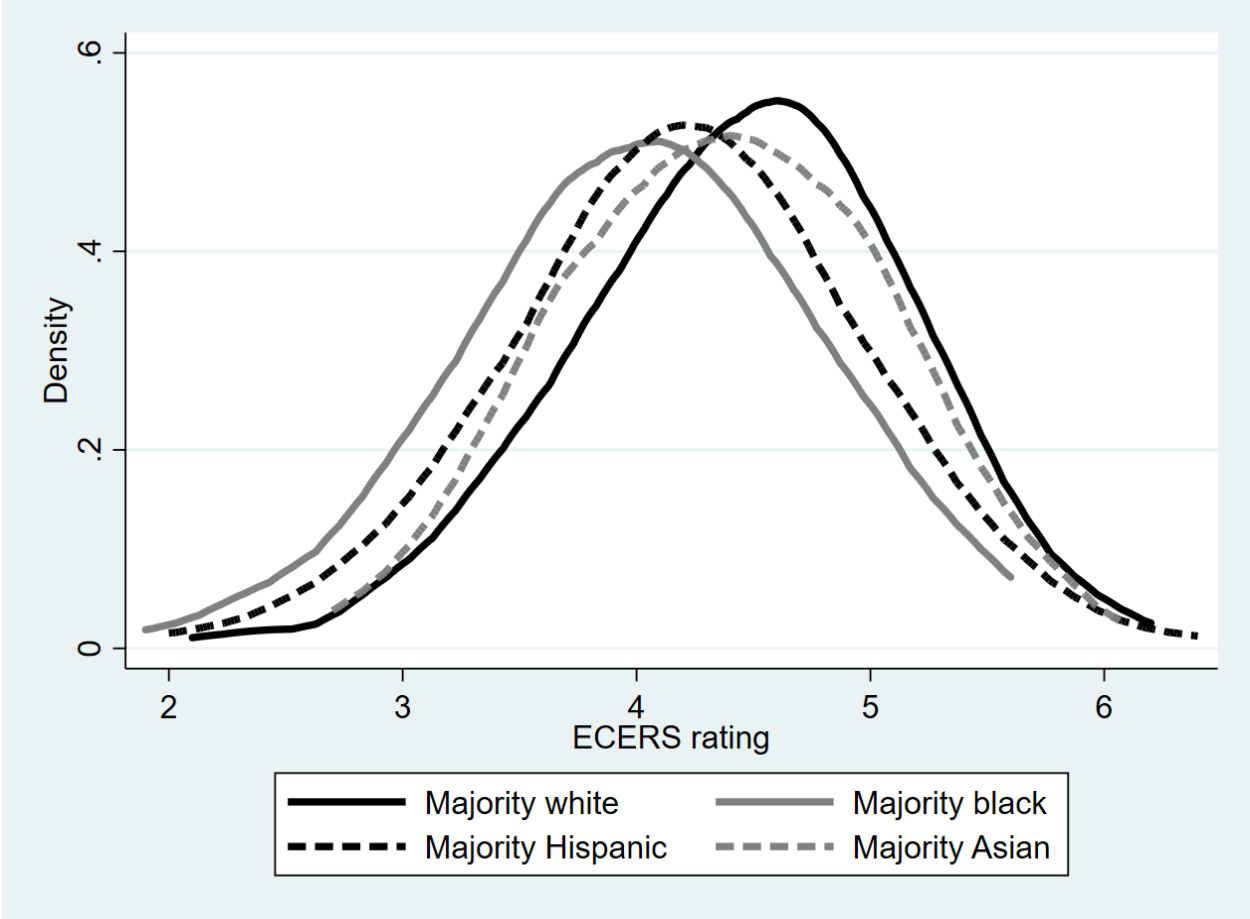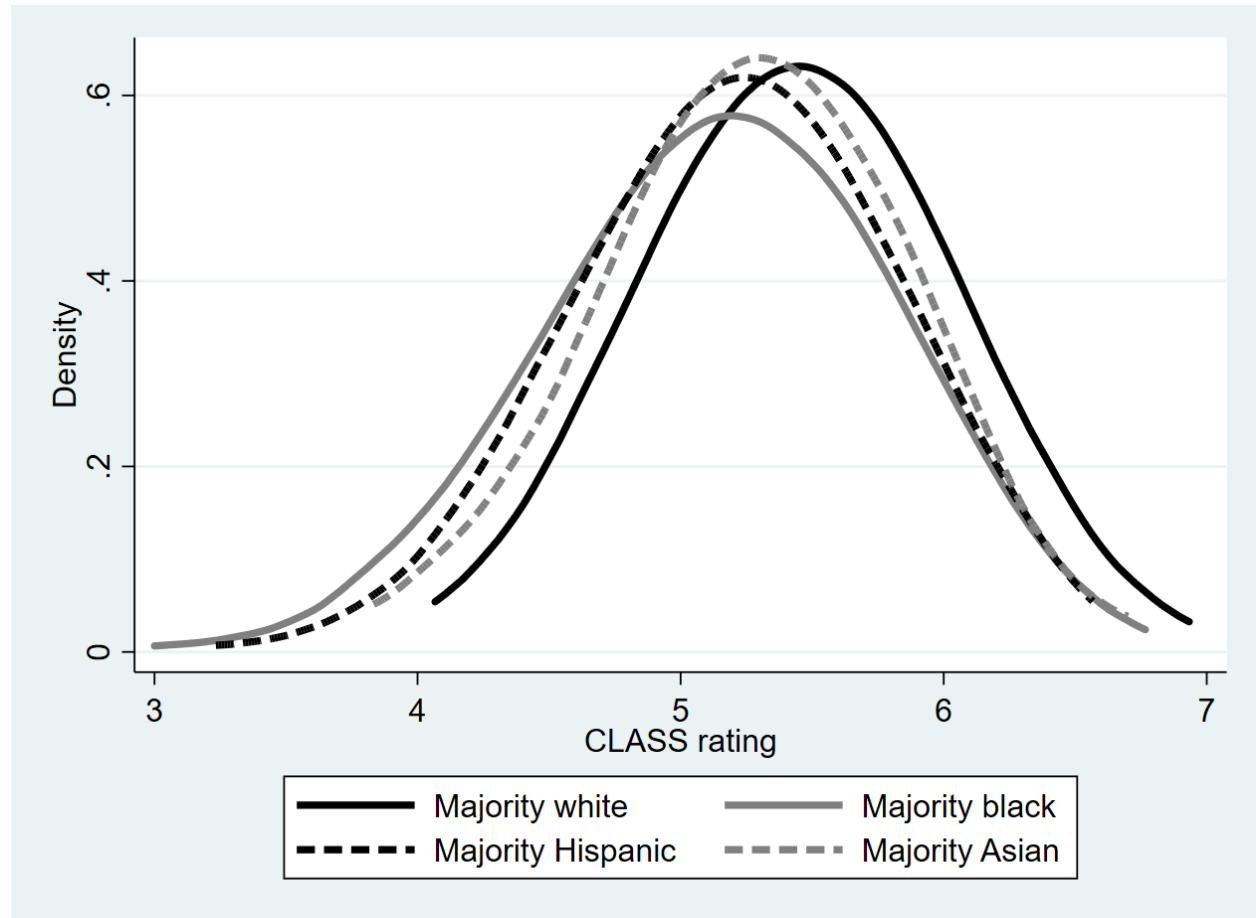
Figure A2. *Distribution of provider CLASS ratings, by majority race/ethnicity of students*

Note. Four providers out of 1545 scored between 1.8 and 3 on the CLASS, and were omitted from this figure. All four providers were majority black.

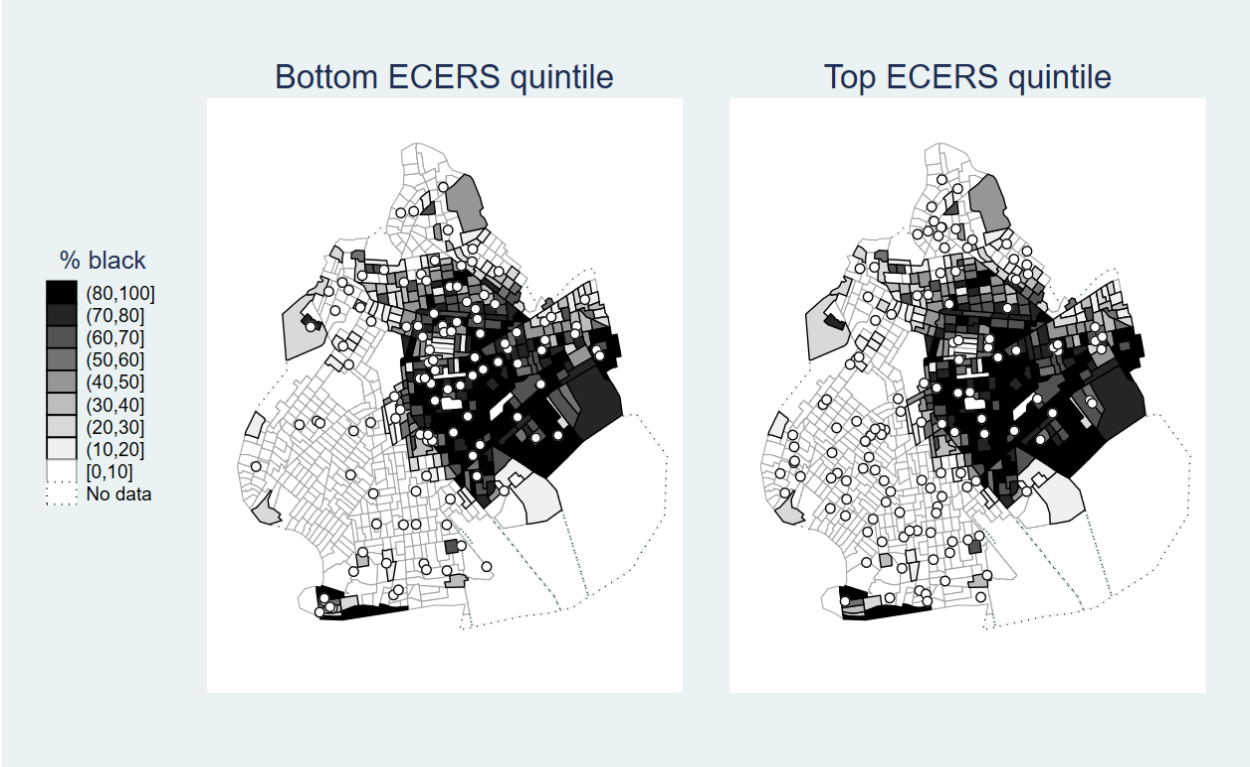**Figure A3.** *High and low-quality UPK providers in Brooklyn*

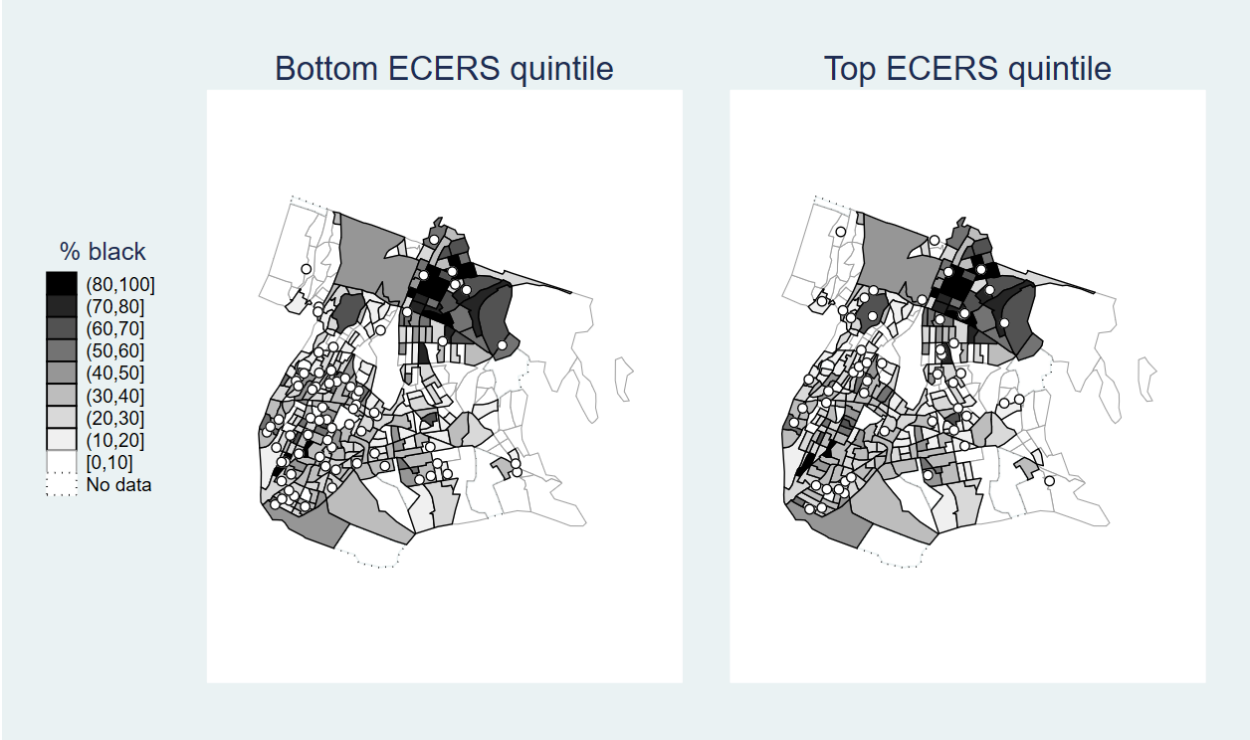**Figure A4.** *High and low-quality UPK providers in The Bronx*

Bottom ECERS quintile   Top ECERS quintile

% black
(80,100]
(70,80]
(60,70]
(50,60]
(40,50]
(30,40]
(20,30]
(10,20]
[0,10]
No data

Figure A5. *High and low-quality UPK providers in Manhattan*
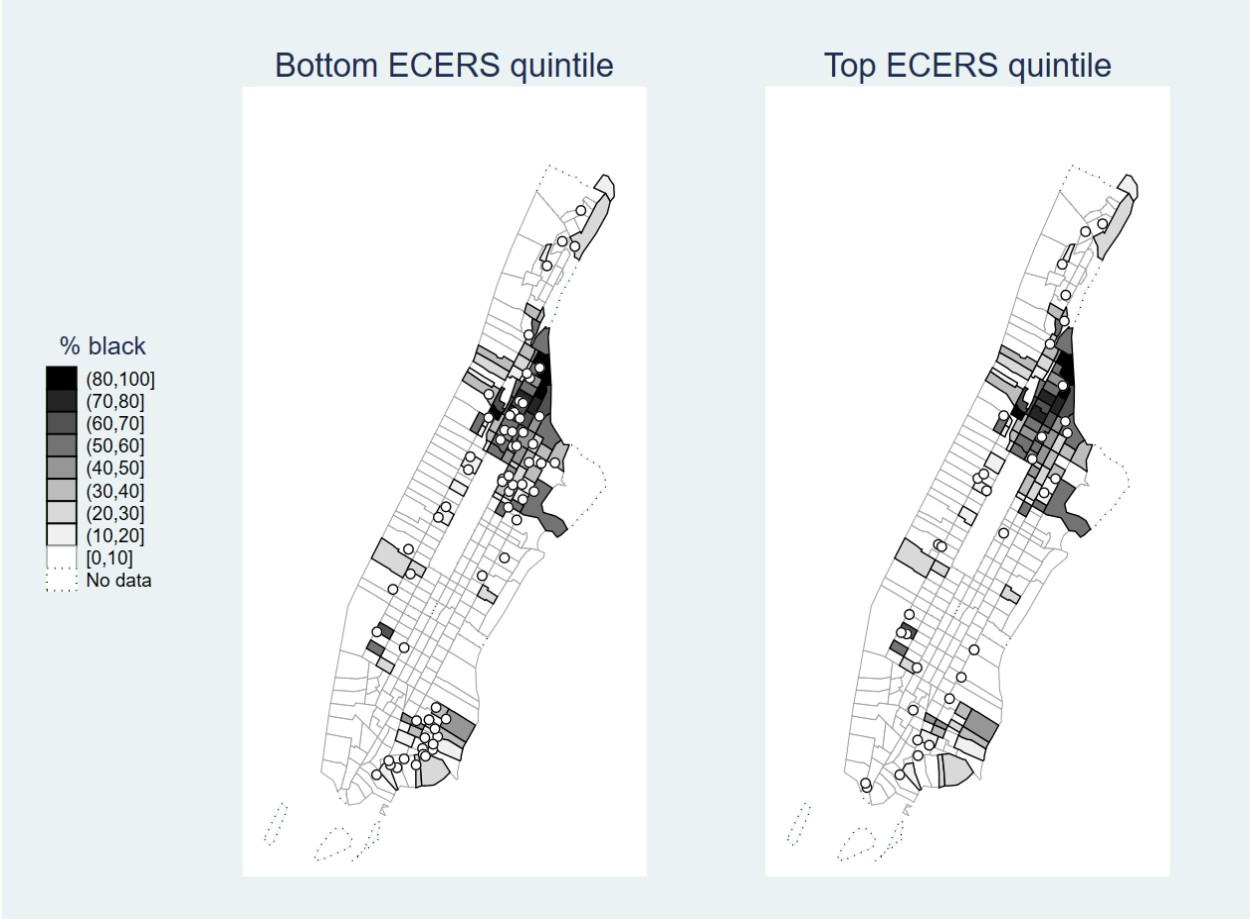
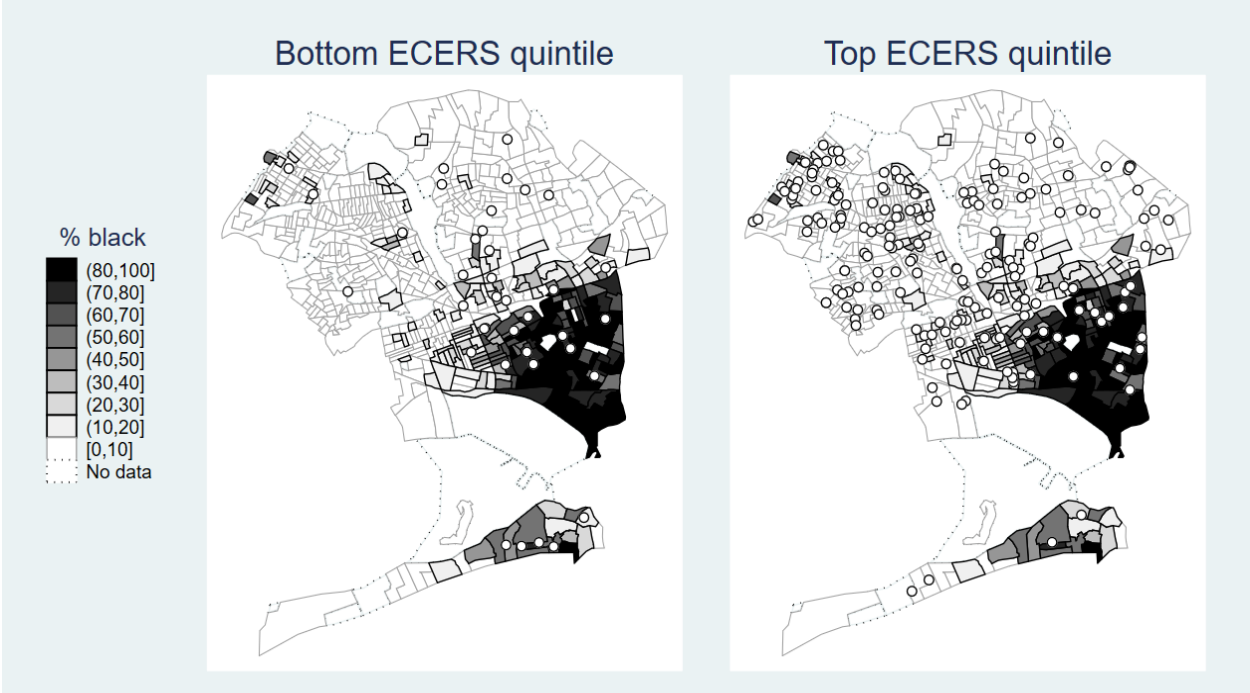**Figure A6.** *High and low-quality UPK providers in Queens*

Figure A7. *High and low-quality UPK providers in Staten Island*