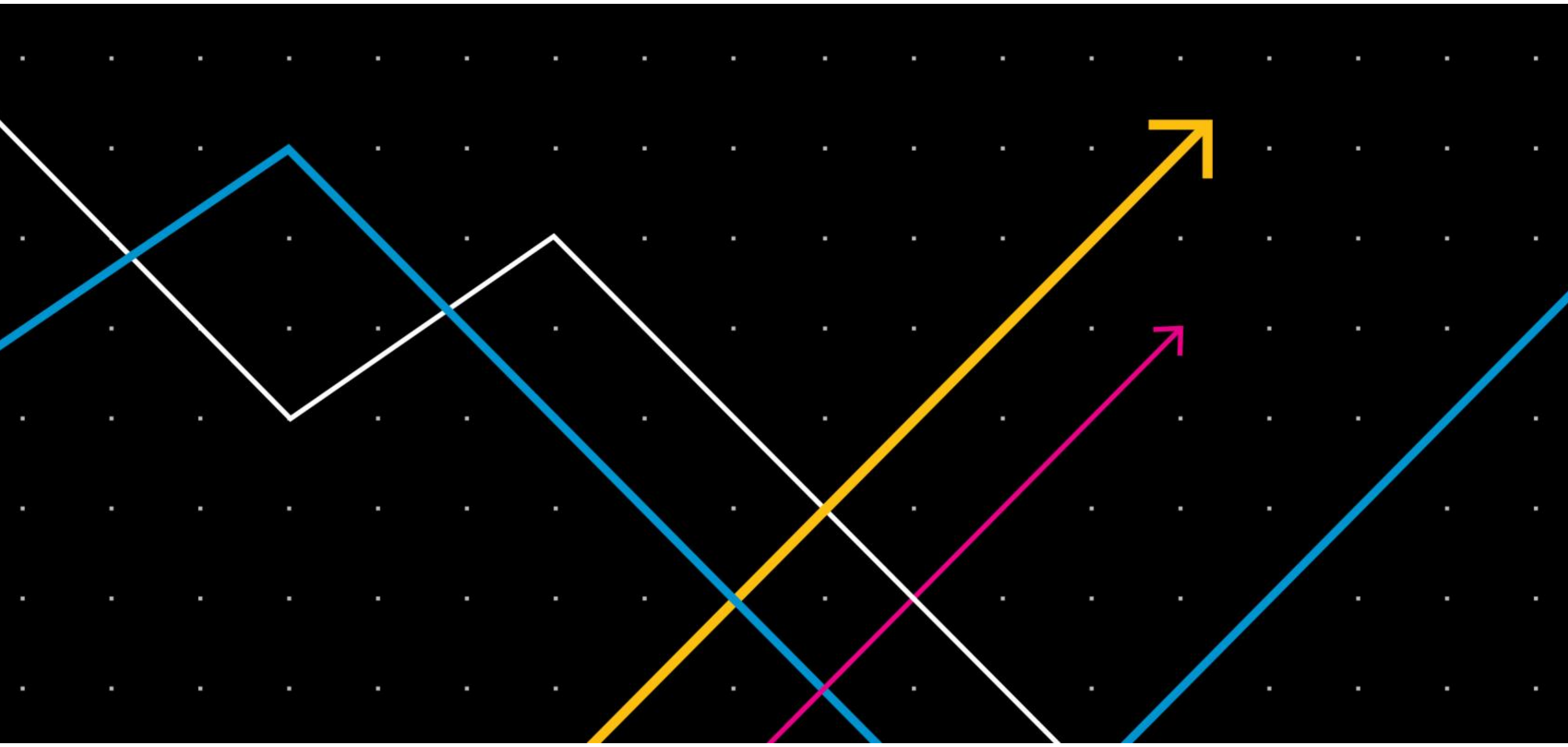


RESEARCH REPORT

Using Centralized Lotteries to Measure Preschool Impact

Insights from the DC Prekindergarten Study

Tomas Monarrez *Erica Greenberg* *Grace Luetmer* *Carina Chien*

October 2020

URBAN INSTITUTE · ELEVATE · THE · DEBATE

ABOUT THE URBAN INSTITUTE

The nonprofit Urban Institute is a leading research organization dedicated to developing evidence-based insights that improve people's lives and strengthen communities. For 50 years, Urban has been the trusted source for rigorous analysis of complex social and economic issues; strategic advice to policymakers, philanthropists, and practitioners; and new, promising ideas that expand opportunities for all. Our work inspires effective decisions that advance fairness and enhance the well-being of people and places.

Contents

Acknowledgments

This report was funded by the Heising-Simons Foundation. We are grateful to them and to all our funders, who make it possible for Urban to advance its mission.

The views expressed are those of the authors and should not be attributed to the Urban Institute, its trustees, or its funders. Funders do not determine research findings or the insights and recommendations of Urban experts. Further information on the Urban Institute's funding principles is available at urban.org/fundingprinciples.

We are grateful to Parag Pathak, Christina Weiland, and Christopher Walters for their extraordinary guidance on this report and the broader study of DC prekindergarten. Our research partners at the District of Columbia Office of the State Superintendent of Education, including Elizabeth Groginsky (now New Mexico's cabinet secretary for early childhood education), Bonnie Mackintosh (now at the US Department of Health and Human Services' Office of Planning, Research, and Evaluation), Gwen Rubinstein, Bradley Quarles, Paul Corbett, Gandhar Kothari, and Evan Kramer, and at My School DC, Catherine Peretti and Michelle Yan, helped us secure a data-sharing agreement and navigate data transfer and provided invaluable insights as the study progressed. Matthew Chingos has been an exemplary senior adviser from project conception onward. Finally, we thank Rebecca Gomez and Andrea Michel, whose support made this study possible and whose reflections enriched it substantially.

Using Centralized Lotteries to Measure Preschool Impact

A growing number of cities organize common application systems for families seeking free public preschool for their children. These cities, including Atlanta, Boston, New Orleans, New York City, and Washington, DC, differ substantially in the parameters of their systems, their preschool eligibility criteria, and the quality and availability of their preschool programs. What they share is the use of an innovative assignment mechanism—the deferred acceptance (DA) algorithm—to assign students to schools that receive more applications than they have seats (Abdulkadiroğlu et al. 2017; Gale and Shapley 1962; and Pathak 2011). The DA algorithm, in turn, provides the basis for a naturally occurring randomized experiment with the potential to revolutionize preschool impact evaluation.

Building on successful applications of DA-based assignment methods in Boston and New Orleans (Abenavoli et al. 2020; Manship, Faria, and Berg 2020; Weiland et al. 2019; Weixler, Lincove, and Valant 2020), this report describes steps taken to use a common application system and centralized lottery to study the District of Columbia's public prekindergarten program. We begin by describing lottery reconstruction and simulation, the first steps toward using lotteries in evaluation. Next, we look at simulation results and their implications for impact evaluation. We then link lottery simulation data and enrollment records to examine compliance and statistical power. We conclude by outlining how these activities set the stage for estimating the effects of public preschool.

We hope that the procedure outlined here will help democratize access to state-of-the-art program evaluation methods for prekindergarten programs that use a centralized school lottery. Research teams, in partnership with school districts and program administrators, can use the evaluation "recipe" provided here to conduct their own evaluations of educational programs. Districts with established and new prekindergarten programs alike, as well as those with new or expanded school lotteries, can use high-quality evidence to improve policymaking and practice. Below, we present the strengths and limitations of DA-based methods so researchers can decide whether and how to implement them in their own contexts.

First, however, we note that DC public prekindergarten and its centralized admissions lottery, known as My School DC, are optimal for using DA-based methods. DC public prekindergarten is a large and diverse program, enrolling 71 percent of 3-year-olds and 87 percent of 4-year-olds across the District of Columbia (Friedman-Krauss et al. 2020). The program received 5,669 applications to its program for 3-year-olds (PK3) and 3,195 applications to its program for 4-year-olds (PK4) in 2018, providing a large sample for study (My School DC 2018). Families can rank up to 12 choices on their applications (leading to oversubscription at nearly all schools), and they are eligible for various preference statuses determined by schools and local education agencies. This generates considerable variation in the likelihood that a family obtains a seat in the program (Greenberg et al. 2020). The assignment algorithm is run once, with all outcomes for all students dependent on a single random draw. These features and others detailed below illustrate the conditions under which common application systems are best suited for program evaluation.

Reconstructing and Simulating the Lottery

During the early 2000s, education leaders in Boston and New York City were struggling to design centralized school choice systems in a way that was fair and not so burdensome to families and school district staff. This caught the interest of a team renowned economists (Parag Pathak, Atila Abdulkadiroğlu, Alvin Roth, and Tayfun Sönmez), who developed a robust mechanism for centralized school choice that prevents families from gaming the system to their advantage (Pathak 2011). This school assignment mechanism, the deferred acceptance algorithm, underpins the school matching systems that dozens of school systems around the country use today.

The DA algorithm is typically called a “school lottery” because schools that are oversubscribed (i.e., schools that have more applicants than open seats) often run a random lottery to determine which students get seats. The system is “centralized” when it has a shared infrastructure that parents use to submit their ranked list of preferred schools. Another essential feature of a centralized system is that the applicant’s lottery number is common across all schools. In addition, schools (and districts) may set priorities over applicants, which are combined with lottery numbers to determine the order in which they admit applicants. Typical priority rules include sibling priorities and in-boundary resident preference.

To use the school lottery for evaluation, it is paramount to understand the underlying assignment algorithm. We summarize the DA algorithm in box 1.[1] The idea is to preliminarily match students to their most preferred school as their names are read off the applicant list. Once schools reach capacity, applicants’ priority scores are checked against the priority scores of preliminarily matched students. Priority scores are defined as a combination of the applicant’s enrollment priority ranking at the school and their lottery number. If the applicant has a better priority score than anyone in the preliminary match list, the applicant replaces the matched individual with the lowest priority score on the list. The applicant who has been rejected is placed back in the unmatched applicant list, and the school that rejected the applicant is removed from the applicant's ranked choice list.

To use the natural randomized experiment taking place in the school lottery for policy evaluation, we must understand the anatomy of the lottery to assess which students are randomized into the program and why. One approach to understanding the lottery is to replicate the student-program match list for previous lottery iterations via simulation. Box 1 provides the steps to turn the lottery inputs into the matched list. Some inputs vary from case to case, but the overall framework of the routine outlined in box 1 is common to all school systems using the DA algorithm in their school assignments.

BOX 1

The Deferred Acceptance Algorithm in Centralized School Choice

The DA algorithm uses four datasets as inputs:

1. Applicants' ranked school lists
2. Applicants' random lottery numbers (ranging from 0 to 1, with 0 being the best possible outcome)
3. Schools' open seat capacities
4. Schools' static priority rankings over applicants (e.g., in-boundary preference or sibling preference), where 1 indicates highest priority; applicants with no priority receive a large integer as a default (e.g., 11111)[a]

To begin, we compute the overall **priority score** at each school the applicant ranked, defined as the common lottery number plus their priority ranking at each school on their list. This is easiest to implement when the data are structured by application (i.e., each row identifies a unique applicant and ranked school). The priority score provides the order in which applicants are matched to seats at a school (with lower priority scores being better).

The DA algorithm can be outlined following the steps below. There are other ways of formulating the algorithm, some of which are more efficient, but we find that this characterization is easiest to understand intuitively.

Although there are students in the unmatched list, here are the steps for each student who is unmatched:

- The applicant applies to their most preferred school from which they have not been rejected, and if the applicant has listed no more schools, they are removed from the unmatched list, and the school moves to the next applicant.

- If the school is not full (it has more available seats than matched students), the school preliminarily accepts the applicant, the applicant is removed from list of unmatched applicants, and the school moves to the next applicant.
- If the school is full, it checks whether the applicant has priority over any preliminarily matched students.

    If so, the school rejects the lowest-priority preliminarily matched applicant and accepts the current applicant, and the rejected applicant is placed back on the unmatched applicant list.

    If not, the applicant is rejected and applies to the next most preferred school (and goes back to the beginning of the process).

The algorithm stops once there are no more unmatched students. The algorithm's output is a list of applicant IDs with their matched school IDs. Applicants with missing school IDs in this list are unmatched.

[a] School static priorities or preferences are those that do not change as the DA assignment algorithm is churning. Applicants' sibling crosswalks may also be needed to account for dynamic preferences, which are those that change as the algorithm iterates and are thus harder to account for.

We implemented our lottery simulation in DC for PK3 and PK4 for each lottery year from 2014 to 2018.[2] The data we acquired from the DC lottery administration was well suited for replication and simulation. The application files were ordered by an "application ID" that linked applicants to each school they ranked (in order of preference). Furthermore, the file included a column indicating the school's static priority ranking for the applicant, as well as the applicant's lottery number. This facilitated the data preparation process, as we did not have to create a mapping from student characteristics to school priorities (these vary by school in DC, and some are complicated), which would have led to measurement error. For simplicity, we use only static priority rankings in our lottery simulations.[3]

Table 1 presents summary statistics on the rate at which our simulations replicate the applicant match files provided by the DC lottery administration. For PK3 in 2018, our simulation successfully replicated 5,549 of 5,669 applications, or 97.9 percent, meaning that our algorithm successfully matched applicants to the school they were actually matched to by the DC lottery or placed on the waiting list. Out of 120 failed replications, 83 applications were matched to the wrong school, 21 were matched but wait-listed in the real data, and 16 were wait-listed in the simulation but matched in the real data. Our replication rate for PK3 and PK4 ranges from 96 percent to 98 percent for every lottery year. We consider these replication rates to be sufficiently high for policy evaluation.

TABLE 1

Replication Rate of DC Prekindergarten Lottery Matches

By year and grade

	Number of applicants in the real data	Replicated correctly in our simulation	Share correct	Total errors	Matched in the Real Data: Replicated as an unmatch in our simulation	Matched in the Real Data: Replicated as a match (to the wrong school) in our simulation	Unmatched in the Real Data: Replicated as a match in our simulation
PK3							
2014	4,250	4,168	0.981	82	12	56	14
2015	4,925	4,850	0.985	75	12	47	16
2016	5,186	5,098	0.983	88	13	60	15
2017	5,167	5,084	0.984	83	13	55	15
2018	5,669	5,549	0.979	120	21	83	16
PK4							
2014	2,506	2,406	0.960	100	30	46	24
2015	2,998	2,903	0.968	95	20	55	20
2016	3,012	2,921	0.970	91	22	50	19
2017	2,979	2,893	0.971	86	22	41	23
2018	3,195	3,074	0.962	121	28	63	30

Source: My School DC lottery data, 2014–18.

Notes: PK3 = public prekindergarten for 3-year-olds; PK4 = public prekindergarten for 4-year-olds. Lottery replications come from simulations of the lottery according to the algorithm described in box 1.

The Deferred Acceptance Propensity Score Method

Armed with a successful lottery replication code, we turn to the work of Abdulkadiroğlu and coauthors (2017) to leverage the DC prekindergarten lottery for policy evaluation. The basic idea of this research design is that the lottery creates a myriad of naturally occurring randomized experiments in school assignment. Consider a student with no priority that has ranked a school with one seat available first. Whether this student gets her top choice is entirely a function of her random lottery number. Furthermore, students that have applied for this seat and do not get it will move down their ranked list of schools and enter another lottery for another set of limited seats. Students that do not get seats in that round move down their list again and potentially enter yet another lottery, and so on. Thus, the DA mechanism generates a cascade of randomized experiments in school assignment that may affect a large share of applicants.

Importantly, what happens in the lottery for school A can affect the probability of assignment at school B, even for students that applied only to school B. For example, the chances of getting a seat for a student who applies only to school B depends in part on which schools other applicants to school B ranked higher on their list. The student will have a better chance of getting into school B if the other applicants ranked undersubscribed schools higher on their list. Thus, the assignment outcomes of school B depend on not only applicants to school B but the entire application system, making the anatomy of the lottery complex and rich in random variation in school assignment chances.

A key advantage of the method proposed by Abdulkadiroğlu and coauthors is that it seeks to leverage all random variation generated by the school lottery. The authors demonstrate that previous approaches that use the number of applicants randomized at their first-choice school ("first-choice instrument" designs) are inefficient because they do not use random variation in seats at non-first-choice schools. The sample size for these designs is necessarily smaller than the current method, leading to reduced statistical power.[4]

The main insight Abdulkadiroğlu and coauthors make is similar to that of the propensity score theorem invoked in common matching methods. In matching designs, the researcher is interested in comparing treated and untreated individuals with similar observables and runs into the "curse of dimensionality"—it is hard to find individuals in the data with the exact observables and that are both treated and untreated. In DA assignment systems, the probability of obtaining a seat is determined by applicants' "type," determined by their preferences (ranked choice lists) and priority level (which can vary by school). There is also a problem of dimensionality here: it is almost impossible to find students who were both treated and untreated with exactly the same preferences and priority levels. Abdulkadiroğlu and coauthors prove that the probability of obtaining a seat, or the *DA propensity score*, is a sufficient control to ensure that comparisons between treated and untreated groups identify causal effects.

Despite the similar names, the DA propensity score method has a higher degree of credibility than the more common propensity score method. The DA propensity score is based on the random assignment to treatment (the "gold standard" of causal inference) driven by the lottery, and the score ensures that we compare individuals with similar "risk" of being treated. In contrast, common propensity score methods rely on the strong assumption that the treatment can be assumed to be as good as randomly assigned by accounting for observable differences between treatment and control groups. For the current application (e.g., only comparing students who were matched versus unmatched in the prekindergarten lottery data without implementing the DA propensity design), this assumption would entail the untenable assertion that treatment is as good as random for applicants who ranked the same number of schools and who live in the same ward of the District of Columbia.

Abdulkadiroğlu and coauthors (2017) propose three methods for computing the DA propensity score, two of which are based on the structure of the lottery and one based on direct simulation. We opt for the simulation approach for ease of implementation and because it is easier to explain intuitively. We direct readers interested in the other two methods to the paper by Abdulkadiroğlu and coauthors (2017). In essence, computing the DA propensity score involves running the lottery repeatedly, holding constant preferences, priorities, and seat availability, allowing the lottery number to be redrawn each time the lottery is simulated. Changing nothing but the lottery number in each simulation is key, as it means that differences in student assignment are driven only by random chance. The lottery replication code we describe above is well suited for this type of simulation. It is only necessary to write a routine that executes the algorithm repeatedly with varying lottery number draws.[5]

We do this with our code and simulate the lottery 25,000 times for each grade and lottery year. The result of these simulations is a spreadsheet of student identifiers and columns denoting each school in the system. The elements of the spreadsheet count the number of times the applicant was matched to a given school over the 25,000 simulated lottery runs. For some students, all simulated rounds match to the same school (or to no school). These applicants do not face any risk in the lottery and are not subject to any randomization in school assignment. For example, a student whose first-choice school has more seats than applicants will always be matched to that school, regardless of their lottery number.

But for many students, variation in their random lottery numbers affects their school match, and the spreadsheet lists multiple schools that they get matched to across our simulations. By dividing these match counts by 25,000 (and scaling them to range between 0 and 100), we obtain our estimate of the DA propensity score for each applicant and school.

Because our main research question is not about assessing the impact of certain types of schools relative to others, but instead about the causal impact of DC prekindergarten overall, we add up the DA propensity score across all schools for each applicant, resulting in a score that measures the probability that the applicant gets matched to *any* prekindergarten program. This is the score we use in the analysis. Intuitively, applicants with similar risk of not being matched to any program are in the same stratum of the experiment. Any differences in applicant outcomes within groups with similar scores can be interpreted causally. For example, suppose that applicant A and applicant B have a 75 percent DA propensity score of being matched to a program. Whether applicant A or applicant B gets matched is entirely a function of a weighted coin flip (a 75 percent chance of being matched and a 25 percent change of not being matched). If applicant A gets a match and applicant B does not, we can compare the outcomes of applicant A and applicant B and interpret any differences causally.

This example illustrates the need to simulate the lottery as many times as possible to obtain an accurate estimate of the DA propensity score. More simulations create a more precise score. Ideally, we would like to compare observations within 1 percentage-point "buckets" of the DA propensity score, meaning we need to simulate the lottery a sufficient number of times to credibly detect the difference between, say, a 75 percent propensity score to be matched to a program versus a 76 percent propensity score. To be conservative, we simulate the lottery at least 10^4 times, which gives us score estimates precise up to four decimal places. We settled on slightly more than double that, or 25,000 lottery simulations.

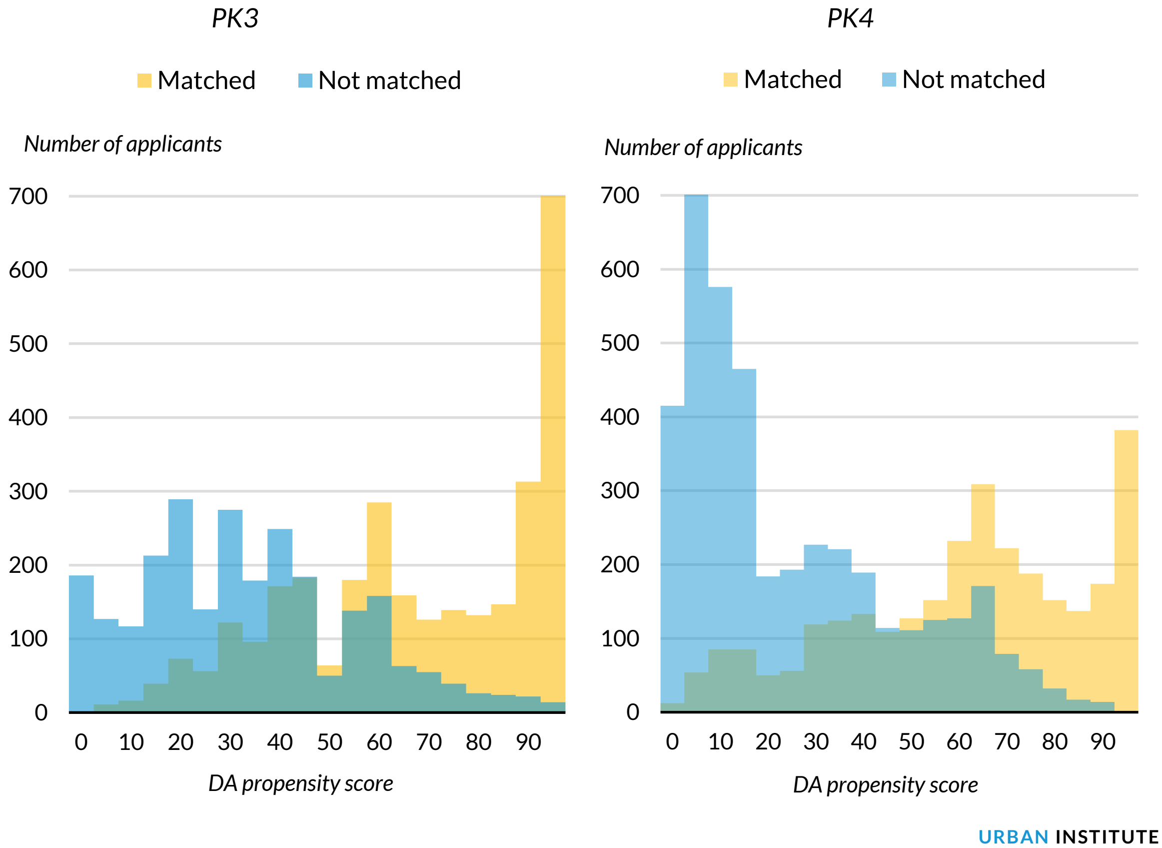
Figure 1 presents histograms of our estimated DA propensity scores for PK3 (left panel) and PK4 (right panel). Observations are pooled across lottery years and split by applicant match status. Matched applicants are shown in yellow, unmatched applicants are shown in blue, and the overlap between them is shown in green. The sample used from here on out is restricted to applicants with *nondegenerate risk* of being unmatched at the end of the lottery run (i.e., whether they are matched depends on their lottery number).[6] Applicants that get matched to a school are more likely to have higher propensity scores than those that are eventually unmatched. This is to be expected. It simply means that those with a greater chance of being matched tend to be matched in greater numbers, resulting in a distribution of the propensity score that is shifted to the right. The opposite is true for the unmatched group, whose distribution mass is shifted to the left.

There is also considerable mass at propensity levels away from the extremes. Hundreds of applicants across both grades face a significant risk of ending up unmatched, as delineated by the overlap in the histograms. For these applicants, the outcome of the lottery draw has a great impact on the likelihood of being matched. These are the observations from which we draw the bulk of our statistical power for causal evaluation. Take PK4 applicants with between a 40 percent and a 60 percent chance of obtaining a match. There are a few hundred applicants both unmatched and matched with propensity scores in this range. In essence, the DA algorithm makes comparisons between the outcomes of matched and unmatched applicants with similar propensity scores, the reason being that these applicants *flip the same weighted coin* to find out whether they will receive a match.

FIGURE 1

DA Propensity Score, by Match Status

Experimental sample, by grade, pooling lottery years



Source: My School DC lottery data, 2014–18.

Note: DA = deferred acceptance; PK3 = public prekindergarten for 3-year-olds; PK4 = public prekindergarten for 4-year-olds.

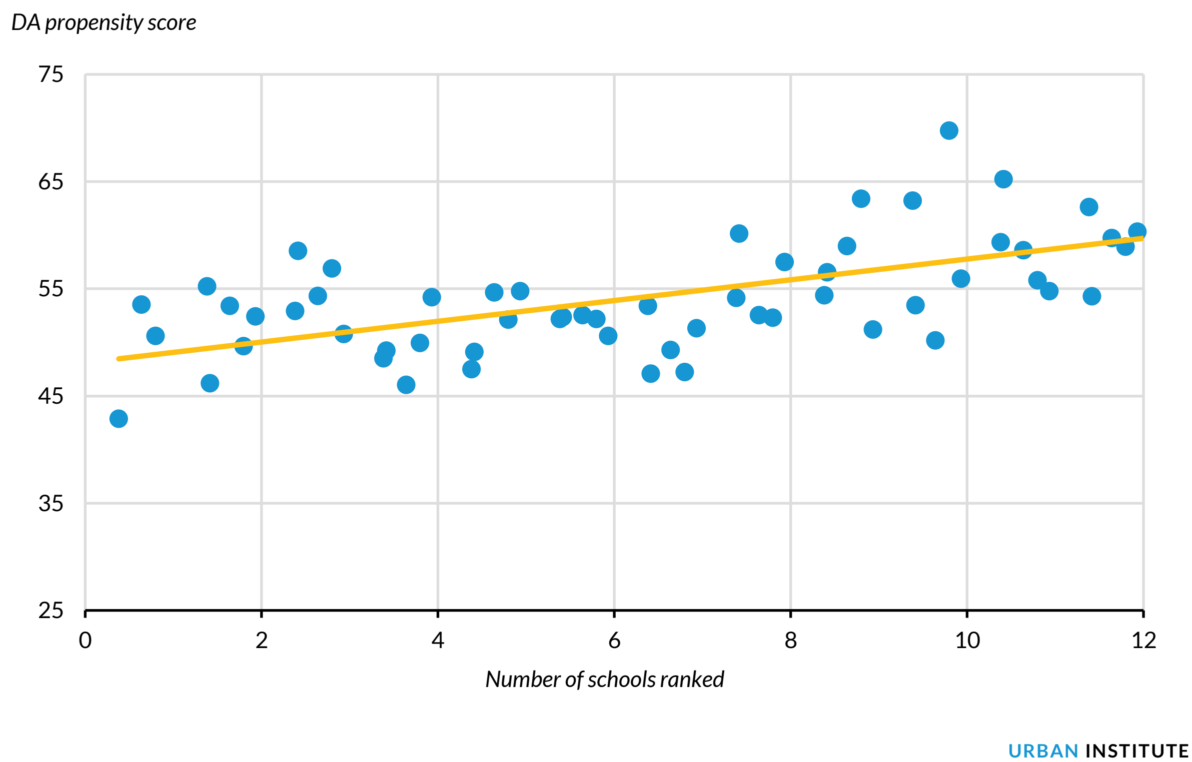
What determines an applicant's DA propensity score? The precise answer is complicated because it is a function of the combination of all applicants' preferences, their corresponding priority ranking at listed schools, and schools' capacity constraints. One way to begin unpacking the score is to correlate it with simple statistics on applicants' ranked lists. One of these is the total number of schools that applicants rank, which can be interpreted as a proxy for how "choosy" an applicant is. If the applicant lists only one school, this suggests strong preference for specific programs. If the applicant lists many schools (say, the maximum of 12), it signals that the applicant may have preference for a specific program but is more concerned with obtaining a slot somewhere.

Figure 2 shows a binned scatterplot of the PK3 DA propensity scores (vertical axis) against the number of schools ranked by applicants (horizontal axis).[7] There is a strong positive relationship between the score and the number of schools ranked. On average, listing one more school is associated with an increase in the likelihood of being matched of about 1.1 percentage points. Thus, the average difference in match probability between someone who ranked a single school and someone who ranked 12 schools is about 11.8 percentage points, which is substantial. Interestingly, these patterns do not hold for PK4, for which we find no correlation between number of schools ranked and the likelihood of a match (appendix figure A.3). The difference in this relationship across grades is likely the product of grade differences in the types of applicants that apply to oversubscribed schools, as well as different programs' priorities and capacities.

FIGURE 2

DA Propensity Score and Number of Schools Ranked

Experimental sample of PK3 applicants, pooling lottery years



Source: My School DC lottery data, 2014–18.

Notes: DA = deferred acceptance; PK3 = public prekindergarten for 3-year-olds. The horizontal axis shows 100 quantiles ("bins") of the residualized distribution of number of school ranked, after controlling for lottery year effects. The vertical axis shows average DA propensity score within these bins.

Appendix table A.2 pushes the examination of the DA propensity score one step further by estimating models of the score as a function of the number of schools ranked, as well as indicators for the applicants' ward of residence in the District of Columbia.[8] The probability of a match varies significantly between wards and by grade, reflecting that preferences and program seat availability vary significantly by ward and between entry grades (all PK3 applicants) and nonentry grades (many PK4 applicants). These patterns are relevant for understanding the impact of the lottery design on applicants of different backgrounds, as proxied by ward.

Lottery Compliance

The next step in developing our evaluation is to examine how much random variation in match status translates to random variation in program take-up. This is important because from a policy perspective, we are interested in the causal impact of enrolling in the prekindergarten program, not the impact of being matched to the program. The key issue we must grapple with is that applicants who are unmatched in the lottery can subsequently gain a seat off a waiting list when families that were matched through the lottery decide not to enroll at a given school (perhaps because they got off the waiting list at a higher-ranked school). Likewise, families who were offered a seat can decline to participate in the program (perhaps because they decide to keep their child in a private preschool or day care). In the jargon of randomized experiments, this means there is a considerable amount of *noncompliance* in the control group.

In other words, a large share of unmatched applicants eventually enroll in DC prekindergarten, which is not surprising, given the universal nature of the program. But noncompliance limits our analysis in two ways. First, it reduces the evaluation's statistical power by limiting the ability of our instrumental variable (match status) to shift the likelihood that applicants take up the program (enroll in prekindergarten). The relative severity of this problem depends largely on sample size, which is not a problem for us (as we show below).

Second, experimental noncompliance changes the nature of the causal effects we will eventually estimate. Parents' decisions to enroll in the program is endogenous to other household characteristics. This is precisely what invalidates causal inference using simple comparisons by prekindergarten enrollment in the first place. Match status is random, conditional on the DA propensity score, so we can use it to tease out random variation in enrollment provoked by the lottery. But random variation in match status changes the enrollment outcome only for *lottery compliers*, applicants who enroll because they were matched and otherwise would not have enrolled. Thus, recalling the sample restrictions above, any causal estimates we can estimate apply to the subpopulation of lottery compliers in the population of applicants facing lottery risk, not to the entire population eligible for the program. In DC, the subpopulation of lottery compliers is generally more advantaged than the population as a whole (table 2). It is important to keep this caveat in mind as we conduct the evaluation, as it may limit the external validity of our findings.

TABLE 2

Select Characteristics of Children in Washington, DC

	PK3			PK4		
	PK-eligible children	Students included in experimental sample	Students not included in experimental sample	PK-eligible children	Students included in experimental sample	Students not included in experimental sample
Child's race or ethnicity						
Black	54%	41%	61%	51%	37%	48%
Hispanic	14%	21%	14%	23%	26%	22%
White	22%	26%	16%	19%	26%	21%
Asian	4%	5%	3%	2%	3%	3%
Other race or multiracial	6%	6%	5%	6%	7%	6%
Family composition						
One parent	43%	29%	48%	45%	34%	44%
Two parents	52%	64%	47%	49%	62%	51%
No parents	5%	7%	6%	6%	5%	6%
At least one parent is an immigrant	24%	34%	23%	31%	38%	31%
Language spoken at home[a]						
English only	69%	58%	71%	64%	59%	64%
Spanish	12%	16%	11%	19%	21%	18%
Other languages	14%	19%	12%	11%	15%	12%
Disability[b]						
Parental disability	9%	6%	10%	9%	7%	8%
Child disability	0%	0%	0%	1%	1%	1%
Parental work status[c]						
One-parent household, not working	16%	10%	18%	12%	9%	12%
One-parent household, part time	9%	6%	10%	6%	4%	5%
One-parent household, full time	18%	13%	19%	28%	20%	26%
Two-parent household, not working	0%	0%	1%	0%	0%	0%
Two-parent household, one part time	1%	1%	1%	2%	2%	1%
Two-parent household, one full time	14%	18%	14%	13%	16%	14%
Two-parent household, both part time	1%	2%	1%	0%	0%	0%
Two-parent household, one part time one full time	7%	9%	6%	11%	15%	12%
Two-parent household, full time	28%	35%	24%	22%	29%	24%
Poverty, family income below 100% of FPL	24%	16%	27%	18%	12%	18%
Low income, family income below 200% of FPL	36%	25%	41%	45%	35%	43%

	PK3			PK4		
	PK-eligible children	Students included in experimental sample	Students not included in experimental sample	PK-eligible children	Students included in experimental sample	Students not included in experimental sample
Not low income, family income at least 200% of FPL	63%	75%	59%	55%	65%	57%
Food stamp recipients	37%	27%	42%	34%	25%	32%
Parents' highest educational attainment						
Some high school	12%	11%	14%	14%	12%	12%
High school diploma or some college	45%	36%	50%	45%	33%	42%
Four-year college degree or more	43%	52%	36%	40%	54%	44%
Family has access to at least one vehicle	74%	83%	71%	68%	74%	69%

Source: Estimates using 2013–17 American Community Survey Public Use Microdata Samples downloaded from IPUMS-USA.

Notes: FPL = federal poverty level; PK3 = public prekindergarten for 3-year-olds; PK4 = public prekindergarten for 4-year-olds. Percentage totals may not sum to 100 percent because of rounding and nonresponse.

[a] This variable reflects parents' primary language. If one parent speaks a non-English language, we use that language.

[b] Any reported conditions involving an individual's visual, auditory, and physical abilities are included in this variable. Individuals who have a physical, mental, or emotional difficulty that limits the ability to live alone are also included.

[c] We define this variable using caregiver work status in place of parents when there are no parents in the household. Children eligible for PK3 and PK4 are included in the table. The table does not include demographic data for children living outside, but attending prekindergarten within, the District of Columbia (195 children over five years in our study). The table does not present data from the Office of the State Superintendent of Education but uses students' lottery outcomes to describe average community characteristics of matched and wait-listed students, respectively.

Lottery compliance and risk in match status are separate issues, both of which generate selection into our analytic sample and somewhat limit the evaluation's external validity. First, we restrict the sample to applicants facing the risk of not being matched to any program. This experimental sample excludes applicants who were always going to get a seat, no matter their lottery number, and those who were never going to get a seat. Appendix table A.5 shows that excluding applicants with degenerate risk from the sample means that the experimental sample is somewhat more affluent than the population of all program-eligible children. Then, *within the experimental sample*, we have the issue of lottery compliance we describe above. Thus, our program effect estimates correspond to lottery compliers, who are a subset of applicants with nondegenerate risk, who themselves are a subset of the total population of lottery applicants and eligible children.

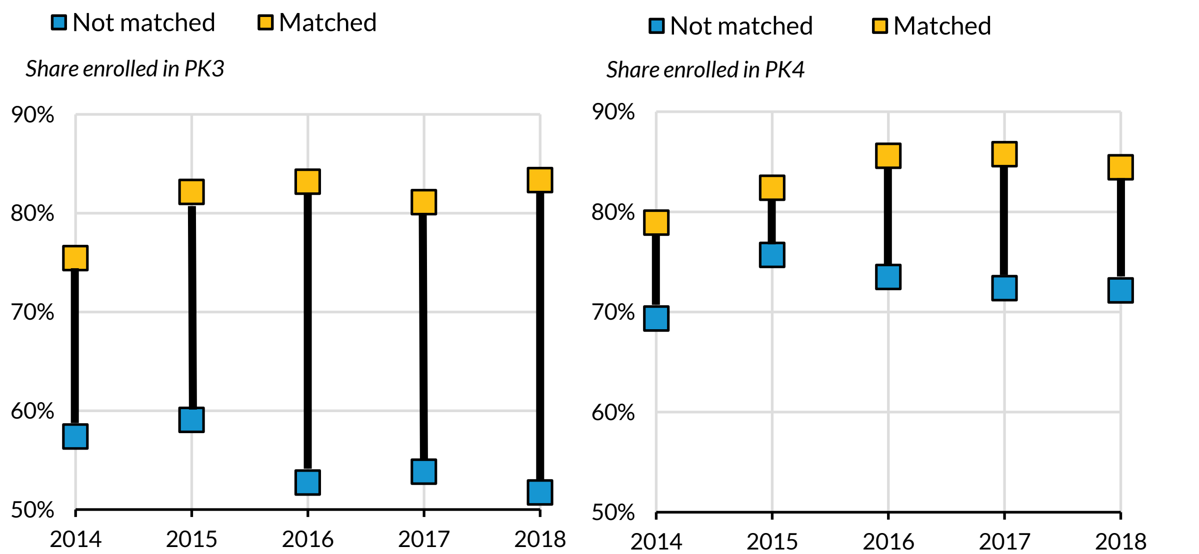
We measure enrollment using administrative student records from the Office of the State Superintendent of Education (OSSE). Our data span the same years as the lottery (2014 to 2018), and we can observe student-level enrollment files across both prekindergarten grades as well as K–3 schools. These files include information on school attended, participation in special education and English language learner programs, attendance, and other characteristics. Administrators at OSSE helped us match lottery application identifiers to student identifiers in the enrollment data using student names (which our research team cannot access), birthdays, and addresses. Our research team confirmed that this procedure does a good job of linking these two data sources, via several verification exercises.[9]

Figure 3 shows the share of our sample that enrolled in the prekindergarten program following the lottery, separated by match status, lottery year, and grade. In PK3, about 80 percent of applicants that were matched enrolled in the program, and between 50 and 60 percent of unmatched applicants eventually enrolled. This means that, depending on the year, the first-round match had an approximate *first-stage* effect of increasing the enrollment rate by 20 to 30 percentage points, the gap in enrollment rates between matched and unmatched applicants.[10] Enrollment among matched PK4 applicants looks similar to PK3 applicants, though unmatched PK4 applicants enroll at higher rates, reducing the gap in enrollment rates by match status (compared with PK3).

FIGURE 3

Enrollment Rate between Matched and Unmatched Applicants

Experimental sample, by grade and application year



Source: My School DC lottery data, 2014–18.

Notes: PK3 = public prekindergarten for 3-year-olds; PK4 = public prekindergarten for 4-year-olds. Sample restricted to applicants with a nondegenerate risk of being unmatched to a program.

Table 3 presents models estimating the first-stage effect more rigorously, using multivariate regressions controlling for the DA propensity score, year effects, and applicant characteristics. The estimates in the first two columns are simply weighted averages of the gaps reported in figure 3. For instance, the PK3 first-stage impact is about 26 percentage points before adding controls. Once we add DA propensity controls, this impact is reduced to about 19 percentage points, implying that the enrollment decision is correlated with the chance of getting a match. Once more, this pattern reinforces the notion that controlling for the DA propensity score is crucial. Still, our model has sufficient statistical power to detect this modest first-stage effect with precision.[11] Adding additional controls (for number of schools ranked and zip code indicators) does little to our first-stage estimates, which is encouraging, given that the key control variable is the DA propensity score. Similar results hold for PK4, which has a smaller first-stage effect, about 13 percentage points once we add the key controls. In the appendix, we present estimates of the first stage separately by year (appendix table A.3).

Taken together, the first-stage results establish that "winning the lottery" has a large causal impact on the probability that an applicant enrolls in DC's prekindergarten program. Because we leverage the random lottery aspect of the assignment algorithm, these results provide compelling evidence that the

DC centralized school lottery generates natural randomization into program enrollment for a subset of applicants. This means it is possible to evaluate several school-level policies using the methodology described here. In principle, it is possible to estimate the separate impact of each individual school or school grouping. We are interested in grouping all schools together and investigating how enrolling in any school affects outcomes (on average, for a selected subset of applicants).

TABLE 3

First-Stage Models: The Impact of a Match on Likelihood of Program Enrollment

Grades PK3 and PK4, pooling lottery years

	(1)	(2)	(3)	(4)
PK3				
Applicant matched	0.265*** (0.012)	0.264*** (0.012)	0.192*** (0.015)	0.193*** (0.015)
Year effects		X	X	X
DA p-score x year effects			X	X
Controls for number of schools ranked and zip code fixed effects				X
F-statistic	477.435	472.548	168.726	172.121
R^2	0.08	0.08	0.09	0.13
N	5,631	5,631	5,631	5,604
PK4				
Applicant matched	0.108*** (0.010)	0.109*** (0.010)	0.129*** (0.013)	0.126*** (0.012)
Year effects		X	X	X
DA p-score x year effects			X	X
Controls for number of schools ranked and zip code fixed effects				X
F-statistics	121.386	122.685	104.018	102.483
R^2	0.02	0.02	0.02	0.07
N	6,944	6,944	6,944	6,905

Sources: My School DC lottery data and Office of the State Superintendent of Education enrollment data, 2014–18.

Notes: DA = deferred acceptance; PK3 = public prekindergarten for 3-year-olds; PK4 = public prekindergarten for 4-year-olds. This table uses ordinary least squares models. The outcome is enrollment in the prekindergarten program. Robust standard errors apply in all models. The f-statistic corresponds to the test of significance of the match status indicator.

*** $p < 0.01$.

Measuring Impact on Student Outcomes

The DA propensity score method is essentially a complex randomized controlled trial. As is typical in randomized controlled trial studies, we begin the impact analysis by conducting a balance test showing that treatment and control units look similar at baseline. Table 4 presents estimates of regression models of the lottery match (the "treatment") on applicants' key characteristics: the number of schools listed and the location of their residence (proxied by ward indicators). These models restrict attention to applicants reporting residence in the District of Columbia.[12] Columns 1 and 2 for PK3 (and columns 4 and 5 for PK4) show that applicant characteristics are significantly correlated with the likelihood of being matched. For instance, PK3 applicants residing in Ward 3 are 18.5 percentage points less likely to be matched than Ward 1 applicants. Controlling for the lottery year does little to remedy this link. This establishes that comparing outcomes by applicant match status would not be an “apples-to-apples” comparison. It would effectively entail comparisons between applicants with different preferences and living in different parts of the city. It would require an untenable mental leap to attribute causality to such comparisons.

Columns 3 and 6 demonstrate that controlling for the DA propensity score eliminates applicant differences in the likelihood of being matched. Because different lottery years differ along many dimensions (including total applicant number and composition, seat availability, and features of the assignment algorithm), it is necessary to make sure DA propensity controls are implemented separately by lottery year. These models therefore control for the DA propensity score interacted with lottery year indicators. In essence, this restricts comparisons between applicants in the same lottery year and with similar propensity to being matched to a program. The propensity is a function of applicants’ characteristics and their ranked list of schools. Notably, none of the coefficients in these models is statistically different from zero. Furthermore, the point estimates for these coefficients are of a very small magnitude relative to the other models, nearly zero in most cases.

In sum, the DA propensity controls have done their jobs. This is exactly what we would expect to see if match status is random, conditional on the DA propensity score. If match status is truly random, it should be uncorrelated with applicant characteristics, as well as any other observed or unobserved applicant attributes that could, in principle, be correlated with student outcomes. We cannot test whether match status is uncorrelated with unobserved attributes such as socioeconomic status, but this evidence strongly suggests this is the case. (As in table 2, additional controls do not change these findings. For brevity, we omit models with additional controls.)

TABLE 4

Balance Test: Probability of a Match as a Function of Applicant Characteristics

Grades PK3 and PK4, pooling lottery years

	PK3			PK4		
	(1)	(2)	(3)	(4)	(5)	(6)
Number of schools ranked	0.012***	0.013***	-0.000	0.002	0.002	0.001
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)
Ward 2	-0.036	-0.033	-0.019	-0.020	-0.022	-0.010
	(0.031)	(0.031)	(0.024)	(0.036)	(0.036)	(0.031)
Ward 3	-0.185***	-0.180***	0.001	0.202***	0.203***	-0.001
	(0.029)	(0.029)	(0.026)	(0.022)	(0.022)	(0.019)
Ward 4	0.039*	0.046**	0.003	0.042*	0.044**	-0.016
	(0.021)	(0.021)	(0.017)	(0.022)	(0.022)	(0.018)
Ward 5	-0.136***	-0.130***	-0.029	-0.046*	-0.043*	0.012
	(0.025)	(0.025)	(0.021)	(0.023)	(0.023)	(0.019)
Ward 6	-0.049**	-0.045**	-0.011	-0.053**	-0.051**	-0.010
	(0.022)	(0.022)	(0.018)	(0.023)	(0.023)	(0.018)
Ward 7	0.033	0.033	-0.018	0.074***	0.074***	-0.000
	(0.029)	(0.029)	(0.023)	(0.027)	(0.027)	(0.021)
Ward 8	0.095***	0.100***	0.017	0.209***	0.206***	0.007
	(0.032)	(0.032)	(0.025)	(0.027)	(0.027)	(0.021)
Year effects		X	X		X	X
DA p-score x year effects			X			X
Joint f-statistic	21.504	22.465	0.697	36.38	35.789	0.522
R^2	0.03	0.03	0.36	0.04	0.04	0.37
N	5,592	5,592	5,592	6,883	6,883	6,883

Sources: My School DC lottery data and Office of the State Superintendent of Education enrollment data, 2014–18.

Notes: DA = deferred acceptance; PK3 = public prekindergarten for 3-year-olds; PK4 = public prekindergarten for 4-year-olds. This table uses ordinary least squares models. Robust standard errors apply in all models. The outcome is an indicator for program match status. The sample is restricted to applicants reporting residence in the District of Columbia. The joint f-statistic corresponds to a test of joint significance of the coefficients on ward indicators and number of schools ranked.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

We are now ready to estimate causal impacts of prekindergarten program enrollment on student outcomes. With currently available data, we have a limited set of outcomes to explore. We cannot look at student achievement in this setting (students are not given standardized assessments until third grade), but we can examine whether prekindergarten enrollment has an impact on the likelihood that students remain in District public schools (including the District of Columbia Public Schools, or DCPS, system and public charter schools) for elementary school grades. This outcome is important in its own right, especially when considering interest in keeping affluent students in public schools (Weiland et al. 2019).

Table 5 presents program impact estimates for two outcomes associated with persistence in DC public schools: enrollment in the following grade and continuous enrollment in the system (for as long as we can tell with the available data). The first panel shows results for PK3. Being matched to the program causes a 5 percentage-point increase in the likelihood that the student enrolls in PK4 the following year, off of a baseline of 55 percent in the unmatched group. This model is known as the *reduced form* (RF) or *intent-to-treat* (ITT) model because it computes the impact of being assigned to the program, not the impact of actually taking it up.

To measure the impact of program enrollment, we must compute the *instrumental variables* (IV) estimate, which effectively scales the RF effect by dividing it by the magnitude of the first-stage effect (the compliance models reported in table 2). For our PK3 models, this entails dividing the RF/ITT estimates by 0.2. The IV estimate can be interpreted as a *local average treatment effect* (LATE) because the estimates apply to lottery compliers, as discussed above (lottery compliers are a subset of applicants with nondegenerate risk). Column 2 shows that the IV/LATE effect of enrolling in PK3 is to increase the likelihood of enrolling in PK4 the next year by 30 percentage points, or 55 percent relative to the baseline.

Similar findings hold when the outcome is continuous enrollment. Continuity is defined as the student being enrolled in every grade we could observe given their lottery year and the constraints in our data. For 2014 PK3 applicants, continuity entails enrolling in PK4 in 2015, kindergarten in 2016, first grade in 2017, and second grade in 2018. For other cohorts, an equivalent definition applies. As such, column 4 of table 5 shows that enrollment in PK3 leads to a 28.5 percentage-point increase in the likelihood of enrollment continuity off a baseline of 48 percent.

Estimated impacts for PK4 enrollment are of a smaller magnitude compared with those for PK3, in part because the first-stage effect of a lottery match is smaller (as discussed above). We can reject the hypothesis that PK4 enrollment has no impact on the likelihood of persistence into public kindergarten. Our estimates indicate that PK4 enrollment leads to a statistically significant 24 percentage-point increase in the probability that the student enrolls in public kindergarten the following year. Additionally, the instrumental variable models suggest that PK4 enrollment leads to a 19 percentage-point increase in the likelihood of continuous public school enrollment, though the estimates lack the precision necessary to reject that this effect is zero.

The magnitude and precision of these impacts provide compelling evidence that enrollment in the city's public prekindergarten program has a large effect on the persistence of families in public schools. The DA algorithm and steps for reconstructing and simulating the lottery (accounting for compliance) and for estimating instrumental variables models provide confidence that these impacts are causal. By leveraging the centralized school assignment lottery in the District of Columbia, we can generate new insights on DC prekindergarten and inform efforts to deliver effective early education programs around the country.

TABLE 5

Impact of the DC Prekindergarten Program on the Probability of Enrolling in Public Schools in Grades K–3

Grades PK3 and PK4, pooling lottery years

	Enrolls in PK4		Continuous Enrollment	
	(1) RF/ITT	(2) IV/LATE	(3) RF/ITT	(4) IV/LATE
PK3				
Matched	0.053*** (0.016)		0.049*** (0.018)	
Enrolls in program		0.304*** (0.079)		0.285*** (0.094)
Lottery year fixed effects	X	X	X	X
DA p-score x year fixed effects	X	X	X	X
Controls for number of schools ranked and zip code fixed effects	X	X	X	X
N	4,389	4,389	4,389	4,389

	Enrolls in Kindergarten		Continuous Enrollment	
	(1) RF/ITT	(2) IV/LATE	(3) RF/ITT	(4) IV/LATE
PK4				
Matched	0.029** (0.014)		0.022 (0.016)	
Enrolls in program		0.239** (0.109)		0.186 (0.126)
Lottery year fixed effects	X	X	X	X
DA p-score x year fixed effects	X	X	X	X
Controls for number of schools ranked and zip code fixed effects	X	X	X	X
N	5,335	5,335	5,335	5,335

Sources: My School DC lottery data and Office of the State Superintendent of Education enrollment data, 2014–18.

Notes: DA = deferred acceptance; IV/LATE = instrumental variables or local average treatment effect; PK3 = public prekindergarten for 3-year-olds; PK4 = public prekindergarten for 4-year-olds; RF/ITT = reduced form or intent-to-treat model. Robust standard errors apply in all models. Columns 1 and 3 use ordinary least squares models. The instrumental variable in the two-stage least squares models in columns 2 and 4 is the match status indicators.

** $p < 0.05$; *** $p < 0.01$.

Conclusion

This report describes how to use a common application system and centralized lottery to conduct applied policy research on public preschool. It follows a report on lottery applicants and application patterns in DC (Greenberg et al. 2020) and will be followed by an expanded set of impact findings, building on the persistence effects identified here. We offer step-by-step descriptions here as a resource for researchers partnering with a growing number of cities that organize lotteries similar to the one in the District of Columbia. We hope this report can lower the barrier to entry to using these complex methods to improve the quality of evidence that guides early education policy and practice.

Appendix

TABLE A.1

Summary Statistics of Experimental Analysis Sample

Grades PK3 and PK4, pooling lottery years

	PK3				PK4			
	Matched		Unmatched		Matched		Unmatched	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Enrolled in prekindergarten	0.81	(0.39)	0.55	(0.50)	0.83	(0.37)	0.73	(0.45)
DA propensity score (%)	70.50	(24.85)	35.37	(21.66)	63.12	(25.62)	26.31	(22.05)
Number of schools ranked	6.51	(4.08)	5.86	(3.92)	5.35	(3.83)	5.50	(3.78)
Enrolls in PK4	0.66	(0.47)	0.55	(0.50)	0.84	(0.37)	0.73	(0.45)
Enrolls in kindergarten	0.46	(0.50)	0.41	(0.49)	0.63	(0.48)	0.59	(0.49)
Enrolls continuously	0.57	(0.49)	0.48	(0.50)	0.54	(0.50)	0.52	(0.50)
Ever in special ed.	0.10	(0.30)	0.08	(0.28)	0.09	(0.29)	0.09	(0.29)
Continuously in special ed.	0.03	(0.17)	0.02	(0.15)	0.03	(0.17)	0.03	(0.16)
Ever retained	0.03	(0.17)	0.03	(0.16)	0.03	(0.16)	0.02	(0.14)
Lottery year								
2014	0.17	(0.38)	0.16	(0.37)	0.18	(0.38)	0.16	(0.36)
2015	0.24	(0.43)	0.19	(0.40)	0.22	(0.41)	0.19	(0.39)
2016	0.18	(0.38)	0.19	(0.39)	0.20	(0.40)	0.18	(0.39)
2017	0.21	(0.41)	0.22	(0.41)	0.20	(0.40)	0.23	(0.42)
2018	0.20	(0.40)	0.24	(0.43)	0.21	(0.41)	0.24	(0.43)
Ward of residence								
Ward 1	0.17	(0.38)	0.15	(0.36)	0.09	(0.29)	0.12	(0.33)
Ward 2	0.06	(0.24)	0.07	(0.25)	0.03	(0.16)	0.04	(0.19)
Ward 3	0.05	(0.22)	0.09	(0.29)	0.26	(0.44)	0.15	(0.36)
Ward 4	0.26	(0.44)	0.20	(0.40)	0.20	(0.40)	0.21	(0.41)
Ward 5	0.10	(0.30)	0.16	(0.36)	0.10	(0.30)	0.15	(0.36)
Ward 6	0.20	(0.40)	0.21	(0.41)	0.11	(0.31)	0.18	(0.38)
Ward 7	0.08	(0.28)	0.07	(0.26)	0.09	(0.28)	0.08	(0.28)
Ward 8	0.07	(0.25)	0.05	(0.21)	0.13	(0.33)	0.07	(0.25)
Observations	3,083		2,548		2,902		4,042	

Sources: My School DC lottery data and Office of the State Superintendent of Education enrollment data, 2014–18.
Notes: DA = deferred acceptance; PK3 = public prekindergarten for 3-year-olds; PK4 = public prekindergarten for 4-year-olds; SD = standard deviation. Experimental sample consists of applicants that face nondegenerate risk.

TABLE A.2

Correlates of the DA Propensity Score

Grades PK3 and PK4, pooling lottery years

	(1) PK3	(2) PK4
Number of schools ranked	1.265***	0.148
	(0.096)	(0.092)
Ward 1	0.000	0.000
	(.)	(.)
Ward 2	-1.888	-0.451
	(1.848)	(1.955)
Ward 3	-17.763***	20.199***
	(1.523)	(1.202)
Ward 4	4.258***	5.986***
	(1.238)	(1.262)
Ward 5	-9.869***	-5.454***
	(1.379)	(1.366)
Ward 6	-3.389***	-4.015***
	(1.194)	(1.320)
Ward 7	5.069***	7.292***
	(1.720)	(1.637)
Ward 8	8.195***	19.714***
	(1.963)	(1.717)
2014 lottery	0.000	0.000
	(.)	(.)
2015 lottery	4.613***	0.797
	(1.203)	(1.131)
2016 lottery	-2.087*	-1.244
	(1.267)	(1.155)
2017 lottery	-1.360	-5.379***
	(1.226)	(1.083)
2018 lottery	-5.503***	-4.451***
	(1.182)	(1.111)
Constant	49.081***	36.775***
	(1.387)	(1.420)
R^2	0.08	0.11
N	5,581	6,871

Source: My School DC lottery data and Office of the State Superintendent for Education enrollment data, 2014–18.
Notes: DA = deferred acceptance; PK3 = public preschool for 3-year-olds; PK4 = public preschool for 4-year-olds. Robust standard errors apply in all models.
*** $p < 0.01$.

TABLE A.3

First-Stage Models, by Lottery Year

Grades PK3 and PK4, experimental applicant sample

	PK3				
	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018
Applicant matched	0.121*** (0.037)	0.139*** (0.033)	0.241*** (0.034)	0.200*** (0.033)	0.254*** (0.029)
DA propensity score (%)	0.002** (0.001)	0.001*** (0.001)	0.001* (0.001)	0.002*** (0.001)	0.001* (0.001)
Number of schools ranked	0.014*** (0.004)	0.008*** (0.003)	0.004 (0.003)	0.004 (0.003)	0.002 (0.003)
Constant	0.452*** (0.040)	0.511*** (0.033)	0.478*** (0.038)	0.457*** (0.036)	0.488*** (0.033)
Zip code fixed effects	X	X	X	X	X
F-statistic	10.882	18.082	48.806	36.424	75.18
R^2	0.10	0.14	0.18	0.14	0.18
N	938	1,239	1,009	1,192	1,211

	PK4				
	(1) 2014	(2) 2015	(3) 2016	(4) 2017	(5) 2018
Applicant matched	0.134*** (0.032)	0.109*** (0.026)	0.115*** (0.029)	0.130*** (0.026)	0.145*** (0.027)
DA propensity score (%)	-0.000 (0.001)	-0.000 (0.000)	0.000 (0.000)	0.001 (0.000)	-0.000 (0.000)
Number of schools ranked	0.002 (0.003)	-0.010*** (0.003)	-0.003 (0.003)	-0.010*** (0.003)	-0.000 (0.003)
Constant	0.668*** (0.030)	0.806*** (0.024)	0.736*** (0.026)	0.759*** (0.024)	0.718*** (0.022)
Zip code fixed effects	X	X	X	X	X
F-statistic	17.676	17.191	15.781	24.386	28.388
R^2	0.07	0.07	0.07	0.08	0.10
N	1,138	1,392	1,303	1,498	1,562

Source: My School DC lottery data and Office of the State Superintendent for Education enrollment data, 2014–18.

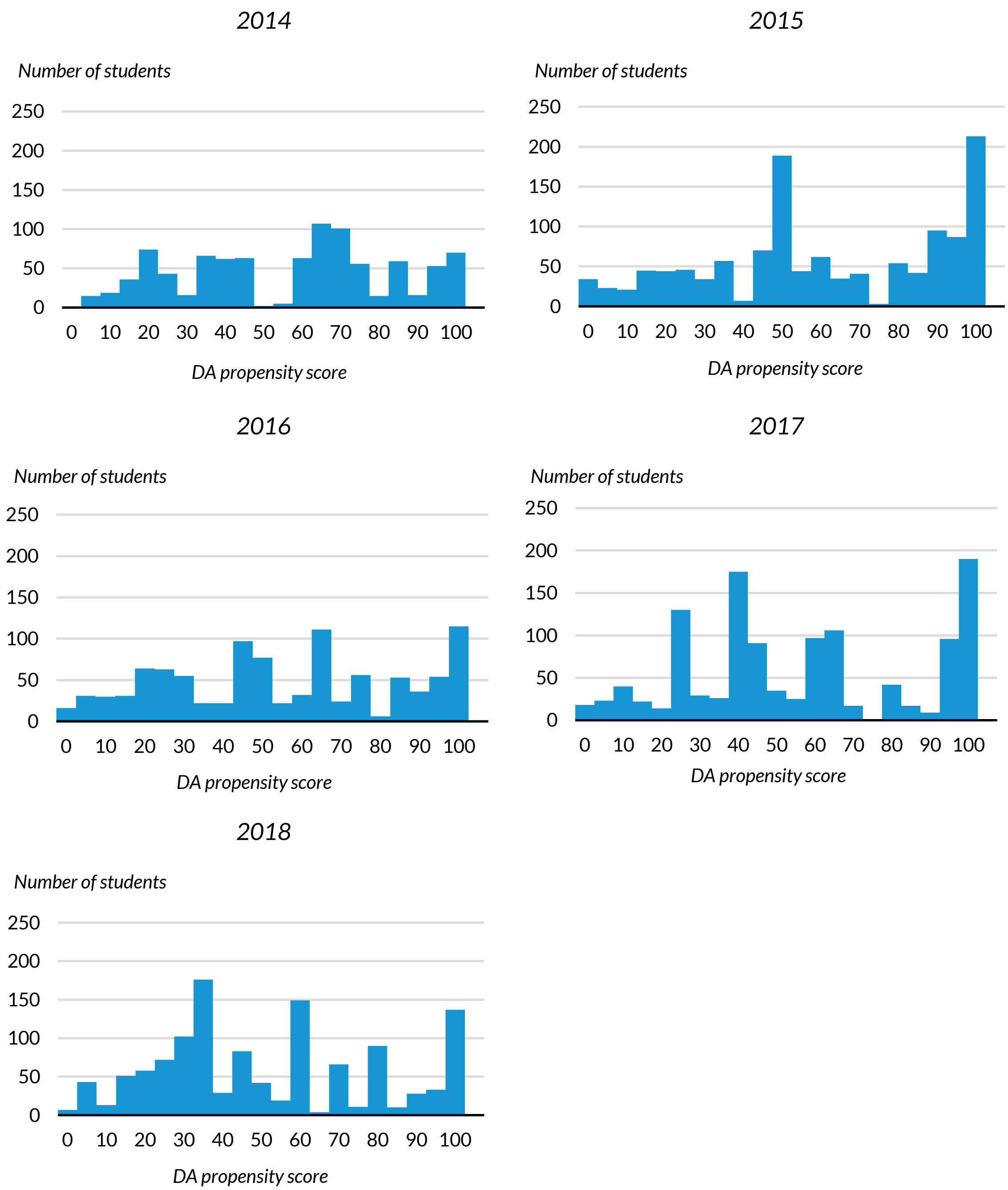
Notes: DA = deferred acceptance; PK3 = public preschool for 3-year-olds; PK4 = public preschool for 4-year-olds Robust standard errors apply in all models.

* p < 0.1; ** p < 0.05; *** p < 0.01.

FIGURE A.1

Distribution of the DA Propensity Score

PK3 applicants, by year



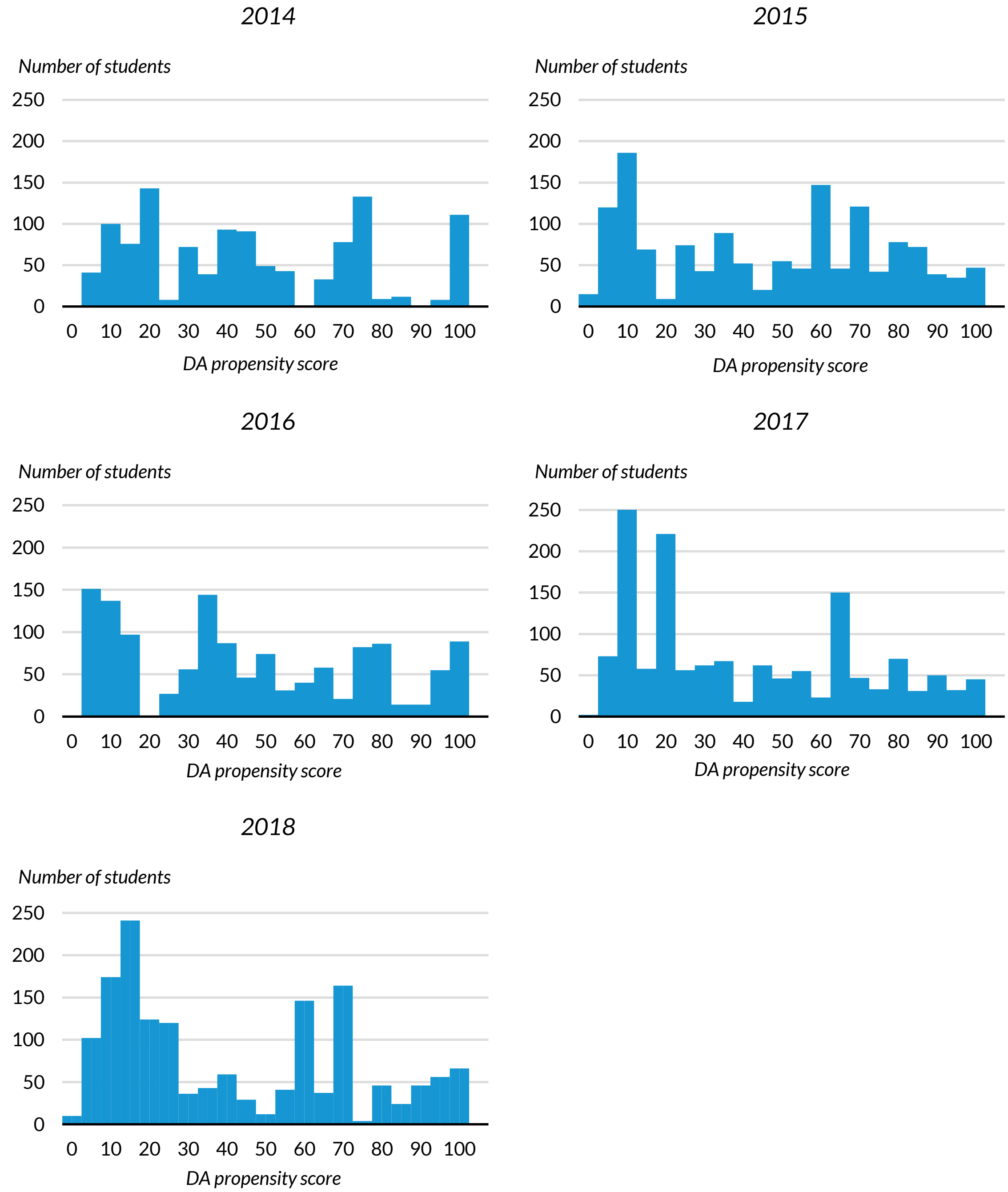
URBAN INSTITUTE

Sources: My School DC lottery data and Office of the State Superintendent of Education enrollment data, 2014–18.
Note: DA = deferred acceptance; PK3 = public prekindergarten for 3-year-olds.

FIGURE A.2

Distribution of the DA Propensity Score

PK4 applicants, by year



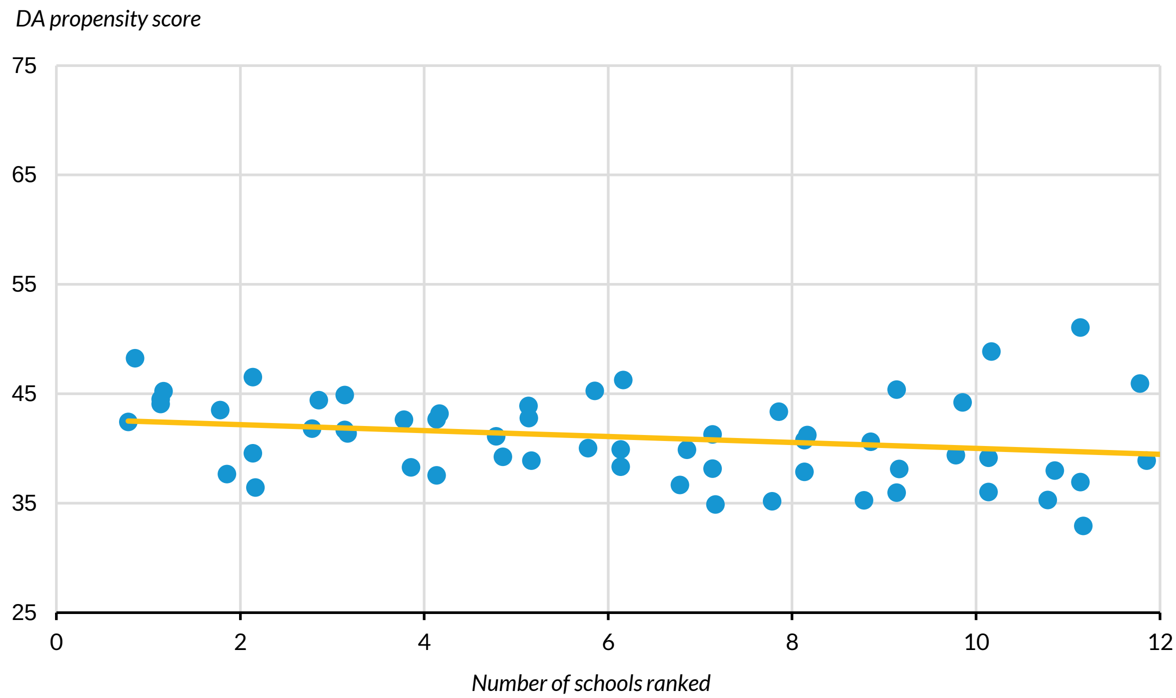
URBAN INSTITUTE

Sources: My School DC lottery data and Office of the State Superintendent of Education enrollment data, 2014–18.
Note: DA = deferred acceptance; PK4 = public prekindergarten for 4-year-olds.

FIGURE A.3

DA Propensity Score and Number of Schools Ranked

PK4 applicants, pooling lottery years



URBAN INSTITUTE

Sources: My School DC lottery data and Office of the State Superintendent of Education enrollment data, 2014–18.
Note: DA = deferred acceptance; PK4 = public prekindergarten for 4-year-olds.

Notes

[1] The basic logic of this algorithm stems from the seminal work of Gale and Shapley (1962) for the solution to the "stable marriage" problem.

[2] We also recommend that readers who are interested in replicating a lottery review easily accessible packages for the Gale-Shapley algorithm, available for most common platforms, including R, Python, and MATLAB. The key difference between the off-the-shelf versions of the Gale-Shapley algorithm and our code is the importance of school capacities for the lottery. Gale-Shapley solves a "marriage problem," in which there is a one-to-one match, but the school lottery matches many applicants to a single school (up to the school's capacity).

[3] There are also school priorities that are updated during the algorithm routine, affecting a small share of applications (about 2 percent). We do not try to model these, as we do not have the necessary data to simulate these edge cases effectively at our disposal. The priorities in the DC lottery (My School DC refers to these as "school preferences") that are not accounted for in our simulation include sibling-related preferences and the "guarantee preference." We do not have the data at our disposal to simulate these correctly. The former requires data on enrollment and lottery applications of prekindergarten applicants' siblings across all grades, but we currently can observe only the PK3 and PK4 lotteries. The latter involves data on "Early Action" schools.

[4] In DC, there is a single lottery run, but other school systems conduct several lotteries. This statement applies only to systems with a single lottery run.

[5] The lottery numbers should be drawn from a uniform distribution defined between 0 and 1 (inclusive).

[6] The term "nondegenerate" refers to probability not being equal to exactly 1 or 0. We define 0 as those with an estimated DA propensity score of less than a 0.01 percent chance of being matched.

[7] Binned scatterplots report the mean of the variable on the vertical axis across bins of the horizontal variable axis that are of equal sample size (100 quantiles). Lottery year effects have been partialed out of the estimates presented in figure 2.

[8] We restrict the sample to applicants who are DC residents.

[9] To confirm that this procedure does a good job of linking these two data sources, we flagged different combinations of unique student identifiers with mismatched grades or years. Then, within each bin, we checked for mismatched birthdays and addresses to determine whether any matched unique student identifiers were incorrectly merged.

[10] This is similar to the compliance rate in the Boston prekindergarten study (Weiland et al. 2019).

[11] We use heteroskedasticity robust standard errors in all models.

[12] Nonresidents can apply to the DC prekindergarten lottery but can enroll only if they move to DC.

References

Abdulkadiroğlu, Atila, and Tayfun Sönmez. 2003. "School Choice: A Mechanism Design Approach." *American Economic Review* 93 (3): 729–47.

Abdulkadiroğlu, Atila, Joshua D. Angrist, Yusuke Narita, and Parag A. Pathak. 2017. "Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation." *Econometrica* 85 (5): 1373–432.

Abenavoli, Rachel, Pamela A. Morris, Rebecca Unterman, and Elise Cappella. Forthcoming. "Leveraging Pre-K Application Lotteries to Evaluate Teacher Professional Development: Weighing the Advantages and Disadvantages." Paper to be presented at the Association of Public Policy Analysis and Management 2020 Fall Research Conference, November 11–13.

Friedman-Krauss, Allison H., W. Steven Barnett, Karin A. Agrver, Katherine S. Hodges, G. G. Weisenfeld, and Beth Ann Gardiner. 2020. *The State of Preschool 2019: State Preschool Yearbook.* New Brunswick, NJ: Rutgers Graduate School of Education, National Institute for Early Education Research.

Gale, D., and L. S. Shapley. 1962. "College Admissions and the Stability of Marriage." *American Mathematical Monthly* 69 (1): 9–15.

Greenberg, Erica, Grace Luetmer, Carina Chien, and Tomas Monarrez. 2020. *Who Wins the Preschool Lottery? Applicants and Application Patterns in DC Public Prekindergarten.* Washington, DC: Urban Institute.

Manship, Karen, Ann-Marie Faria, and Juliette Berg. Forthcoming. "Lottery Structures for Admission to Public Montessori Preschool Programs: Promises and Challenges for Research." Paper to be presented at the Association of Public Policy Analysis and Management 2020 Fall Research Conference, November 11–13.

My School DC. 2018. "2018, Lottery: Number of Applications by Grade." Washington, DC: My School DC.

Pathak, Parag A. 2011. "The Mechanism Design Approach to Student Assignment." *Annual Review of Economics* 3 (1): 513–36.

Weiland, Christina, Rebecca Unterman, Anna Shapiro, Sara Staszak, Shana Rochester, and Eleanor Martin. 2019. "The Effects of Enrolling in Oversubscribed Prekindergarten Programs through Third Grade." *Child Development* 95 (1): 1401–22.

Weixler, Lindsay Bell, Jane Lincove, and Jon Valant. Forthcoming. "Using the New Orleans Early Childhood Education Lottery to Design Causal Impact Studies." Paper to be presented at the Association of Public Policy Analysis and Management 2020 Fall Research Conference, November 11–13.

About the Authors

Tomas Monarrez is a research associate in the Center on Education Data and Policy at the Urban Institute. His research focuses on education policy topics as they relate to economic and racial inequality. Monarrez received bachelor's degrees in economics and mathematics from the University of Texas at Austin and earned his doctoral degree in economics from the University of California, Berkeley.

Erica Greenberg is a senior research associate in the Center on Education Data and Policy. Her research spans early childhood and K–12 education, focusing on programs and policies like public prekindergarten, Head Start, child care subsidies, and home visiting. She also investigates the causes, consequences, and measurement of educational inequality. Greenberg holds a BA from Yale University. She received her MA in political science and her PhD in education policy from Stanford University.

Grace Luetmer is a research analyst in the Center on Education Data and Policy. She graduated from Boston College with a BA in economics and from Carnegie Mellon University with an MS in public policy and management.

Carina Chien is a research assistant in the Center on Education Data and Policy. She graduated from Cornell University with a bachelor's degree in economics and comparative literature.

Statement of Independence

The Urban Institute strives to meet the highest standards of integrity and quality in its research and analyses and in the evidence-based policy recommendations offered by its researchers and experts. We believe that operating consistent with the values of independence, rigor, and transparency is essential to maintaining those standards. As an organization, the Urban Institute does not take positions on issues, but it does empower and support its experts in sharing their own evidence-based views and policy recommendations that have been shaped by scholarship. Funders do not determine our research findings or the insights and recommendations of our experts. Urban scholars and experts are expected to be objective and follow the evidence wherever it may lead.

URBAN INSTITUTE · ELEVATE · THE · DEBATE

500 L'Enfant Plaza SW
Washington, DC 20024

www.urban.org