

Urban Advantage Interim Report

May 2010

Meryle Weinstein

Lila Nazar

Jacob Leos-Urbel

Elizabeth Debraggio

Amy Ellen Schwartz

*Institute for Education and Social Policy
New York University*

I. Introduction

Urban Advantage (UA) is a comprehensive program, managed by the American Museum of Natural History, in partnership with seven New York City science-rich cultural institutions. Designed to improve scientific learning and investigation in middle schools in New York City, UA provides professional development to teachers, school administrators, and parent coordinators along with resources to schools, students, and families. UA takes advantage of the wealth of intellectual and institutional capacity available in the city and facilitates access to those resources for the city's students. This report presents the first results of the study being conducted by the Institute for Education and Social Policy at New York University of the first five years of UA.

II. Average Achievement in UA and Non-UA Schools

Measures and definitions

UA is a program that primarily impacts the teachers who have participated in UA professional development sessions and students who have been taught by UA teachers, therefore, there should be differences in teaching and learning between UA schools that have only one participating teacher compared to a school with four participating teachers. This report uses school-level data obtained from the *New York City Annual Report Cards* and the *New York State School Report Cards* for academic years 2003-04 through 2007-08 to examine the characteristics of schools participating in UA and compare them to non-participating schools. Therefore, the results presented here provide only initial evidence about the impact of any participation in UA by teachers and students in a school on science achievement. To more fully estimate the impact of UA on student achievement, individual student and teacher level data is necessary and these analyses will be undertaken in the next phase of our research.

Our primary outcome measure is average test scores on the New York State Intermediate Level Science (ILS) Examination administered in 8th grade. Performance on this test can be expressed in three ways: scale score, performance levels, and standardized score. A *scale score* is a re-scaled version of the number of correct test answers that the student scored. This scale score ranges from 0 to 100. A school's *average scale score* is simply the average of its students' scale scores on that test.

Performance levels are discrete ranges of scale scores determined by state educational standards: Level 1 (Not Meeting Learning Standards), Level 2 (Partially Meeting Learning Standards), Level 3 (Meeting Learning Standards) and Level 4 (Meeting Learning Standards with Distinction). Students scoring at Level 3 or Level 4 on a given subject exam are considered to be “proficient” in that subject and meeting state standards. A school's *proficiency rate* in a given subject is the percent of students reaching Level 3 or 4 in that subject.

Finally, a *standardized score* (also known as a “z”-score) indicates where a student's scale score falls in the *distribution* of test scores. This score is calculated as the difference between a student's scale score and the city average score, divided by the overall standard deviation in test scores.¹ Its interpretation is straightforward: a student's standardized score tells us how far he/she scored from the city average test-taker in his/her grade, in standard deviation units. A standardized score of 1.5 indicates a student scored 1.5 standard deviations above average. Similarly, a standardized score of -0.3 indicates a student who scored 0.3 standard deviations below average.² A score at grade level average has a standardized score of zero. At the school level, the *average standardized score* in a given subject is simply the average of its students' standardized scores on that test.

¹ The standard deviation is a measure of dispersion, or variation, in scores. Loosely, it can be thought of as how far away from the mean the average student scored. If all students receive the same score, the standard deviation is zero.

² When test scores are distributed normally, roughly 68 percent of students fall between 1 standard deviation below and 1 standard deviation above the mean. Similarly, roughly 95 percent fall between 2 standard deviations below and 2 standard deviations above the mean.

Each of these three measures has advantages and shortcomings. Scale scores are an “absolute” measure of performance, but the citywide (and statewide) average fluctuates from year to year and is sensitive to test inflation or changes in test design. Cross-grade comparisons of scale scores may be problematic in practice, even if they are intended to be comparable. These problems may stem from the lack of vertical scaling, or shifts over time in the grade composition of schools.³ Proficiency rates are an easy way to understand “absolute” measures of performance, and are the most commonly cited student achievement measure as the centerpiece of *No Child Left Behind*. They are also a critical component of the *New York City School Progress Reports*. On the other hand, they mask a great deal of information and frequently provide misleading comparisons of school performance.⁴

While less transparent to the average reader, standardized scores address most of the shortcomings cited above. Their use is standard practice in educational research and evaluation, because of their comparability across tests, across grades, and over time. We report both proficiency rates and standardized scores.

Figure 1 displays the percentage of students scoring at Level 3 or 4 on the ILS and compares the percentage of students meeting the standards (scoring in levels 3 or 4) in UA and non-UA schools. In 2004-05, UA schools had a lower percentage of students meeting the standards on the science assessment compared to non-UA schools. In 2005-06 and 2006-07 the percent of students meeting the standards moved closer to, but still slightly below, those of non-UA schools and this

³ A comparison of average scale scores for two schools at two points in time may make sense if the grade composition of these schools remains constant. But if one school increases its population of 6th graders (for example) relative to the other, and 6th grade scale scores are typically lower than other grades, then the former school will be “penalized” for its growth in its 6th grade population.

⁴ To illustrate, suppose two schools make equal improvements in their students’ scale scores. Assume the first school’s students were originally just below the Level 3 cut score and the second’s were much further below the cut score. Even if the two schools make identical progress, the first school’s proficiency rate is likely to rise much more than the second. This example can be extended to include a third school where many students are already above the Level 3 threshold, but still make the same progress in scale scores as the other two schools. In this case, the first school will appear to have significantly greater “gains,” as measured by proficiency rates, than both the second and third school.

difference is not statistically significant. However, in 2007-08, the percent of students at UA schools who met the standards was 53.5 percent to 48.6 percent among non-UA schools and this difference is statistically significant.

Figure 1. Percent Scoring at Level 3 or 4, UA and Non-UA Schools by Year

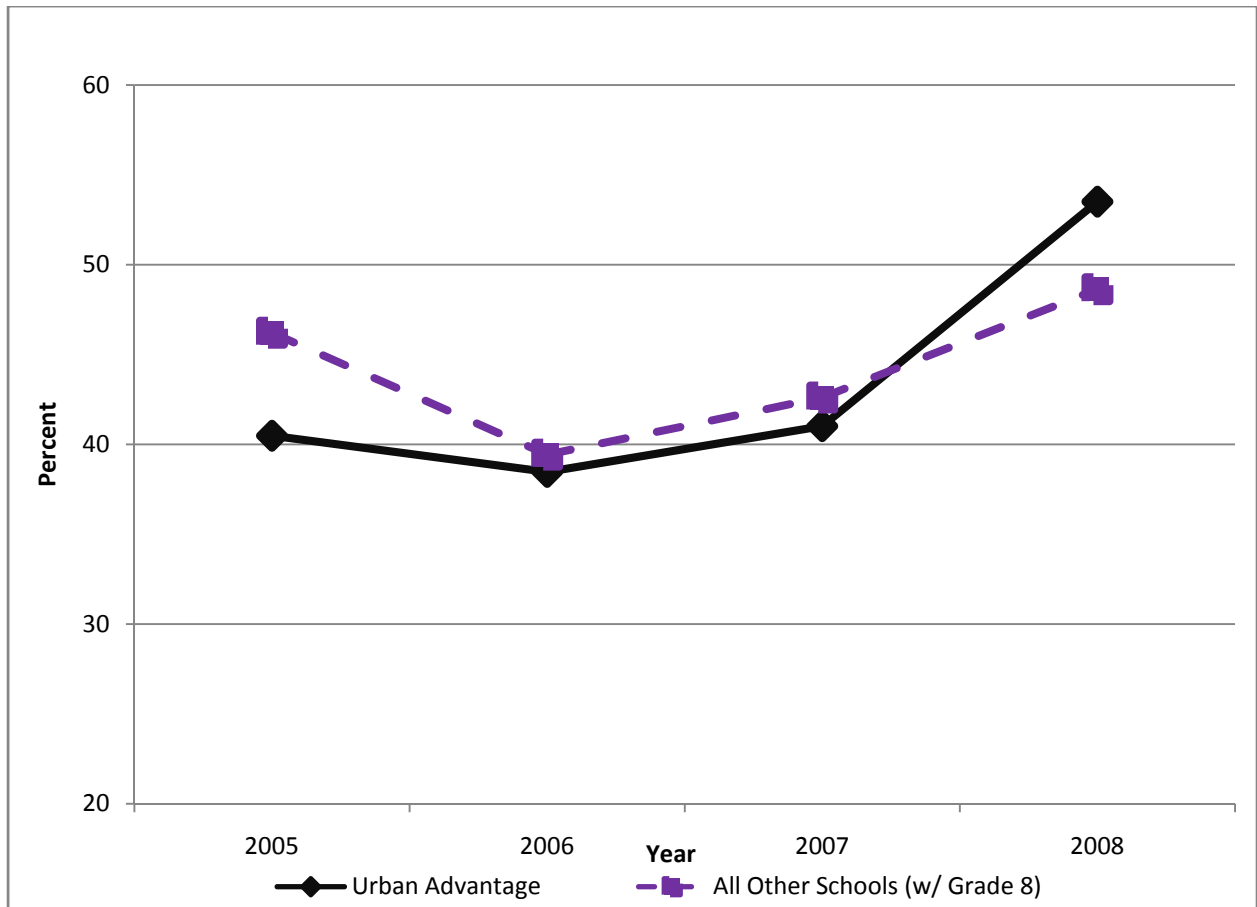
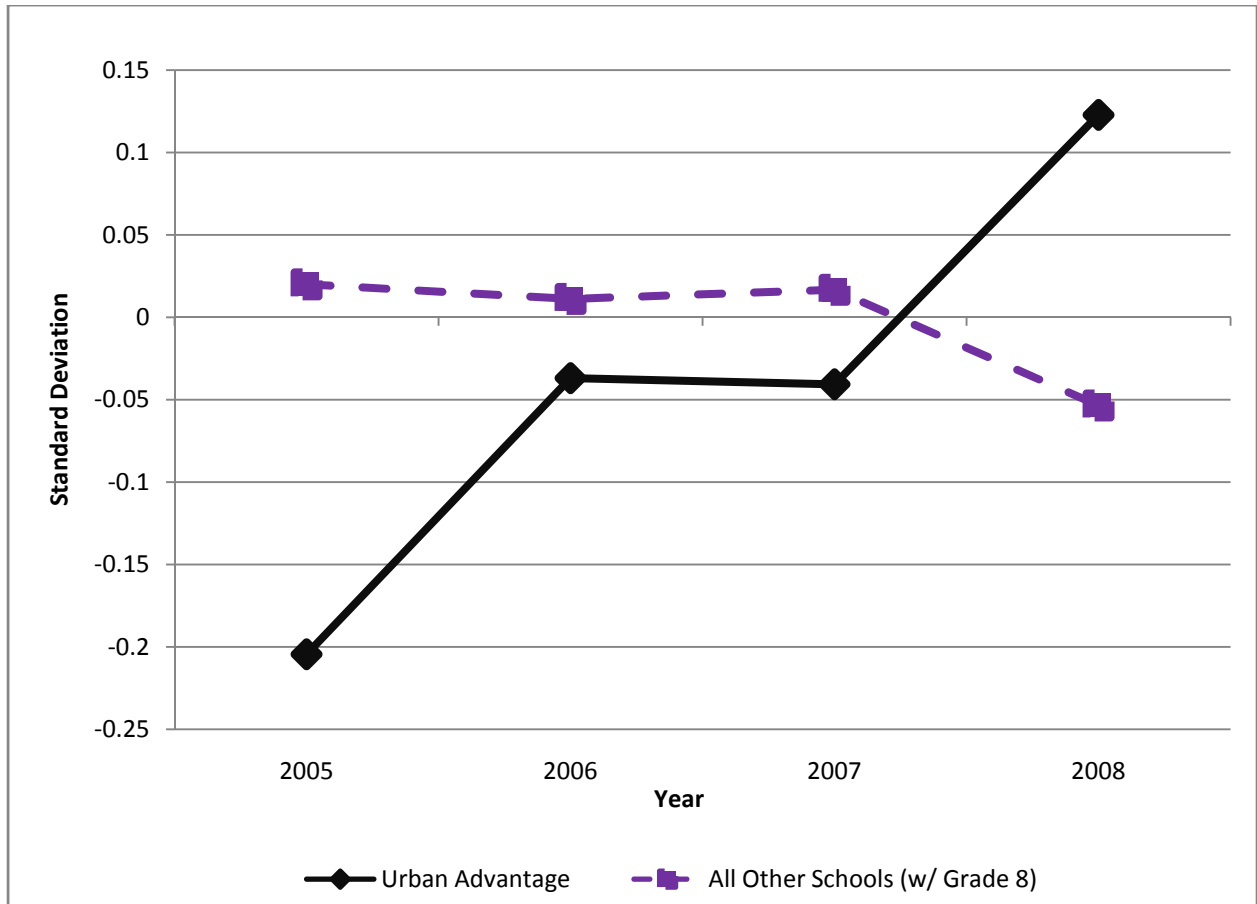


Figure 2 shows schools' average standardized scores on ILS. These scores can be compared to the Figure 1, which showed the percent scoring proficient at levels 3 or 4. Using standardized scores, we again see that UA schools on average scored below non-UA schools in 2004-05, drew

closer to non-UA schools in 2005-06 and 2006-07, and scored higher than non-UA in 2007-08 – and again, this difference is statistically significant.

Figure 2. Z-Scores, 8th Grade Science by UA and non-UA Schools by Year



Finally, we examined the growth in achievement with the number of years the school has participated in UA. Following the research suggesting that it takes at least three years for a reform to show results, we distinguish two cohorts of schools: Cohort 1 includes all those schools that began participating in UA in 2004-05 and remained through 2007-08; Cohort 2 includes those that began in 2005-06 and remained through 2007-08. As shown in Figure 3, schools in Cohort 1 had higher levels of science proficiency in the years before entering Urban Advantage compared to both Cohort 2 and to schools that never participated in Urban Advantage. However, from 2005-06 to

2007-08, proficiency levels are much more similar between the two cohorts, with Cohort 2 moving from slightly below Cohort 1 in the percent of students proficient to slightly above. By 2007-08, schools in Cohort 2 have a higher percentage of students meeting the standards than either those in Cohort 1 or those who have never participated in UA.

Figure 3. Percent Scoring at Levels 3 or 4, UA Cohort by Never UA

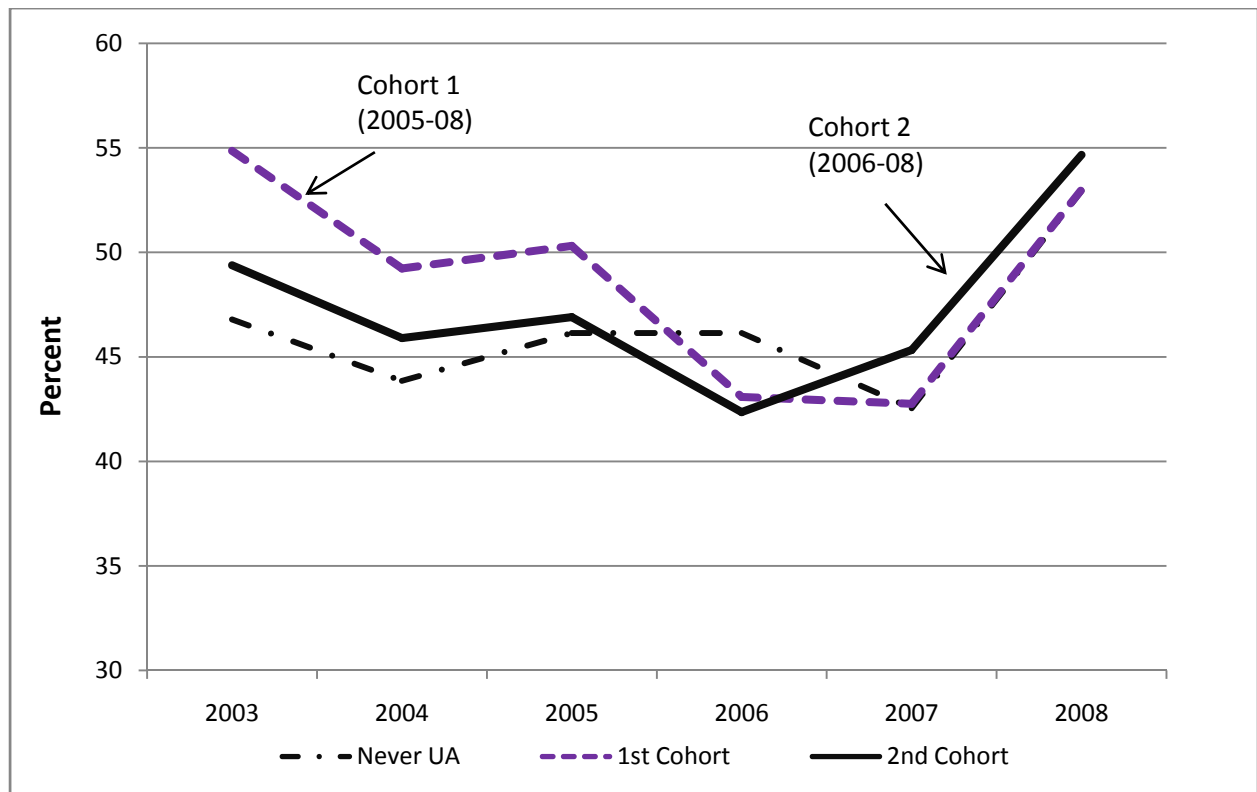
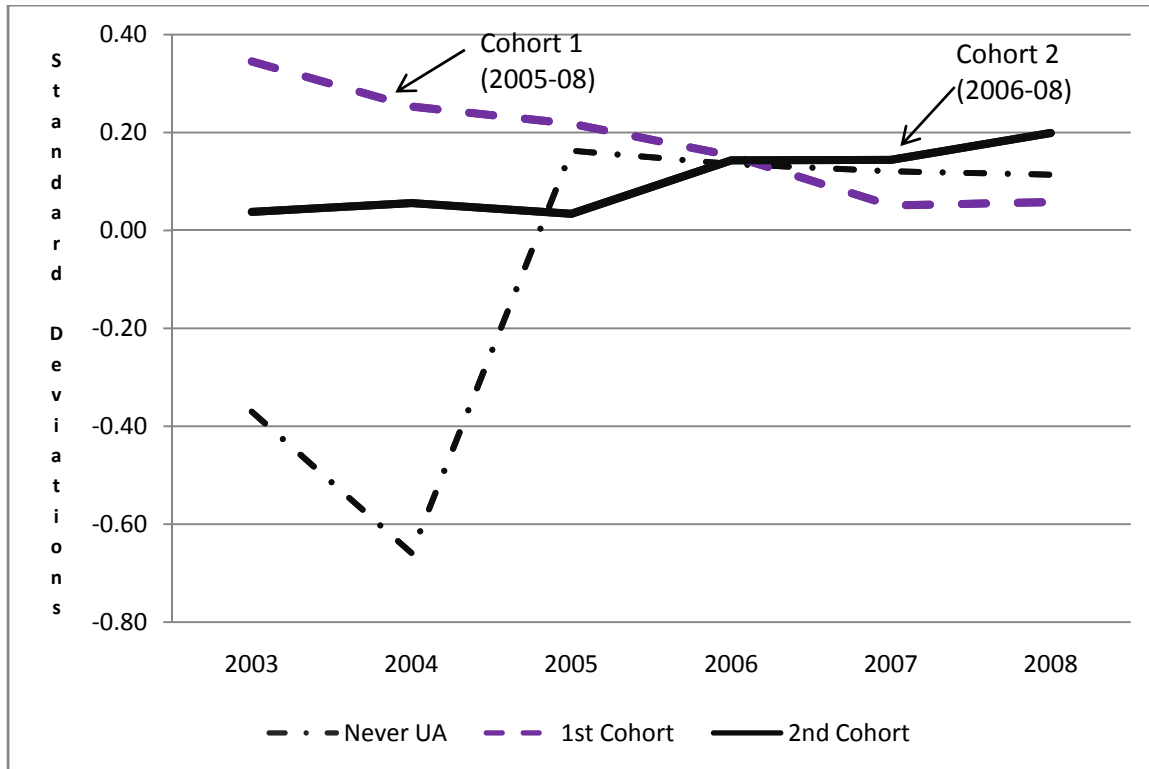


Figure 4 shows the trends in average standardized scores on the ILS for comparison with the average proficiency rates results in Figure 3. Results are similar. The average standardized scores for Cohort 1 are above those of Cohort 2 schools and non-UA schools from 2002-03 through 2004-05. In 2005-06 the lines cross indicating that all three groups have roughly the same average standardized scores, and in 2006-07 and 2007-08 Cohort 2 has higher average standardized scores, followed by non-UA schools and Cohort 1 schools.

Figure 4. Z-Scores, Science Test, UA Cohort by Never UA



Analysis of Preliminary Findings on Academic Achievement

First, the lag between the implementation of UA and changes in performance is not unexpected. Consistent with the views of scholars of school improvement, three years is the minimum amount of time needed to see such results.⁵

Second, although this comparison of test scores suggests good news, that is that UA does contribute to improved science performance for students in participating schools, the results cannot be readily interpreted as measuring the *causal effect* of Urban Advantage, the performance differential between UA and non-UA schools for two primary reasons. First, there may be differences between

⁵ Fullan, Michael and Suzanne M. Stiegelbauer. 1991. *The New Meaning of Educational Change*. New York: Teachers College Press.

the students and schools participating in UA and those that do not, and the data shown in these figures do not adjust for these differences. Second, UA has evolved over time – in its programmatic features as well as its recruitment – and these differences may be important determinants of success. Again, these analyses employ a naïve definition of participation and have not controlled for or examined any features of program implementation, participation, etc. To summarize, the analyses do not control for any characteristics of students or schools or program characteristics which, perforce, limits the interpretation of the results.

III. School Characteristics: Urban Advantage Schools in Context

UA currently serves schools in every borough and City Council District in the city and there is at least one UA school in almost every community school district (See Appendix for distribution by borough and community school district). UA has grown from serving 31 schools and 5,500 students in 8th grade only in 2004-05 to serving 155 schools and 27,000 students in grades 7 and 8 in 2007-08.⁶

As shown in Table 1 UA schools are, in many respects, quite similar to other New York City schools serving 8th graders.⁷ The one consistent difference we found between UA and non-UA schools is size. On average, UA schools are larger than non-UA schools, ranging from over 1000 students in 2004-05 to almost 600 in 2007-08, compared to between 400 and 800 for non-UA schools.

⁶ UA also serves schools in District 75, the citywide special education district. Because these schools typically do not report test scores, they have been excluded from the analysis.

⁷ New York City schools have a variety of grade span configurations that include grade 8. Some schools are traditional middle schools that serve grades 6-8, while others may be K-8 or 6-12.

Table 1: Urban Advantage Schools Compared to Non-Urban Advantage Schools by Year of Entry

	Urban Advantage				All Other Schools with Grade 8			
	2005 N=26	2006 N=61	2007 N=42	2008 N=39	2005 N=307	2006 N=319	2007 N=302	2008 N=257
Total Enrollment	1046 (468)	670 (469)	567 (379)	575 (445)	765 (443)	543 (392)	493 (360)	471 (327)
Percent Black	38.31 (22.75)	35.39 (26.67)	37.57 (29.17)	33 (28.68)	37.57 (29.01)	39.29 (29.23)	40.28 (29.80)	40.34 (29.94)
Percent Hispanic	37.32 (22.89)	42.75 (25.81)	43.4 (28.09)	43.05 (25.07)	40.4 (26.16)	40.64 (26.54)	40.21 (26.52)	40.17 (26.82)
Percent Asian/Other	14.71 (20.72)	9.05 (12.78)	7.31 (12.53)	12.72 (15.72)	9.08 (11.98)	8.1 (12.52)	8.31 (12.84)	8.18 (13.06)
Percent White	9.66 (18.60)	12.05 (18.37)	11.36 (20.80)	10.79 (14.40)	12.94 (19.40)	11.36 (18.56)	10.63 (17.85)	10.85 (27.73)
Percent ELL	11.43 (7.05)	11.71 (11.32)	11.31 (10.70)	12.15 (10.44)	10.77 (11.05)	10.7 (11.23)	10.58 (12.76)	10.72 (11.65)
Percent Free Lunch	71.81 (19.77)	68.32 (26.23)	59.17 (10.70)	65.44 (26.39)	68.24 (22.76)	69.17 (23.21)	67.34 (26.77)	66.37 (25.73)

Standard deviations in parentheses

As UA has grown over time, the profile of UA schools has remained fairly stable, although there have been some changes. For instance, the percentage of students in UA schools who are black decreased from 38.3 percent in the 2005 cohort to 33.0 percent in the 2008 cohort, while the percentage of Hispanic students increased from 37.3 percent among schools entering UA in 2004-05 to 43.0 percent among schools entering in 2007-08. To be clear, none of the differences between UA and non-UA schools displayed in Table 1 were found to be statistically significant; that is, while the characteristics of the schools entering UA have slightly changed over time, they are still serving students who are similar to other students in non-UA schools.

Table 2 compares the mean characteristics of all UA schools in a particular year to non-UA schools in that same year. Again, on average, UA schools are larger than non-UA schools and characteristics of students at UA schools are similar to those at non-UA schools with grade 8.

Table 2: Mean characteristics of UA and Non-UA Schools by Year

	Urban Advantage				All Other Schools with Grade 8			
	2005 N=26	2006 N=86	2007 N=116	2008 N=129	2005 N=307	2006 N=325	2007 N=321	2008 N=301
Total Enrollment	1046 (468)	720 (475)	657 (439)	635 (446)	765 (443)	553 (404)	502 (369)	487 (338)
Percent Black	38.31 (29.75)	36.94 (27.78)	35.24 (27.46)	32.81 (28.34)	37.57 (29.01)	39.31 (29.18)	40.59 (29.83)	40.75 (29.46)
Percent Hispanic	37.32 (22.89)	40.70 (24.96)	43.12 (26.20)	42.91 (26.40)	40.40 (26.16)	40.62 (26.45)	39.95 (26.33)	40.32 (26.30)
Percent Asian/Other	14.71 (20.72)	10.00 (14.84)	9.66 (15.13)	11.59 (16.36)	9.08 (11.98)	8.24 (12.76)	8.44 (12.88)	8.19 (12.92)
Percent White	9.66 (18.60)	11.65 (18.60)	11.51 (19.17)	12.36 (19.27)	12.94 (19.40)	11.23 (18.43)	10.45 (17.65)	10.26 (17.52)
Percent ELL	11.43 (7.05)	11.26 (10.35)	11.72 (12.29)	12.40 (12.01)	10.77 (11.05)	10.68 (11.16)	10.48 (12.55)	10.71 (11.32)
Percent Free Lunch	71.81 (19.77)	69.27 (24.43)	62.57 (29.09)	63.57 (26.86)	68.24 (22.76)	69.19 (23.11)	67.50 (26.63)	67.07 (25.61)

Standard deviations in parentheses

Across UA schools, as with city schools as a whole, there is much variation. As the large standard deviations in Table 1 and Table 2 show, UA serves students in schools that vary in size from very large to very small, from a student body where all are eligible for free lunch to those where only a small proportion are eligible, and schools where majority of students are black or Hispanic to those that have a more balanced mix of students of different ethnicities. This finding is consistent through each cohort of new schools entering participation in the program.

Analysis of Preliminary Findings on School Characteristics

The tables above suggest that UA is serving a student population that is demographically similar to that of all New York City public schools serving 8th graders. Again, it should be noted that any differences in demographic characteristics between UA and non-UA schools or between groups of UA schools by year are not statistically significant.

IV. Next Steps

In the next phase of the analysis, we will further explore the above findings using student-level data. Using student-level data will enable us to conduct a more nuanced examination of the impact of UA, including differences among students both within and between UA schools. We will use regression methods to control for student characteristics including poverty, English language proficiency, as well as prior academic performance on other exams. In addition, we will be able to explore the importance of the length of exposure to UA, for example, to distinguish students who attended a UA school in both 7th and 8th grade and therefore received two “doses” of UA from those who only had a one year dose in 8th grade. We will look at the relationship between teacher participation in UA professional development and student achievement. Additionally, we will expand our analysis to look at subsections of the science exams and questions that specifically address scientific inquiry, which is the main focus of UA.

Appendix

Figure 1A. Number of UA Schools by Borough and Year

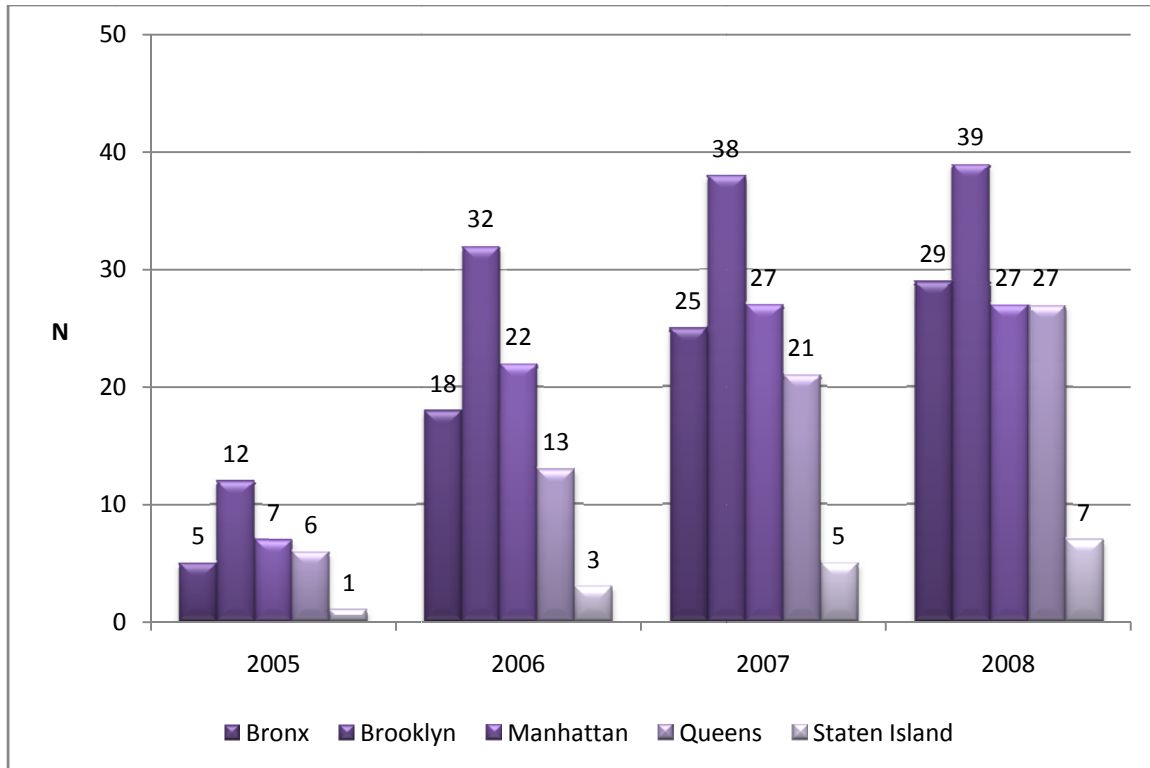
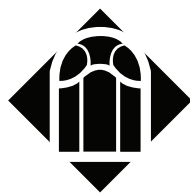


Table 1A. Number of UA Schools by Community School District & Year

CSD	2005	2006	2007	2008	Total N Schools in CSD 2008	% Schools in CSD 2008
1	1	3	4	2	12	16.67
2	2	4	4	4	18	22.22
3	2	4	7	7	18	38.89
4	0	3	4	5	15	33.33
5	2	3	3	3	9	33.33
6	0	5	5	6	20	30.00
7	1	2	4	4	13	30.77
8	1	2	3	5	12	41.67
9	0	2	5	3	23	13.04
10	1	4	6	7	26	26.92
11	1	4	5	6	15	40.00
12	1	4	2	4	19	21.05
13	0	1	2	3	11	27.27
14	1	2	3	2	8	25.00
15	1	4	7	5	14	35.71
16	0	1	0	0	9	0.00
17	2	4	5	6	16	37.50
18	1	1	1	2	8	25.00
19	2	5	6	6	13	46.15
20	1	2	1	4	9	44.44
21	1	4	4	3	15	20.00
22	0	1	1	1	9	11.11
23	2	5	5	4	22	18.18
24	1	2	4	4	9	44.44
25	0	1	3	5	14	35.71
26	1	1	3	4	6	66.67
27	1	1	2	2	16	12.50
28	1	4	2	5	7	71.43
29	1	1	4	3	14	21.43
30	1	3	3	4	12	33.33
31	1	3	5	7	11	63.64
32	1	2	3	3	7	42.86
Total	31	88	116	129	430	30.00



I N S T I T U T E F O R
Education and Social Policy

The Institute for Education and Social Policy is a joint research center of NYU's Wagner and Steinhardt Schools. Founded in 1995, IESP brings the talents of a diverse group of NYU faculty, graduate students and research scientists to bear on questions of education and social policy. We are one of the nation's leading academic research centers addressing urban education issues and a substantial amount of our work focuses on New York City public schools and students. More information is available on our website at <http://steinhardt.nyu.edu/iesp>.