SUPPLEMENT TO "THE ELITE ILLUSION: ACHIEVEMENT EFFECTS AT BOSTON AND NEW YORK EXAM SCHOOLS"
(*Econometrica*, Vol. 82, No. 1, January 2013, 137–196)

By Atila Abdulkadiroğlu, Joshua Angrist, and Parag Pathak

This document describes the processing steps from the original data files to the analysis files used in the main paper.

APPENDIX C: Boston Data

Boston Public Schools is the source for four data sets: the exam school application file, the enrollment file which contains student demographic and school attendance information, the Massachusetts Comprehensive Assessment System (MCAS) test score file, the College Board test file which contains PSAT, SAT, and AP scores, and data from the National Student Clearinghouse. This appendix describes these data sets and the procedures used to construct the analysis sample.

C.1. *Sources*

Exam School Application File

The exam school application file contains a record for each student consisting of a registration number, application year, grade, date of birth, preferences over three exam schools, and scores on the ISEE verbal, quantitative, reading, and math sections. Each record also includes the rank of each student by the exam schools on their preference list and the school where the student receives an offer (if any). This data set covers students in grades 7, 9, and 10 and application years 1999–2008. Since there are a small number of grade 10 applicants, we kept students applying for grades 7 and 9 only.

Table C.I indicates the steps involved in processing the exam application file. We excluded duplicate observations, applicants from private schools, and those who did not rank or were not ranked by any exam school. We also dropped students who obtained an offer at an exam school that is not on their preference list.

For each applicant, the exam school application file indicates whether the student receives an offer at a school to which they've applied. For a given application year, grade, and school, we compute the lowest-ranked student to obtain an offer from that school. Each student is then coded as obtaining an offer at an exam school if her score is above this minimum cutoff for any school that is on her preference list.

TABLE C.I

PROCESSING OF BOSTON EXAM SCHOOL APPLICATION DATA[a]

Application Year (1)	Total Number of Records (1)	Excluding Duplicate Observations (2)	Excluding Applicants From Private Schools (3)	Excluding Students Who Did Not Rank an Exam School (4)	Excluding Students Who Are Not Ranked by an Exam School (5)	Excluding Students Who Obtain an Offer at a School They Do Not Rank (6)	Excluding Students Not Matched to Boston Public Schools at Baseline (7)	Excluding Students Previously Enrolled in Exam School (8)	Excluding Students With No Observed Outcome MCAS Test Scores (9)
				Panel A. 7th Grade Applicants					
1999	2,353	2,353	1,353	1,307	1,307	1,307	1,180	1,180	1,000
2000	2,283	2,283	1,252	1,165	1,165	1,165	1,125	1,125	1,032
2001	2,317	2,317	1,299	1,196	1,196	1,196	1,193	1,193	1,100
2002	2,365	2,365	1,304	1,237	1,236	1,236	1,235	1,235	1,118
2003	2,494	2,494	1,386	1,251	1,251	1,251	1,240	1,240	1,127
2004	2,217	2,217	1,206	1,174	1,174	1,174	1,172	1,172	1,083
2005	2,062	2,062	1,116	1,105	1,105	1,099	1,095	1,095	1,001
2006	2,079	2,079	1,184	1,166	1,166	1,161	1,158	1,158	1,052
2007	1,992	1,992	1,086	1,081	1,080	1,073	1,068	1,068	974
2008	1,874	1,874	1,050	1,049	1,040	1,036	1,013	1,013	908
All years	22,036	22,036	12,236	11,731	11,720	11,698	11,479	11,479	10,395
				Panel B. 9th Grade Applicants					
2001	1,520	1,520	863	787	787	787	783	680	496
2002	1,607	1,607	876	829	828	828	826	755	553
2003	1,750	1,750	951	812	812	812	809	727	546
2004	1,723	1,723	936	918	918	918	912	815	631
2005	1,630	1,630	936	924	924	924	918	832	642
2006	1,729	1,729	992	981	981	981	977	889	677
2007	1,684	1,683	945	936	931	931	930	842	612
All years	11,643	11,642	6,499	6,187	6,181	6,181	6,155	5,540	4,157

[a]This table summarizes the steps going from the raw application data to the analysis sample.

Enrollment File

The BPS enrollment file contains an end-of-year (June) snapshot for each student enrolled in Boston Public Schools, with unique student identifier (the BPS ID), the student's grade and school, and demographic information.

The variables of interest in the enrollment file are grade, year, date of birth, sex, race, special education (SPED) and limited English proficiency (LEP) status, subsidized lunch eligibility, and school. Students are coded as attending an exam school if their year-end enrollment is at an exam school. Years at an exam school is the total number of years where a student is at an exam school at the end of the year. We transform the enrollment file into a wide-format layout for each student where we compute the grade and exam school years attended for a given year. Finally, we keep only students that attend Boston Public Schools in 6th or 8th grade and use their demographic information from that year.

MCAS Test File

Each record in the MCAS test file contains a student identifier (BPS ID) and scores on MCAS tests in a given year. We use data from school years 1999–2000 through 2008–2009. The scores we look at are math and English Language Arts (ELA) for grades 4–10. The MCAS test file contains raw scores for all BPS test takers for math, English Language Arts, writing, and science. As shown in Table C.II, the number of grades tested has increased over time. MCAS Math for grade 8 was the first examination offered in 1999. By the end of our data, there are tests for math and English tests at grades 7, 8, and 10. Baseline scores for grade 7 applicants are from 4th grade MCAS exams. For 9th grade applicants, baseline math is from 8th grade math and baseline English is from 7th grade English, since the 8th grade English exam was first offered in 2006.

We standardize scores to have mean zero and standard deviation 1 within a subject-grade-year among all test-takers in Boston Public Schools. When there is more than one test score for a student for a particular subject, we use the first available score. Boston exam offers appear to boost grade repetition at some schools and cohorts, though not consistently. In any case, increased repetition seems likely to boost scores, if anything, since repetition increases age at test.

College Board Test File

The College Board provides BPS with reports on the test performance of all BPS test-takers from 2004–2005 through 2009–2010. These files come with the name, date of birth, address, gender, school of test, and test year for each exam. BPS matched the PSAT file for October 2004 and October 2005, the SAT file which is available from 2005–2009, and the Advanced Placement test file, available from 2005–2009. The timing of these tests for our applicant cohorts is shown in Table C.II.

The PSAT file is not matched to BPS student IDs for years 2006–2009, so we have to link College Board files to BPS files for these years. The address information in the College Board file is entered by the test-taker and does not

TABLE C.II

DATA STRUCTURE AND TEST OUTCOMES FOR BOSTON[a]

Application Year	Math 7 (1)	Math 8 (2)	Math 10 (3)	English 7 (4)	English 8 (5)	English 10 (6)	PSAT (7)	SAT (8)	AP (9)
Panel A. 7th Grade Applicants									
1999		2001	2003			2003			2005
2000		2002	2004	2001		2004	2004	2006	2006
2001		2003	2005	2002		2005	2005	2007	2007
2002		2004	2006	2003		2006	2006	2008	2008
2003		2005	2007	2004		2007	2007	2009	2009
2004		2006	2008	2005	2006	2008	2008	2010	2010
2005	2006	2007	2009	2006	2007	2009	2009		
2006	2007	2008		2007	2008				
2007	2008	2009		2008	2009				
2008	2009			2009					
Panel B. 9th Grade Applicants									
2001			2003			2003		2005	2005
2002			2004			2004	2004	2006	2006
2003			2005			2005	2005	2007	2007
2004			2006			2006	2006	2008	2008
2005			2007			2007	2007	2009	2009
2006			2008			2008	2008	2010	2010
2007			2009			2009	2009		

[a]This table reports the years for which outcomes are available for each applicant cohort. Application year refers to the fall of application year, while test outcome year refers to the spring of year. Test outcomes are available based on the schedule of the MCAS and availability of SAT, PSAT, and AP score outcomes.

immediately concord with the BPS address system. There also appear to be small errors in the date of birth in the College Board file for similar reasons. Our procedure to match these files to the BPS registration files is as follows. First, we take all unique year, date of birth, gender, school of test, and zip code matches between the BPS registration file and the College Board PSAT file. Among the remaining unmatched PSAT records, we take all unique year, date of birth, gender, and school of test matches between the two files. Finally, for the remaining unmatched PSAT test records, we hand-match the records for these four years to the closest record in the registration file, attempting to correct mismatches due to address misspellings or typos in the date of birth.

BPS students take AP exams across a range of subjects. The tests with 500 or more takers are Calculus AB/BC, Statistics, Biology, Chemistry, Physics B/C, English Language and Composition, English Literature and Composition, European History, U.S. Government and Politics, U.S. History, Microeconomics, and Macroeconomics. The other tests are Art History, Art: Drawing, Art: 2D Design, Art: 3D Design, Chinese Language and Culture, Computer Science A, Computer Science AB, Environmental Science, French Language, French Literature, German Language, Comparative Government and Politics, Latin:

Vergil, Latin: Literature, Music Theory, Spanish Language, Spanish Literature, and World History.

We standardize the PSAT and SAT scores to have mean zero and standard deviation 1 within a year among all test-takers in Boston Public Schools. When there is more than one test score for a student, we use the earliest available one. We only use applicant cohorts for whom we might expect to observe PSAT, SAT, or AP exams; these are summarized in Table C.II.

National Student Clearinghouse

BPS matches data on seniors to National Student Clearinghouse (NSC) files, which record information on enrollment at over 90 percent of American four-year colleges and universities. We match these files to the Barron's Profile of American Colleges, provided to us by David Deming in electronic form, to form our list of competitiveness. The most common colleges attended in our sample are Bunker Hill Community College, University of Massachusetts Boston, University of Massachusetts Amherst, Roxbury Community College, Boston University, Massachusetts Bay Community College, Suffolk University, Boston College, Salem State University, and Northeastern University. Table C.V shows the number of applicants we observe with NSC outcomes relative to the number of students we expect (defined as those enrolled in BPS in grade 12).

C.2. *Matching Data Sets*

The MCAS test file and enrollment files are merged by grade, year, and BPS ID. Any test record that is not matched to the enrollment file is dropped.

The exam applicant file is matched to the enrollment/MCAS file using an auxiliary table that links exam registration number to BPS ID. This table provides a BPS ID for each exam registration number. For a small number of cases, an exam registration number is matched to more than one BPS ID. In these cases, we matched the registration number to the BPS ID where the date of birth is the same between the exam applicant and enrollment file.

Table C.III reports match rates from exam applicant file to the enrollment/MCAS file. The overall match rate is 98.1 percent (11,476 out of 11,698) for grade 7 applicants and 99.6 (6155 out of 6181) for grade 9 applicants. The match rate for offered students in grade 7 is 98.5%, while the match rate for students who were not offered is 97.8%. The lower match rates come from earlier application years 1997–2000. The match rate for not offered is larger than for offered for three of these years, and the differences in match rates are small. For grade 9, where the application cohorts start in 2001, the match rate for offered students is 99.9%, while for non-offered it is 99.5%. Applicants who are not matched to the enrollment file at baseline are dropped, as are applicants who enrolled in an exam school before application. This latter restriction only impacts grade 9 applicants, as can be seen by comparing columns (7) and (8) of Table C.I.

TABLE C.III

Match From Boston Exam Application to Enrollment Data[a]

Application Year	Number of Students (1)	Fraction With Match		
		Total (2)	Offered (3)	Not Offered (4)
Panel A. 7th Grade Applicants				
1999	1,307	0.903	0.919	0.890
2000	1,165	0.966	0.958	0.972
2001	1,196	0.997	0.996	0.998
2002	1,236	0.999	1.000	0.999
2003	1,251	0.991	0.996	0.987
2004	1,174	0.998	1.000	0.997
2005	1,099	0.996	0.996	0.996
2006	1,161	0.997	0.995	1.000
2007	1,073	0.995	1.000	0.991
2008	1,036	0.978	0.989	0.966
All years	11,698	0.981	0.985	0.978
Panel B. 9th Grade Applicants				
2001	787	0.995	1.000	0.993
2002	828	0.998	1.000	0.997
2003	812	0.996	1.000	0.995
2004	918	0.993	1.000	0.992
2005	924	0.994	0.992	0.994
2006	981	0.996	1.000	0.994
2007	931	0.999	1.000	0.999
All years	6,181	0.996	0.999	0.995

[a]This table provides summary statistics on the match between the exam school application data and the Boston Public School enrollment file. The sample in column (1) is the sample in column (6) of Table C.I.

C.3. *Construction of the Analysis Sample*

The size of the final analysis sample is presented in column (8) of Table C.I.

Some of the analysis stacks grades and includes multiple test scores for individual students. For each student in an application year, Table C.IV reports the number of students with at least one followup test score (column (2)). It also presents the number of test scores expected for each cohort, and the number of test scores observed for both math and English. For example, a 7th grade applicant for the 2005–2006 school year contributes math scores in grade 7 (Spring 2006), 8 (Spring 2007), and 10 (Spring 2009). Hence, we expect 3285 math scores from the 1095 applicants for this cohort, and we observe at least one score for 1001 students, which corresponds to a total of 2650 student-score observations. On the other hand, a 7th grade applicant for the previous school year contributes one fewer test score (no grade 7 math). Table C.V shows a related analysis of expected followup for PSAT, SAT, and AP scores, and the NSC outcomes.

TABLE C.IV

TEST OUTCOME DATA FOR BOSTON EXAM SCHOOL APPLICANTS[a]

Application Year	Number of Students (1)	Number With an Observed Test Score (2)	Number of Math Test Scores Expected (3)	Math Test Scores Observed (4)	Number of English Test Scores Expected (5)	English Test Scores Observed (6)
			Panel A. 7th Grade			
1999	1,180	1,000	2,360	1,765	1,180	800
2000	1,125	1,032	2,250	1,776	2,250	1,792
2001	1,193	1,100	2,386	1,843	2,386	1,897
2002	1,235	1,118	2,470	1,894	2,470	1,945
2003	1,240	1,127	2,480	1,897	2,480	2,006
2004	1,172	1,083	2,344	1,842	3,516	2,890
2005	1,095	1,001	3,285	2,650	3,285	2,650
2006	1,158	1,052	2,316	2,039	2,316	2,038
2007	1,068	974	2,136	1,884	2,136	1,879
2008	1,013	908	1,013	895	1,013	907
All years	11,479	10,395	23,040	18,485	23,032	18,804
			Panel B. 9th Grade			
2001	680	496	680	496	680	495
2002	755	553	755	551	755	550
2003	727	546	727	545	727	543
2004	815	631	815	621	815	630
2005	832	642	832	630	832	636
2006	889	677	889	662	889	673
2007	842	612	842	603	842	610
All years	5,540	4,157	5,540	4,108	5,540	4,137

[a]This table summarizes the observed test score outcomes for exam school applicants. The sample is restricted to students in column (8) of Table C.I.

APPENDIX D: NEW YORK DATA

The New York City Department of Education is the source for three data sets: the exam school application and Student Enrollment Office (formerly, OSEPO) files which contain demographic information, the registration file which contains school attendance information, and the NYSED and Regents test score file. This appendix describes these data sets and details the procedures used to construct the analysis sample.

D.1. *Sources*

Exam School Application and Enrollment Office Files

The exam school application file is maintained by the Enrollment Office, which runs high school admissions. All applicants must take the Specialized High School Admissions Test (SHSAT) to apply to an exam school. On test day, students also submit a ranking of exam schools. At a later date, students

TABLE C.V

MATCHING OF COLLEGE BOARD TEST OUTCOME DATA FOR BOSTON APPLICANTS[a]

Application Year (1)	Number of Applicants (1)	Number With an Observed PSAT Test Score (2)	Number With an Expected PSAT Test Score (Enrolled as of Grade 11) (3)	Number With an Observed SAT Test Score (4)	Number With an Expected SAT Test Score (Enrolled as of Grade 11) (5)	Number With an Observed AP Test Score (6)	Number With an Expected AP Test Score (Enrolled as of Grade 12) (7)	Number With an Observed NSC Outcomes (8)	Number With an Expected NSC Outcomes (Enrolled as of Grade 12) (9)
				Panel A. 7th Grade					
1999	1,180	50	0	640	0	291	1,180	840	1,180
2000	1,125	707	1,125	647	1,125	341	1,125	879	1,125
2001	1,193	826	1,193	710	1,193	432	1,193	959	1,193
2002	1,235	834	1,235	683	1,235	427	1,235	1,053	1,235
2003	1,240	844	1,240	687	1,240	481	1,240	1,059	1,240
2004	1,172	788	1,172	679	1,172	499	1,172	949	1,172
2005	1,095	664	1,095	3	0	345	0	202	0
2006	1,158	10	0	0	0	14	0	137	0
2007	1,068	0	0	0	0	0	0	115	0
2008	1,013	0	0	0	0	0	0	0	0
All years	11,479	4,723	7,060	4,049	5,965	2,830	7,145	6,193	7,145
				Panel B. 9th Grade					
2001	680	22	0	374	680	113	680	522	680
2002	755	462	755	413	755	159	755	602	755
2003	727	520	727	426	727	177	727	618	727
2004	815	635	815	478	815	235	815	786	815
2005	832	598	832	454	832	255	832	805	832
2006	889	612	889	481	889	290	889	797	889
2007	842	528	842	2	0	142	0	201	0
All years	5,540	3,377	4,860	2,628	4,698	1,371	4,698	4,331	4,698

[a]This table summarizes the observed College Board test score outcomes for exam school applicants. The sample is restricted to students in column (8) of Table C.I.

are also required to complete a New York City Public High School Admissions Application and submit it to their guidance counselor.

Several Enrollment Office files are used in the analysis. The first contains a record for each student indexed by their ID number (OSISID) and their score on the SHSAT. For each student, the exam school offer file contains a list of the schools ranked and an indicator for the school at which the student obtains an offer (if any). The Enrollment Office student file has demographic information such as grade, sex, race, home language code, and borough of residence for each student. There are also separate files indicating special education and limited English proficiency for each student. Each file for a given application year contains an OSISID number for each student, which allows us to merge the files together.

Registration and Enrollment Files

The NYC registration file is from the Office of School Performance and Accountability and is available as part of data underlying school progress reports. The registration and enrollment cover all public school students in grades 9 to 12 for school years 2002–2003 through 2008–2009. This data set includes each student's OSISID, grade, and current school as of October in the school year. The registration data are used to determine whether and for how many years a student enrolls in an exam school, where a student who is enrolled in October is counted as enrolling for the entire year. Starting in 2004–2005, there is a separate file that contains a list of all students who obtain a subsidized school lunch in that year. This variable is used to code subsidized lunch status for applicants using the application year. For applicants in 2003–2004, 2004–2005, and 2005–2006, we use the lunch status record from 2004–2005. For application cohort in 2006–2007, we use the lunch status record from 2005–2006.

Table D.I indicates the steps involved in processing the exam application file and merging it with the Student Enrollment files. From the file of exam applicants, we eliminate private school applicants (based on whether their OSISID starts with the letter "A") and those who do not submit a New York City Public School Admissions Application (based on the Round 1 HS ranking file). The 4000–5000 private school applicants are excluded because these students do not have a NYC ID at the time of application, they do not have baseline information, and the relevant counterfactual for this population is unlikely to be a regular NYC public high school. We also exclude students who did not rank at least one of the three original academic exam schools: Bronx Science, Brooklyn Technical, and Stuyvesant.

Baseline Test Files

The NYC Department of Education also provided us with NYSED grade 8 standardized exams in math and English Language Arts for all public school students for years 2002–2003 through 2007–2008. These tests are taken in the

TABLE D.I

PROCESSING OF NYC EXAM SCHOOL APPLICATION DATA[a]

Application Year	Total Number of Records (1)	Excluding Applicants From Private Schools (2)	Excluding Applicants Not in Round 1 of the Application Process (3)	Excluding Students Who Did Not Rank an Exam School (4)	Excluding Students Who Did Not Rank Brooklyn Tech, Bronx Science, or Stuyvesant (5)
2003–2004	28,136	23,637	22,293	22,287	22,205
2004–2005	28,279	24,123	22,894	22,859	22,776
2005–2006	28,442	23,971	22,810	22,810	22,376
2006–2007	26,616	22,377	21,278	21,278	20,824
All years	111,473	94,108	89,275	89,234	88,181

Application Year	Excluding Students Not Matched to Student File (6)	Excluding 9th Graders (7)	Excluding Students Who Took SHSAT in Previous Years (8)	Excluding Students Without Post-Assignment Numeric Outcome Test Scores at All (9)
2003–2004	22,108	21,091	21,091	18,361
2004–2005	22,776	21,883	21,880	19,106
2005–2006	22,376	21,448	21,446	18,842
2006–2007	20,824	20,124	20,122	17,431
All years	88,084	84,546	84,539	73,740

[a]This table summarizes the steps going from raw application data to the analysis sample.

winter of grade 8 and are required of all public school students in the state. These tests serve as our baseline math and English scores.

Regents Test File

The NYC Regents test file contains the date and raw score for each tested student. Regents exams are mandatory state examinations where performance determines whether a student is eligible for a high school diploma in New York. There are Regents examinations in English, Global History, U.S. History, and multiple exams in math and science. A Regents exam typically has a multiple choice section and a long answer or essay component, and each exam usually lasts for three hours. The English exam, however, consists of two three-hour pieces over two days. The exam has a locally graded component and test scores bunch near performance thresholds.

The New York State Board of Regents governs and designs the Regents exams. Starting in 2005, they started to modify the math exams. At the beginning of our sample, the two math exams were Elementary Algebra and Planar Geometry (Math A) and Intermediate Algebra and Trigonometry (Math B). Two new math exams, Integrated Algebra I (Math E) and Geometry (Math G), have since been phased in. Since students typically either take Math A or Math E, we focus on the score on the test taken first, taking the Math A score when both are contemporaneous. Likewise, students typically either take Math B or Math G, so we focus on the score that comes first, taking the Math B score when both are contemporaneous. We denote the first test outcome as "Math" and the second outcome as "Advanced Math." There are Regents science exams in Earth Science, Living Environment, Chemistry, and Physics. The science outcome we focus on is Living Environment because it is the only Regents science exam required to obtain a state high school diploma.

In Table D.II, for each test, we report the number of applicants and the number of test scores we observe. English and U.S. History Regents exams are typically taken in 11th grade. For the 2006–2007 applicants, we expect to observe these scores in 2009–2010, a year after the Regents test score file's last date. Even though there are a small number of students who take these exams before the 11th grade, we do not examine Regents English and U.S. History outcomes for the 2006–2007 applicant cohort, since the vast majority do not.

Since students may take Regents exams multiple times, there can be multiple test scores per student in the Regents test file. Table D.II presents the number of students who have taken each exam more than once among the exam applicant sample. This fraction is about 10%, with slightly higher retake rates for math and Global History. Some students may also take Regents exams before exam school enrollment. Table D.II shows the fraction of students who take exams before enrolling in an exam school. A large fraction of exam school applicants take math before enrolling. Most Regents exams are offered in January, June, and August, with most students usually taking tests in June.

TABLE D.II

MATCH FROM NYC EXAM APPLICANTS TO REGENTS TEST SCORE OUTCOMES[a]

	Application School Year				
	2003–2004	2004–2005	2005–2006	2006–2007	All Years
Record Availability	(1)	(2)	(3)	(4)	(5)
I. Math					
Number of applicants	21,091	21,880	21,446	20,122	84,539
Number with score observed before treatment	2,685	3,157	3,673	3,975	13,490
Number with score observed after treatment	15,055	15,307	14,206	12,492	57,060
Number with different multiple scores observed after treatment	1,795	2,354	2,000	1,821	7,970
Number with different multiple scores observed after treatment, on first date	3	20	10	2	35
Number with score observed on most common date	5,822	5,873	6,078	8,033	25,806
Number with score observed before most common date	3,522	3,875	4,348	2,022	13,767
Number with score observed after most common date	5,711	5,559	3,779	2,437	17,486
II. Advanced Math					
Number of applicants	21,091	21,880	21,446	20,122	84,539
Number with score observed before treatment	7	9	13	29	58
Number with score observed after treatment	10,375	10,691	10,939	12,130	44,135
Number with different multiple scores observed after treatment	1,469	1,750	899	234	4,352
Number with different multiple scores observed after treatment, on first date	13	4	0	0	17
Number with score observed on most common date	3,913	3,938	5,496	11,177	24,524
Number with score observed before most common date	4,310	4,671	5,443	953	15,377
Number with score observed after most common date	2,152	2,082	0	0	4,234

(*Continues*)

Record Availability	Application School Year				
	2003–2004 (1)	2004–2005 (2)	2005–2006 (3)	2006–2007 (4)	All Years (5)
III. English					
Number of applicants	21,091	21,880	21,446	20,122	84,539
Number with score observed before treatment	2	1	0	11	14
Number with score observed after treatment	16,847	17,322	17,202	3,039	54,410
Number with different multiple scores observed after treatment	1,979	2,024	1,501	137	5,641
Number with different multiple scores observed after treatment, on first date	11	3	0	0	14
Number with score observed on most common date	9,333	8,614	8,985	2,457	29,389
Number with score observed before most common date	1,829	2,587	2,705	582	7,703
Number with score observed after most common date	5,685	6,120	5,512	0	17,317
IV. Global History					
Number of applicants	21,091	21,880	21,446	20,122	84,539
Number with score observed before treatment	3	19	18	8	48
Number with score observed after treatment	17,057	17,735	16,434	15,429	66,655
Number with different multiple scores observed after treatment	2,321	2,882	1,771	203	7,177
Number with different multiple scores observed after treatment, on first date	19	59	1	0	79
Number with score observed on most common date	13,746	13,471	13,100	14,328	54,645
Number with score observed before most common date	796	844	1,037	1,101	3,778
Number with score observed after most common date	2,514	3,420	2,296	0	8,230

(*Continues*)

TABLE D.II—*Continued*

Record Availability	Application School Year				
	2003–2004 (1)	2004–2005 (2)	2005–2006 (3)	2006–2007 (4)	All Years (5)
V. U.S. History					
Number of applicants	21,091	21,880	21,446	20,122	84,539
Number with score observed before treatment	41	23	91	101	256
Number with score observed after treatment	15,766	16,015	14,270	1,855	47,906
Number with different multiple scores observed after treatment	1,152	1,102	496	212	2,962
Number with different multiple scores observed after treatment, on first date	20	0	3	0	23
Number with score observed on most common date	10,252	10,365	11,844	970	33,431
Number with score observed before most common date	1,464	2,068	2,426	55	6,013
Number with score observed after most common date	4,049	3,582	0	830	8,461
VI. Living Environment					
Number of applicants	21,091	21,880	21,446	20,122	84,539
Number with score observed before treatment	440	878	894	922	3,134
Number with score observed after treatment	16,562	16,807	16,310	14,102	63,781
Number with different multiple scores observed after treatment	1,356	1,807	1,484	977	5,624
Number with different multiple scores observed after treatment, on first date	2	7	8	0	17
Number with score observed on most common date	11,601	11,455	11,286	11,071	45,413
Number with score observed before most common date	207	324	344	209	1,084
Number with score observed after most common date	4,754	5,027	4,679	2,822	17,282

[a]This table summarizes the match between Regents test score outcomes and exam school applicants. The sample is restricted to students in column (8) of Table D.I.

For some subjects, such as Global History, most applicants take the test at the end of 10th grade. For other tests, such as math (Math A or E), many students take the exam before entering high school and some students take the exam multiple times. The exact number of students who take the exam before 9th grade, the number who take the exam more than once after 9th grade, and the number who take the exam on a date other than the most common date are presented in Table D.II. For each test where there is a retake, we only use the first test outcome. New York exam offers do not consistently boost grade repetition.

For each test, students who have scores before the 9th grade are omitted because they tested prior to potential exam enrollment. If a student takes the test more than once after 9th grade, we use the test score from the earliest date. There are a small number of cases where there is more than one score on the same date, and this date is the first date after entering 9th grade. In some of these cases, there are two different test codes, where one code ends with a "2." We use the score corresponding to the test that does not end with a "2." Otherwise, we treat the score as missing.

For each subject, we standardize scores to have mean zero and standard deviation 1 within year-semester-subject among the universe of students: 8th graders from public school who participated in Round 1 of the HS Admissions process, have valid demographic information, and did not take the SHSAT test in a previous year.

D.2. *Matching Data Sets*

We match the exam application to the student file using the OSISID. Table D.III shows the match rates. Nearly every student who has applied to an exam school can be matched to the corresponding Enrollment Office student file. The student file allows us to identify whether an applicant is in grade 8 or 9. Since there are a limited number of 9th grade applicants for grade 10 spots, we keep only students applying for grade 9. Finally, our sample is limited to first-time SHSAT takers.

For each exam school and applicant year, the exam school offer file indicates the school at which a student obtains an offer (if any). The offered school is the student's most preferred school where a student has a high enough SHSAT score. For each school, we compute the minimum score needed to obtain an offer using the offer indicators and SHSAT scores. We code anyone with an SHSAT score above the lowest score offered as having received an offer.

Students are coded as attending an exam school if they are enrolled at an exam school in the registration file.

D.3. *Construction of the Analysis Sample*

After processing the exam application file, we next match it to the registration file for grade 9. An exam applicant may not match to the registration file

TABLE D.III

MATCH FROM NYC EXAM APPLICATION TO STUDENT DATA[a]

Application Year	Number of Students (1)	Fraction With Match		
		Total (2)	Offered (3)	Not Offered (4)
2003–2004	22,205	0.996	0.997	0.995
2004–2005	22,776	1	1	1
2005–2006	22,376	1	1	1
2006–2007	20,824	1	1	1
All years	88,181	0.999	0.999	0.999

[a]This table reports the fraction of applicants with a match between the exam application file and the student demographic file. The sample corresponds to column (5) of Table D.I.

if she leaves New York City's Public Schools following application. Such an applicant would not contribute any followup scores.

To generate the final analysis data set, we merge the student registration and test file with the exam application file. The exam application file contains the OSISID, a list of exam schools that students have ranked, and the student's raw SHSAT test score. These data span four years: 2003–2004 through 2006–2007.

Next, we merge baseline scores for students for whom they are available. Finally, we merge the data set of cleaned Regents outcome scores. For each test, we compute the implied years of exam school attendance based on the test date and enrollment status. If a student took a Regents test in the fall semester, we compute years assuming the exam date is January 31st. Otherwise, we compute years assuming the exam date is June 1st. The resulting file is our analysis sample. An applicant who is matched to the registration file for grade 9 may not contribute followup scores if the applicant leaves New York City's public schools before taking a Regents exam. The last column of Table D.I indicates the sample of students who contribute at least one followup score.

Dept. of Economics, Duke University, Durham, NC 27708, U.S.A.; atila.abdulkadiroglu@duke.edu,

Dept. of Economics, MIT, 77 Massachusetts Ave., Cambridge, MA 02142, U.S.A. and NBER; angrist@mit.edu,

and

Dept. of Economics, MIT, 77 Massachusetts Ave., Cambridge, MA 02142, U.S.A. and NBER; ppathak@mit.edu.