New Mexico Charter School Study: Technical Appendix

April 30, 2021

Rachel McCormick, Eleanor Harvill, Jacqueline Mendez, and Meghan Shea, Abt Associates

In partnership with:
Matthew Pahl, *Public Charter Schools of New Mexico*;
Michael Rodríguez and David Rogers, *Dual Language Education of New Mexico*;
Elizabeth Howard, *University of Connecticut;* and
the *New Mexico Public Education Department*.

Abt Associates
6130 Executive Boulevard
Rockville, MD 20852

ACKNOWLEDGEMENTS

We appreciate the many contributions to this project from: David Rodriguez, Juanita Jaramillo, Wendy McRae, Marina Kosareva, and David Dupree (currently or formerly of ***Abt Associates***), Yanira Gurrola Valenzuela (of ***Dual Language Education of New Mexico***), and Zhanna Galochkina, Richard Trujillo, and Alan Brauer (currently or formerly of the ***New Mexico Public Education Department)***. Special thanks to Barbara Goodson of ***Abt Associates*** for her endless guidance and enthusiasm.

We also acknowledge the commitment to rigorous research, the sharing of data, and the helpful feedback on this project from the 21 schools (and their staff) who participated in this study.

Finally, we gratefully acknowledge the support from ***Arnold Ventures*** that made this research possible and Amanda Moderson-Kox for her collaboration and always thoughtful input.

CONTENTS

CONTENTS

Overview of Study Design

This technical appendix supports the New Mexico Charter School Study: Findings Report, providing detailed information on statistical methods and results. The Charter School Lottery Study section describes the experimental methods and detailed findings. The Dual Language and Charter School Matching Study section describes the quasi-experimental methods and findings.

Research Questions

The evaluation answers the following three confirmatory research questions with three experimental studies, focusing on the impacts on student outcomes:

RQ1. What is the overall impact of oversubscribed elementary charter schools in New Mexico on grade 3 math and ELA achievement?[1]

RQ2. What is the overall impact of oversubscribed charter middle schools in New Mexico on grade 8 math and ELA achievement?[2]

RQ3. What is the overall impact of oversubscribed charter high schools in New Mexico on college enrollment?

In addition, we answer one secondary research question with quasi-experimental study:

RQ4. What is the impact of oversubscribed dual language immersion charter and magnet schools in New Mexico on grade 5 English Language Arts (ELA) achievement?

We also answer the first three questions using the quasi-experimental approach described below. The QED analyses are secondary rather than confirmatory.

School Recruitment and Lottery Data Collection

Of the 97 charter schools in New Mexico, 30 were sufficiently oversubscribed and eligible to participate in the study.[3] Twenty-one of the 30 eligible schools agreed to participate in the study. Three schools declined to participate because lotteries for the relevant time-period were unavailable. Six schools did not respond or declined to participate.

All 21 schools that agreed to participate provided admissions records, though not all of these records could be used to identify the results of the initial lottery. Ten participating schools provided lottery records that allowed us to conduct the experiment: i.e. they allowed us to (1) identify the students that were admitted in the initial lottery and (2) follow the students from application through outcome measurement. These schools are included in multiple studies if students were admitted by lottery at multiple grade levels and if lottery records were sufficient to support the experiment. Of the 10 schools in the experimental sample, eight cover multiple educational levels: five K-8 schools and three middle-high schools. We only have two schools with traditional grade spans: one middle school and one high school.

---

1 We conducted multiple comparison adjustments across the math and ELA impacts addressing RQ1.

2 We conducted multiple comparison adjustments across the math and ELA impacts addressing RQ2.

3 School oversubscription was determined via a survey of charter school principals and, if survey responses were not available, discussions with principals and charter school field experts. Schools were eligible for the study if they had at least 20 students on the wait list at an entry grade at the time of survey response (during the 2018-19 academic year).

New Mexico Public Education Department (NMPED) Administrative Data

The NMPED provided student-level, longitudinal administrative data on student demographics and education outcomes as follows:

1. NMPED provided a limited number of variables for all students in the state across the time period covered. These variables included **key student identifiers and enrollment information**.
2. Abt used these data to: (1) match school lottery records to PED data, (2) define the study sample of schools, and (3) create the **finder file which identifies the study sample at the school-by-grade-by-year-level**.

For the Lottery Study studies, the finder file was constructed to follow all students who appear on lottery applications from the year prior to their application to the end of the follow-up period (the 2020-2021 school year).

For the QED studies, the finder file was constructed to request data on treatment and comparison schools. Our goal in identifying comparison schools is to capture the alternative options parents consider when enrolling their child in a treatment school. These are the schools that might appear in the control group of an experimental lottery study. Ideally, we would use the same process we use to identify the experimental sample of individuals (described above) and define comparison schools based on the schools attended by lottery applicants. However, the purpose of the quasi-experimental study is to expand the sample of treatment schools beyond those for whom it is possible to identify lottery applicants. Therefore, we also use students who switch into or out of the treatment school to identify comparisons. Comparison schools are the schools most commonly selected by lottery applicants and the schools most commonly attended by students who also attend the treatment school.

3. Then, Abt requested data on **outcomes, covariates, and descriptive measures for the students in the study sample.**

These data are obtained from the NMPED School Accountability System, which houses state standardized test scores. Outcomes are not available for students who were enrolled in schools that do not participate in the Accountability System (such as those schools operated by county juvenile justice systems).

Pre-specification of Analysis

We registered our Pre-Analysis plan initially in May 2019 and updated the plan to reflect the results of lottery data collection in January 2021 before we conducted impact analyses. The registered plans are available on the Open Science Foundation page for the Lottery Study of New Mexico's Oversubscribed Dual Language & Traditional Charter Schools project: https://osf.io/kryzb/.

We followed our final plan to the greatest extent feasible and have noted places where we were not able to execute the study as planned. For example, our plan specified that we match the DL sample on Kindergarten test scores. However, there were only 12 eligible treatment students with Kindergarten test scores. We therefore matched on first grade test scores instead.

Charter School Lottery Study

This evaluation includes three experimental studies, focusing on oversubscribed elementary, middle school, and high school charter schools. The random assignment mechanism is the same across studies: Students apply to a charter school for entry in a particular grade. If there are more applicants than seats available, the school admits students via lottery. We use lottery records from participating schools to identify treatment and control groups of students. Treatment group students are defined as those offered admission to the charter school in the initial lottery, while control group students are those not offered admission in the initial lottery.

In an ideal version of this experiment, all treatment group students would enroll in the charter school and none of the control group students would do so. Then, comparing treatment group outcomes to control group outcomes would yield the effect of the charter school compared to whatever other schools the control group students attend. Following this logic, we will refer to participating charter schools throughout as treatment schools. However, treatment group students may decline offers of enrollment, and control group students may ultimately accept waitlist offers.

The intended treatment for the three experimental studies is as follows:

- The treatment for *the elementary study* consists of English-language charter schools that serve students in grades K-3, and the treatment group will be students who are admitted by lottery into those schools in kindergarten. For this study, we seek to measure the effect of four years of enrollment in a charter school (K-3) on student achievement.
- The treatment for *the middle school study* consists of all types of charter schools that serve students in grades 7-8, and the treatment group will be students who are admitted by lottery into those schools, typically in grade 6. For this study, we seek to measure the effect of at least two years (7-8), typically three years (6-8) of enrollment in a charter school on student achievement.
- The treatment for *the high school study* consists of all types of charter schools that serve students in grades 10-12, and the treatment group will be students who are admitted by lottery into those schools, typically in grade 9. For this study, we seek to measure the effect of at least three years (10-12), typically four years (9-12) of enrollment in a charter school on college enrollment.

These definitions reflect the reality that charter schools in New Mexico tend to serve multiple grade levels. As a result, relatively few schools in our study serve traditional grade ranges. For example, among the six schools in the middle school study, only one is a traditional middle school serving grades 6-8. The other schools serve grades K-8 (two schools), 6-12 (two schools) or 7-12 (one school).

Data Sources

The evaluation draws on two different data sources to construct outcomes and baseline covariates:

- National Student Clearinghouse (NSC) data on college enrollment. (High School study only.) The NSC defines college as a two- or four-year title IV postsecondary institution that offers Associate's or Bachelor's degrees, according to Integrated Postsecondary Education Data System (IPEDS).
- NMPED administrative data on demographics, enrollment, state standardized test scores, and high school graduation.

Sample and Confirmatory Outcomes

The charter Lottery Study student samples were constructed from entry lottery application records from four elementary schools, six middle schools and four high schools. Schools that serve multiple grade ranges are included in multiple studies. The randomized sample consists of all students who appear on any lottery application record. The analysis sample consists of all students for whom we observe confirmatory outcomes.

Exhibit 1. Sample and Primary Outcomes for Charter Lottery Studies

Lottery Study	Number of Charter Schools	Entry Grade	Confirmatory Outcome(s)	Students Included in Analysis Sample
Elementary	4	K	3rd Grade Math & ELA	NM administrative data contain scores for at least one confirmatory outcome
Middle	6	6 or 7 (1 school)	8th grade Math & ELA	NM administrative data contain scores for at least one confirmatory outcome
High School	4	9 or 10 (1 school)	College Enrollment	Lottery records and/or NM administrative data provide sufficient information to request National Student Clearinghouse data on college enrollment (Observe name and date of birth)

Estimation and Statistical Methodology

We model the outcome y_{is} of individual i who applies to the lottery for school s as follows:

$$y_{is} = \alpha_0 + \beta_1 T_{is} + \sum_j \gamma_j X_{sj} + \sum_k \delta_k Z_{isk} + \sum_m \lambda_m W_{ism} + \mu_s T_{is} + \nu_s + \varepsilon_{is},$$

where T_{is} is the Treatment group indicator, X_{sj} are school characteristics measured prior to the lottery, Z_{isk} are individual baseline characteristics, and W_{ism} are indicators for year of outcome measurement.[4] To reduce the variability of the outcome and increase statistical power, we included a baseline school-level test score and, if possible, a baseline individual test score. For students who enter multiple lotteries, we included them in a single lottery. Baseline covariates and students entering multiple lotteries are discussed further below.

The coefficient of the Treatment indicator, β_1, captures the impact of the average charter in our sample—this is the estimate we use to answer our confirmatory research questions. The regression adjusted control group mean is given by the intercept term, α_0. The coefficients γ_j and δ_k capture the relationships between school and individual characteristics and outcomes. The fixed effects for year of outcome measurement, λ_m, capture differences in tests across years. The random coefficient on Treatment, μ_s, captures variability in impacts across charters in our sample. The school-level random effect, ν_s, captures contextual factors that affect outcomes for all students who enter the lottery for school s, including students who are not admitted. The individual-level error term, ε_{is}, captures individual-level variation in outcomes not explained by the model.

We used inverse probability weighting to account for differential probabilities of random assignment across lotteries.

Baseline covariates

We include two types of baseline covariates in the model: school characteristics and individual student baseline characteristics.

[4] Year indicators are not included in the High School analysis as they are not defined for students who were not enrolled.

Exhibit 2. Individual Baseline Covariates in Lottery Study Analyses

Baseline Characteristic	High School	Middle School	Elementary School
Race/Ethnicity Indicators (White, Black, Hispanic, Native, Other)	✓	✓	✓
EL Status	✓	✓	✓
Free Reduced Price Lunch	✓	✓	✓
Baseline ELA Z Score	✓	✓	✓
Baseline Math Z Score		✓	
Baseline Math Proficiency Level	✓		

The set of school characteristics include the school level demographic and achievement measures (as listed above) for an earlier cohort. We use the most recent cohort of students that does not appear in the analysis sample. For example, if the analysis sample includes students who enter the kindergarten lottery in 2010-2013, we used the students who entered kindergarten in 2009 to define the school characteristics.

Students applying to multiple lotteries

Every analysis of charter school lotteries must address technical issues around student applications to multiple lotteries. Because we use a multi-level modeling analysis framework and estimate heterogeneity in treatment effects across charters, we must associate each student with a single lottery and use that lottery to define treatment status. Ideally, we would use the lottery for the school that the student most prefers to define treatment. The idea is that admission to this lottery is most highly correlated with attending a treatment school. If we select the wrong lottery, the student may be randomly admitted to their preferred treatment school but be in the control group for another school.

In other studies, researchers select the lottery used to define treatment in a variety of ways, including using student rankings of schools and the distance between school and home (c.f. Bloom et al., 2010; Unterman, 2017). These approaches are possible because the lotteries in the study were coordinated by a single administrative body that had a clear approach to randomly assigning students to schools, making it straightforward to identify the lottery most correlated with enrolling in a treatment school. Unfortunately, in New Mexico, each charter school runs their lotteries independently. For the confirmatory analysis, we use the results of the lottery that occurs earliest in time to define the treatment indicator (using data submitted with lottery records on the date of each lottery to determine which lottery occurred earliest in time). We refer to this as the students' "initial lottery."

Where possible, we used the earliest lottery because we are concerned that the timing of lotteries has the potential to bias estimates. Depending on the timing of the lotteries, a student might receive the results from an early lottery and then decide whether to apply to additional lotteries. If we did *not* use the earliest lottery to define the treatment indicator, individual choice would potentially affect the school to which the student is assigned and the treatment indicator, thereby biasing the impact estimate. Using the earliest lottery to define the treatment indicator addresses this concern and yields unbiased estimates. However, the earliest lottery may or may not be for the school most preferred by the student.

In cases where we were not able to determine the timing of the lotteries to determine which occurred earlier, we randomly selected one record to retain. Below, we report detailed of sample formation for each Lottery Study. The number of duplicate observations removed from each sample is included. The elementary study had the highest rate proportion of students applying to multiple schools, with approximately 12% of students applying to multiple lotteries.

Treatment of missing data

We did not impute outcome data, limiting the sample to the individuals for whom we observe outcomes.

We used dummy-variable imputation to address missing individual-level baseline covariates. (Puma et al, 2009)

Calculating attrition

We calculate attrition separately for each confirmatory, secondary and exploratory analysis as follows:

$$Attrition = \frac{\#\ Randomized - \#\ Analysis}{\#\ Randomized}$$

For each analysis, lottery records define the randomized sample. The analysis sample includes students who appear on the lottery records, for whom lottery records were matched to administrative data records, and who are not missing the relevant outcome data. This calculation considers students that appear in lottery records and were not matched to administrative data and students who are missing relevant outcome data to have attrited from the sample. We calculate attrition for the overall sample and for the treatment and control groups separately.

For each study, we report the details of the formation of the treatment and control groups for the study as a whole and, separately, calculate attrition for each analysis. In the attrition calculations, the number of students in the randomized sample may differ across analyses because we exclude students from the randomized sample if exogenous characteristics render them ineligible for the analysis. For example, for the high school study, PARCC scores are only available for students applying to enter 9th grade in 2013-2014 or later and applying to enter 10th grade in 2014-2015 or later. The randomized sample size for analyses of Algebra II PARCC scores excludes students applying to earlier lotteries.

Multiple comparison corrections

To see why multiple comparison corrections are necessary, imagine that there are no true differences between the treatment and control groups and you run 20 independent hypothesis tests at a 5% level of significance. Statistically, we would expect that you reject the null hypothesis once and draw one false conclusion. Multiple comparisons corrections are designed to limit the extent to which running many tests increases the probability of spurious findings.

We committed in our pre-analysis plan to perform multiple comparison corrections for confirmatory outcomes within each study. Because the high school study has only one confirmatory outcome, no multiple comparison correction is required. The elementary and middle school studies have two confirmatory outcomes (math and ELA achievement) and require multiple comparison corrections.

We intended to use a multiple comparisons procedure that controlled the family-wise error rate and accounted for correlation across impacts. These tests are more powerful than simpler procedures that assume tests are independent. We planned to implement the Romano & Wolf stepwise procedure introduced in 2005 and refined in 2016 (Romano & Wolf, 2005; Romano & Wolf, 2016), using the Stata package rwolf (Clarke, 2016). However, the rwolf package is not compatible with our multi-level modeling approach.

We therefore used a simple Bonferroni correction. This approach ignores correlation across outcomes and is likely overly conservative. However, given that none of the Lottery Study findings are significant before correcting, the loss of power does not affect the conclusions drawn from the study. We include 90% confidence intervals that reflect corrected p-values in table notes.

Charter High School Lottery Study

Outcomes

All outcomes are well-defined for all students, including those who drop out. If we used an outcome like achievement that is only measured for students who are enrolled, any differences between the groups in dropout/persistence would bias the comparison. This is an especially important consideration given that some charter high schools are designed to serve students who would otherwise drop out.

The analysis sample for NSC outcomes (including our confirmatory outcome) consists of students whose lottery application records and/or NM administrative data provide sufficient information to request NSC data on college enrollment—i.e., we observe name and date of birth. Students for whom we requested data and did not appear in the NSC data file summarizing college enrollment records are assumed to be not enrolled. NSC enrollment data covers approximately 95% of enrollment in eligible institutions in New Mexico over the follow-up period (NSC, 2021).[5]

[5] Coverage is slightly higher for all institutions in the United States.

The analysis sample for NMPED administrative data outcomes consists of all students who enroll in an NM school after their application for a charter high school. Students who never enroll in a NM charter or public school are assumed to have enrolled in a private school, to have enrolled in a Bureau of Indian Education school, or to have left the state. High school graduation and other outcomes are treated as missing for these students. Students who enroll in 9th or 10th grade and then are not subsequently enrolled are assumed to have dropped out. We take this approach because several of the charter schools told us that their applicants frequently enroll in private schools.

For seven students in the analysis sample for NMPED administrative outcomes, NMPED enrollment data indicates that they were enrolled in traditional charter or public high schools for which we did not obtain outcomes data in addition to being enrolled in schools for which we did obtain outcomes data.

Primary outcome: College enrollment in the fall after on-time high school graduation. Enrollment is a binary outcome defined using data from the NSC. A student is considered to have been enrolled in the fall after on-time high school graduation (outcome takes on a value of 1) if NSC data indicate that the student was enrolled for any length of time in a term that began between August 1 and October 31, inclusive. If the NSC data does not indicate that the student was enrolled, the measure takes on a value of zero, indicating that the student was not enrolled. On-time high school graduation is defined as four years after their application for 9th grade entry or three years after their application for 10th grade entry.

Exploratory outcomes

- ***College Retention*** is defined as continued enrollment (or degree completion) within the **same** higher education institution in the fall semesters of a student's first and second year. This measure is constructed from NSC data and the definition of the measure follows the NSC definition.[6]
- ***College Persistence*** is defined as continued enrollment (or degree completion) at **any** higher education institution — including one different from the institution of initial enrollment — in the fall semesters of a student's first and second year. As described above, this measure is constructed from NSC data and the definition of the measure follows the NSC definition.
- ***On-time High School Graduation.*** This measure is a binary measure based on NM state administrative data that takes on a value of 1 if the student is reported to have received a New Mexico High School Diploma or NM High School Diploma of Excellence within four years of their application for 9th grade entry or within three years of their application for 10th grade entry, and a value of zero otherwise.
- ***Passed Algebra II or equivalent (PARCC test)***, a binary indicator that takes on a value of 1 if the student took and passed one of the following PARCC tests (Algebra II, Integrated Math III) in the entry grade or later. (Note that these are the highest-level PARCC math tests). If a student is not enrolled or is missing test score data for another reason, we assume they did not take the test and therefore did not pass. Sample is limited to students applying to enter 9th grade in 2013-2014 or later and applying to enter 10th grade in 2014-2015 or later so that PARCC scores are available.
- ***Proficient on ELA Achievement Test in grade 10***, a binary indicator that takes on a value of 1 if the student took and passed the ELA achievement test. If a student is not enrolled, we assume they did not take the test.
- ***Proficient on ELA Achievement Test in grade 11***, a binary indicator that takes on a value of 1 if the student took and passed the ELA achievement test. If a student is not enrolled, we assume they did not take the test.

[6] See the definitions of persistence and retention here: https://nscresearchcenter.org/snapshotreport33-first-year-persistence-and-retention/.

Sample Formation, Baseline Characteristics, and Compliance with Random Assignment

Exhibit 3. High School Charter Lottery Study: Sample Sizes and Sample Formation

Sample	Treatment (N)	Control (N)	Total (N)
Randomized Sample: All applicants	297	422	719
Analysis Sample: All applicants with observed confirmatory outcomes	283	401	684
Details of Sample Formation			
Records in Lottery Files	297	423	720
Duplicate Lottery Records Removed	0	1	1
Randomized Sample: All applicants	297	422	719
Analysis Sample: All applicants with observed confirmatory outcomes	283	401	684
Matched to NMPED Enrollment Data	275	392	667
Analysis Sample with Baseline Test Data	209	273	482

Source: School Lottery Records & NMPED Administrative Data

Exhibit 4. High School Charter Lottery Study: Baseline Equivalence on State Achievement Tests

	Treatment group			Control Group			Treatment - Control	
Measure	Mean	Std. Dev.	Sample Size	Mean	Std. Dev.	Sample Size	Diff. (SE)	Effect Size
ELA (z-score)	0.273	0.969	207	0.201	1.024	270	0.072 (0.092)	0.072
Math (z-score)	0.233	0.960	207	0.127	1.036	262	-0.106 (0.093)	0.105

Source: NMPED Administrative Data
Note: Student baseline test scores were standardized using state-level means and standard deviations for that test, year, and grade-level. None of the differences are statistically significant at the 10% level.

Exhibit 5. High School Charter Lottery Study: Baseline Equivalence on Demographics

	Treatment Group		Control Group		Treatment - Control	
Measure	Mean	Sample Size	Mean	Sample Size	Diff. (SE)	Cox Index Effect Size
White (%)	33.7	279	30.6	385	3.0 (3.7)	0.030
Hispanic (%)	61.3	279	62.1	385	-0.8 (3.8)	0.005
Ever EL (%)	18.6	279	24.2	385	-5.5* (3.2)	-0.114
FRPL	65.0	183	66.9	257	-1.9 (4.6)	0.016

Source: NMPED Administrative Data
Note: Comparisons for race categories that make up less than 3% of the sample are not reported.
Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%.
Free/reduced price lunch data are only available for students with non-missing baseline test scores. Among students with non-missing test score data, FRLP data is consistently missing for the 2010-2011 school year and frequently missing for the 2013-2014 school year.

Exhibit 6. High School Charter Lottery Study: Compliance with Random Assignment

	Treatment (%)	Control (%)	Total (%)
Proportion compliant with random assignment	71.1	70.3	70.7
Proportion compliant with random assignment in least compliant school	79.9	36.6	60.6
Proportion compliant with random assignment in most compliant school	66.7	94.1	93.0

Source: NMPED Administrative Data
Sample: Compliance calculations include 638 students (267 treatment and 371 control) who enrolled in an NMPED school at any point during the follow-up period.
Notes: For the treatment group, a student is compliant with random assignment if they enrolled in their associated treatment school at any point during the follow-up period. For the control group, a student is compliant with random assignment if they never enrolled in their associated treatment school at any point during the follow-up period. The most compliant school is the school with the highest total compliance rate and the least compliant school is the school with the lowest total compliance rate.

Findings

Exhibit 7. High School Lottery Study: Impact Estimates

Outcome (%)		Treatment Mean	Control Mean	Difference (Impact)	Standard Error	90 % Confidence interval	Treatment N	Control N
NSC Outcomes								
College Enrollment	**Confirmatory**	**64.7**	**49.6**	**15.1**	**13.7**	**(-7.5, 37.7)**	**283**	**401**
College Persistence	Exploratory	39.9	39.2	0.7	5.5	(-8.3, 9.7)	283	401
College Retention	Exploratory	34.8	33.4	1.4	4.9	(-6.7, 9.4)	283	401
NMPED Administrative Outcomes								
On-time High School Graduation	Exploratory	71.4	68.3	3.1	2.5	(-1.0, 7.2)	266	366
Passed Algebra II PARCC Test	Exploratory	14.2	14.0	0.1	2.0	(-3.2, 3.4)	181	299
Passed Algebra II PARCC Test among those who took the test	*Descriptive*	*21.1*	*18.8*	*2.3*	*3.2*	*(-2.9, 7.6)*	*138*	*224*
ELA Proficiency (grade 10)	Exploratory	24.7	29.4	-4.7	4.8	(-12.6, 3.2)	266	367
ELA Proficiency (grade 10) among those who took the test	*Descriptive*	*43.1*	*45.2*	*-2.1*	*6.7*	*(-13.1, 8.9)*	*155*	*239*
ELA Proficiency (grade 11)	Exploratory	36.3	39.8	-3.5	4.6	(-11.1, 4.1)	267	367
ELA Proficiency (grade 11) among those who took the test	*Descriptive*	*53.0*	*56.4*	*-3.4*	*5.3*	*(-12.2, 5.5)*	*184*	*259*

Source(s): NSC and NMPED Administrative Data
Sample: Non-experimental contrasts (italicized and labeled descriptive) include only students with observed test scores.
Notes: Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%. None of the impacts are significantly different than zero.
All confirmatory and exploratory comparisons in the table are experimental. Students in the sample were not enrolled are included with 0 values as described in the outcomes section above.

Exhibit 8. High School Lottery Study: Variation in Impacts Across Schools

Outcome (%)	Impact of Average School	Variance of Impacts Across School	Standard Error	Predicted Impact of 25th %ile School	Predicted Impact of 75th %ile School
NSC Outcomes					
College Enrollment	**15.1**	**5.7**	**4.3**	**Not Applicable**	**Not Applicable**
College Persistence	0.7	0.5	0.9	Not Applicable	Not Applicable
College Retention	1.4	0.4	1.1	Not Applicable	Not Applicable
NMPED Administrative Outcomes					
On-time High School Graduation	3.1	0.0		Not Applicable	Not Applicable
Passed Algebra II PARCC Test	0.1	0.0	0.0	Not Applicable	Not Applicable
ELA Proficiency (grade 10)	-4.7	0.5	0.5	Not Applicable	Not Applicable
ELA Proficiency (grade 11)	-3.5	0.6	0.5	Not Applicable	Not Applicable

Source(s): NSC and NMPED Administrative Data
Notes: Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%. Precited impacts of 25th and 75th %ile schools are reported for comparisons with statistically significant variation in impacts across schools. None of the comparisons show statistically significant average impacts or variation in impacts. The mixed procedure was not able to calculate the standard error for the variance of impacts across schools.

Exhibit 9. High School Lottery Study: Analysis-specific Attrition Calculations

Outcome	Randomized Treatment N	Randomized Control N	Analysis Treatment N	Analysis Control N	Treatment Attrition	Control Attrition	Overall Attrition	Differential Attrition
NSC Outcomes								
College Enrollment	**297**	**422**	**283**	**401**	**4.7**	**5.0**	**4.9**	**0.3**
College Persistence	297	422	283	401	4.7	5.0	4.9	0.3
College Retention	297	422	283	401	4.7	5.0	4.9	0.3
NMPED Administrative Outcomes								
On-time High School Graduation	297	422	266	366	10.4	13.3	12.1	2.8
Passed Algebra II PARCC Test	197	331	181	299	8.1	9.7	9.1	1.5
ELA Proficiency (grade 10)	297	422	266	367	10.4	13.0	12.0	2.6
ELA Proficiency (grade 11)	297	422	267	367	10.1	13.0	11.8	2.9

Source(s): School Lottery Records, NSC & NMPED Administrative Data

Charter Middle School Lottery Study

All outcomes in this study are constructed from NM administrative data. The analysis sample includes all students who applied to treatment school in an eligible year for whom we observe one or both confirmatory outcomes.

Confirmatory outcomes

- ***ELA achievement (grade 8)***, a continuous z score. We construct the z-score for student i as,

$$z_i = \frac{s_i - \mu}{\sigma},$$

  where s_i is the student's raw score, μ is the mean raw score for all students in NM who took the 8th grade ELA test in that year, and σ is the standard deviation of raw test scores for the same sample of students. The z-score will allow for comparable measurement across years, even though the state fielded several different assessments over the follow-up period.

- ***Passed Algebra I or higher (PARCC test, grade 8)***, a binary indicator that takes on a value of 1 if the student took and passed one of the following PARCC tests (Algebra I, Algebra II, Integrated Math I, Integrated Math II, Integrated Math III, Geometry I) in 8th grade. Students who took the 8th grade math PARCC test (approximately Pre-Algebra level) receive a zero. Sample is limited to cohorts expected to be in 8th grade in 2014-2015 or later so that PARCC scores are available.

Exploratory outcome

- ***Science achievement (grade 7),*** measured as a continuous z-score, as described above. The sample will be restricted to cohorts for whom 7th grade test scores are observed.

All lotteries

The full middle school sample had high overall and differential attrition. As described in our Pre-Analysis Plan, we inspected the attrition rates at the lottery level (school by application year) and dropped the lotteries with high attrition. We dropped lotteries where the difference between the attrition rates for the treatment group and the control group was greater than 18.5 percentage points. This threshold is the highest that results in a sample that meets the WWC attrition standards. We retained at least one lottery from all six schools in the original sample. Out of 21 lotteries, we dropped seven. In the Pre-Analysis Plan, we said that we would not accept a reduction in power of more than 0.5 SD units. This change to the sample meets that criterion. This process was informed by whether outcomes were observed, but not by the actual outcomes for the treatment and control groups. We made the decision to drop lotteries before estimating impacts or calculating treatment and control group means.

Below, we provide the lottery-level attrition data we used to make the decision.

Exhibit 10. All Lotteries Collected for Middle School Charter Lottery Study: Sample Sizes and Sample Formation

Sample	Treatment (N)	Control (N)	Total (N)
Randomized Sample: All applicants	603	2,818	3,421
Analysis Sample: All applicants with observed confirmatory outcomes	517	1,983	2,500
Details of Sample Formation			
Records in Lottery Files	610	2,878	3,488
Duplicate Lottery Records Removed	7	60	67
Randomized Sample: All applicants	603	2,818	3,421
Matched to NMPED Enrollment Data	586	2,681	3,267
Analysis Sample: All applicants with observed confirmatory outcomes	517	1,983	2,500

Source: School Lottery Records & NMPED Administrative Data

Exhibit 11. Middle School Charter Lottery Study: Lottery-level Attrition Calculations

		Randomized		Analysis		Attrition				
School	Application Year	Treatment N	Control N	Treatment N	Control N	Treatment %	Control %	Differential %age Pt	Overall %	Drop Lottery
A	2017	5	12	4	9	20.0	25.0	5.0	23.5	0
B	2010	10	61	9	24	10.0	60.7	50.7	53.5	1
B	2011	12	41	7	28	41.7	31.7	10.0	34.0	0
B	2012	14	49	11	29	21.4	40.8	19.4	36.5	1
B	2013	15	64	11	47	26.7	26.6	0.1	26.6	0
B	2014	11	46	11	33	0.0	28.3	28.3	22.8	1
B	2015	13	48	12	39	7.7	18.8	11.1	16.4	0
B	2016	15	56	13	39	13.3	30.4	17.0	26.8	0
C	2015	35	185	31	129	11.4	30.3	18.8	27.3	1
C	2016	49	197	44	127	10.2	35.5	25.3	30.5	1
C	2017	44	245	42	160	4.5	34.7	30.1	30.1	1
C	2018	54	278	42	177	22.2	36.3	14.1	34.0	0
D	2015	22	175	22	152	0.0	13.1	13.1	11.7	0
D	2016	27	135	27	127	0.0	5.9	5.9	4.9	0
D	2017	39	122	36	112	7.7	8.2	0.5	8.1	0
E	2012	10	14	10	12	0.0	14.3	14.3	8.3	0
E	2014	12	19	9	16	25.0	15.8	9.2	19.4	0
F	2013	50	231	38	158	24.0	31.6	7.6	30.2	0
F	2014	53	241	48	152	9.4	36.9	27.5	32.0	1
F	2015	66	294	49	203	25.8	31.0	5.2	30.0	0
F	2017	47	305	41	210	12.8	31.1	18.4	28.7	0

Source: School Lottery Records & NMPED Administrative Data
Note: The Drop Lottery column contains an indicator that takes on a value of 1 for the lotteries we dropped from the analysis and a value of 0 for the lotteries we retained.

Sample Formation, Baseline Characteristics, and Compliance with Random Assignment

Exhibit 12. Middle School Charter Lottery Study: Sample Sizes

Sample	Treatment (N)	Control (N)	Total (N)
Randomized Sample: Eligible applicants	387	1,794	2,181
Analysis Sample: Eligible applicants with observed confirmatory outcomes	321	1,329	1,650

Source: School Lottery Records & NMPED Administrative Data

Exhibit 13. Middle School Charter Lottery Study: Baseline Equivalence on State Achievement Tests

	Treatment group			Control Group			Treatment - Control	
Measure	Mean	Std. Dev.	Sample Size	Mean	Std. Dev.	Sample Size	Diff. (SE)	Effect Size
Math Achievement (z-score)	0.466	0.949	295	0.358	0.965	1245	0.108* (0.062)	0.112
Reading Achievement (z-score)	0.433	0.977	292	0.405	0.979	1238	0.029 (0.064)	0.030

Source: NMPED Administrative Data
Note: Student baseline test scores were standardized using state-level means and standard deviations for that test, year, and grade-level. Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%.

Exhibit 14. Middle School Charter Lottery Study: Baseline Equivalence on Demographics

	Treatment Group		Control Group		Treatment - Control	
Measure	Mean	Sample Size	Mean	Sample Size	Diff. (SE)	Cox Index Effect Size
White (%)	37.1	321	31.2	1329	5.9** (2.9)	0.051
Hispanic (%)	56.4	321	64.1	1329	-7.7** (3.0)	0.040
Ever EL (%)	20.9	321	22.6	1329	-1.7 (2.6)	-0.034
FRPL	57.0	286	59.7	1236	-2.7 (3.2)	0.011

Source: NMPED Administrative Data
Note: Comparisons for race categories that make up less than 3% of the sample are not reported.
Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%.
Free/reduced price lunch data are only available for students with non-missing baseline test scores. Among students with non-missing test score data, FRLP data is consistently missing for the 2010-2011 school year and frequently missing for the 2013-2014 school year.

Exhibit 15. Middle School Charter Lottery Study: Compliance with Random Assignment

	Treatment (%)	Control (%)	Total (%)
Proportion compliant with random assignment	80.3	84.8	83.9
Proportion compliant with random assignment in least compliant school*	78.9	78.6	78.7
Proportion compliant with random assignment in most compliant school	85.7	96.6	94.5

Source: NMPED Administrative Data
Sample: Compliance calculations include 1,650 students (321 treatment and 1,329 control) in the confirmatory analysis sample.
Note: For the treatment group, a student is compliant with random assignment if they enrolled in their associated treatment school at any point during the follow-up period. For the control group, a student is compliant with random assignment if they never enrolled in their associated treatment school at any point during the follow-up period. The most compliant school is the school with the highest total compliance rate and the least compliant school is the school with the lowest total compliance rate. For this sample, we restrict attention to the schools with more than 10 students in each group when identifying the least compliant schools—this restriction excludes one school.

Findings

Exhibit 16. Middle School Lottery Study: Impact Estimates

Outcome		Treatment Mean	Control Mean	Difference (Impact)	Standard Error	90% Confidence interval	Treatment N	Control N
ELA Achievement (grade 8, z-score)	**Confirmatory**	**0.414**	**0.426**	**-0.012**	**0.071**	**(-0.129, 0.104)**	**320**	**1,323**
Passed Algebra I PARCC test (grade 8, %)	**Confirmatory**	**24.4**	**22.6**	**1.7**	**4.6**	**(-5.8, 9.2)**	**304**	**1,281**
Science Achievement (grade 7, z-score)	Exploratory	0.534	0.400	0.134	0.107	(-0.042, 0.310)	300	1,251

Source(s): NMPED Administrative Data
Notes: Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%. None of the impacts are significantly different than zero. After correcting for multiple comparisons, the 90% confidence interval for ELA achievement is (-0.151, 0.127) and the 90% confidence interval for Algebra II is (-7.2, 10.7).

Exhibit 17. Middle School Lottery Study: Variation in Impacts Across Schools

Outcome (%)	Impact of Average School	Variance of Impacts Across School	Std Error of Variance	Predicted Impact of 25th %ile School	Predicted Impact of 75th %ile School
ELA Achievement (grade 8, z-score)	**-0.012**	0.016	0.330	**Not Applicable**	**Not Applicable**
Passed Algebra I PARCC test (grade 8, %)	**1.7**	0.6	1.4	**Not Applicable**	**Not Applicable**
Science Achievement (grade 7, z-score)	0.134	0.049	1.041	Not Applicable	Not Applicable

Source(s): NMPED Administrative Data
Notes: Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%. Precited impacts of 25th and 75th %ile schools are reported for comparisons with statistically significant variation in impacts across schools. None of the comparisons show statistically significant average impacts or variation in impacts.

Exhibit 18. Middle School Lottery Study: Analysis-specific Attrition Calculations

Outcome	Randomized Treatment N	Randomized Control N	Analysis Treatment N	Analysis Control N	Treatment Attrition	Control Attrition	Overall Attrition	Differential Attrition
ELA Achievement	**387**	**1,794**	**320**	**1,323**	**17.3**	**26.3**	**24.7**	**8.9**
Passed Algebra I	**365**	**1,736**	**304**	**1,281**	**16.7**	**26.2**	**24.6**	**9.5**
Science Achievement	375	1,753	300	1,251	20.0	28.6	27.1	8.6

Source(s): Source: School Lottery Records & NMPED Administrative Data

Charter Elementary School Lottery Study

All outcomes in this study are constructed from NM administrative data. The sample includes all students who applied to a treatment school in an eligible year for whom we observe one or both confirmatory outcomes.

Confirmatory outcomes:

- ***Math achievement (grade 3),*** measured as a continuous z-score, as described for the middle school sample.
- ***ELA achievement (grade 3),*** measured as a continuous z-score**.**

Exploratory outcome:

- ***Science achievement (grade 4),*** measured as a continuous z-score. The sample will be restricted to cohorts for whom 4th grade test scores are observed.

Sample Formation, Baseline Characteristics, and Compliance with Random Assignment

Exhibit 19. Elementary School Charter Lottery Study: Sample Sizes and Sample Formation

Sample	Treatment (N)	Control (N)	Total (N)
Randomized Sample: All applicants	146	500	646
Analysis Sample: All applicants with observed confirmatory outcomes	115	337	452
Details of Sample Formation			
Records in Lottery Files	160	561	721
Duplicate Lottery Records Removed	14	61	75
Randomized Sample: All applicants	146	500	646
Matched to NMPED Enrollment Data	134	431	565
Analysis Sample: All applicants with observed confirmatory outcomes	115	337	452

Source: School Lottery Records & NMPED Administrative Data

Exhibit 20. Elementary School Charter Lottery Study: Baseline Equivalence on Demographics

	Treatment Group		Control Group		Treatment - Control	
Measure	Mean	Sample Size	Mean	Sample Size	Diff. (SE)	Cox Index Effect Size
White (%)	53.0	115	46.3	337	6.8 (5.4)	0.211
Hispanic (%)	40.9	115	47.2	337	-6.3 (5.4)	0.242
Ever EL (%)	0.9	115	3.6	337	-2.7 (1.8)	0.137

Source: NMPED Administrative Data
Note: Comparisons for race categories that make up less than 3% of the sample are not reported.
None of the differences are statistically significant at the 10% level.

Exhibit 21. Elementary School Charter Lottery Study: Compliance with Random Assignment

	Treatment (%)	Control (%)	Total (%)
Proportion compliant with random assignment	74.8	80.1	78.8
Proportion compliant with random assignment in least compliant school	81.6	63.6	69.0
Proportion compliant with random assignment in most compliant school	97.6	82.8	86.4

Source: NMPED Administrative Data
Sample: Compliance calculations include 452 students (115 treatment and 337 control) in the confirmatory analysis sample.
Notes: For the treatment group, a student is compliant with random assignment if they enrolled in their associated treatment school at any point during the follow-up period. For the control group, a student is compliant with random assignment if they never enrolled in their associated treatment school at any point during the follow-up period. The most compliant school is the school with the highest total compliance rate and the least compliant school is the school with the lowest total compliance rate.

Findings

Exhibit 22. Elementary School Lottery Study: Impact Estimates

Outcome		Treatment Mean	Control Mean	Difference (Impact)	Standard Error	90% Confidence interval	Treatment N	Control N
ELA Achievement (grade 3, z-score)	**Confirmatory**	0.609	0.489	0.120	0.145	(-0.118, 0.359)	115	337
Math Achievement (grade 3, z-score)	**Confirmatory**	0.422	0.359	0.063	0.117	(-0.129, 0.256)	115	337
Science Achievement (grade 4, z-score)	Exploratory	0.639	0.640	0.000	0.147	(-0.243, 0.242)	89	273

Source: NMPED Administrative Data
Notes: Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%. None of the impacts are significantly different than zero.
After correcting for multiple comparisons, the 90% confidence interval for ELA achievement is (-0.163, 0.404) and the 90% confidence interval for math achievement is (-0.166, 0.292).

Exhibit 23. Elementary School Lottery Study: Variation in Impacts Across Schools

Outcome (%)	Impact of Average School	Variance of Impacts Across School	Std Error of Variance	Predicted Impact of 25th %ile School	Predicted Impact of 75th %ile School
ELA Achievement (grade 3, z-score)	**0.120**	**0.037**	**0.032**	**Not Applicable**	**Not Applicable**
Math Achievement (grade 3, z-score)	**0.063**	**0.018**	**0.021**	**Not Applicable**	**Not Applicable**
Science Achievement (grade 4, z-score)	0.000	0.024	1.778	Not Applicable	Not Applicable

Source: NMPED Administrative Data
Notes: Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%. Precited impacts of 25th and 75th %ile schools are reported for comparisons with statistically significant variation in impacts across schools. None of the comparisons show statistically significant average impacts or variation in impacts.

Exhibit 24. Elementary School Lottery Study: Analysis-specific Attrition Calculations

Outcome	Randomized Treatment N	Randomized Control N	Analysis Treatment N	Analysis Control N	Treatment Attrition	Control Attrition	Overall Attrition	Differential Attrition
ELA Achievement (grade 3)	146	500	115	337	21.2	32.6	30.0	11.4
Math Achievement (grade 3)	146	500	115	337	21.2	32.6	30.0	11.4
Science Achievement (grade 4)	136	486	89	273	34.6	43.8	41.8	9.3

Source: School Lottery Records & NMPED Administrative Data
Sample: For science analyses, randomized sample sizes are limited to lotteries prior to 2015-2016 so that it is possible to observe 4th grade outcomes during the follow-up period.

Dual Language and Charter School Matching Study

We constructed the matched sample through three steps: (1) identify comparison schools, (2) identify eligible treatment and comparison students, and (3) match comparison group students to treatment group students. This process was undertaken separately for each QED study and sub-study.[7] Study-specific school and student eligibility requirements are discussed in the study sections below.

Identifying eligible schools

Our goal in identifying comparison schools is to capture the alternative options parents consider when enrolling their child in a treatment school. These are the schools that might appear in the control group of an experimental lottery study. Ideally, we would use the same process we used to identify the experimental sample of individuals (described above) and define comparison schools based on the schools attended by lottery applicants. However, the purpose of the quasi-experimental study is to expand the sample of treatment schools beyond those for whom it is possible to identify lottery applicants. Therefore, we also use students who switch schools to identify comparison schools.

For the QED, we identified comparison schools through the following steps for each school:

1. Identified students who ever applied to or enrolled in the treatment school using a combination of school records and state enrollment records. School records included all available lottery records, waitlists, and application records, so long as they could be matched to state administrative data on enrollment. Such records were not available for all treatment schools.

2. For each grade offered by the treatment school, we listed the schools attended by the identified students. This defined the list of candidate schools. This approach followed any student who ever moved across the state. As such, there were candidate schools on the list that could not truly be considered a standard alternative.

3. We selected the schools most commonly attended by identified students. We calculated the proportion of identified students attending that school and included the most commonly attended schools. The threshold we used to define "most commonly attended" varied by treatment school. Our goal was to cover a sufficient number of schools to cover at least 50% of identified students. However, patterns of attendance varied widely across the sample. For one treatment school, we were able to cover 80% of identified students by selecting two comparison schools. These schools tended to be geographically isolated. For other treatment schools, we were included 12 comparison schools and covered 50% of identified students. These schools tended to be in geographically dense locations with a robust charter sector.

As a result of this process, the same treatment school may have a different set of comparison schools at different grades. (This was intentional, as so many treatment schools cover multiple educational levels.) For example, consider a K-8 treatment school that is included in the elementary QED and the middle school QED. For the elementary study, the eligible comparison schools for this treatment school are those that can serve as comparisons for focal grades 1-3. Similarly, for the middle school study, the eligible comparison schools are those that can serve as comparisons for focal grades 7-8.

Moreover, the same school may be identified as an eligible comparison school for multiple treatment schools.

---

[7] For example, matching is performed separately for the main DL study and the DL study focusing exclusively on kindergarten EL students.

DL study. After we identified comparison schools for the five DL schools, we pooled comparison schools for the DL analysis. This approach was necessary because the records we received from several of the DL schools did not allow us to identify students who applied and did not enroll. Therefore, we were only able to identify comparison schools from a relatively few students who switched into or out of the DL school. We believe that this approach is reasonable because the schools in the study are clustered geographically and many schools are identified as comparison schools for several of the DL schools.

Identifying matched comparison group

Separately for each comparison within each study, we perform one-to-one matching without replacement as follows:

1. Randomly order eligible treatment group students. Eligibility requirements for each study sample are described in the section for that study below.

2. For each treatment group student in succession, we identified their matched comparison as follows:

    a. From among the unmatched eligible comparison students, we identified the set of students that (a) attend a comparison school linked to the treatment school, and (b) exactly match on EL status (binary), race/ethnicity (white, Hispanic, native, and other), and eligible for free or reduced price lunch (binary). If no such eligible comparison students exist, these treatment students were dropped from the analysis.

    b. From this restricted set of potential matches, we chose the match with the most similar baseline test scores. We used the Mahalanobis distance metric to judge which potential comparison student has the most similar scores.[8] The choice of metric only affects the middle school study, which includes multiple baseline test scores (math and ELA tests). The other studies have a single baseline test score, in which case there is no difference between the Mahalanobis and Euclidean metrics.

3. We repeated step 2 until all treatment students had been matched or until we ran out of comparison students.

We use 1:1 matching with replacement because this matching approach allows us to associate each comparison group student with a single treatment student. Therefore, we can think of outcome data as nested, with students nested within a matched sample for each treatment school. The nested data structure allows us to estimate the variance in treatment effects across schools, which is consistently found to be important in the charter school literature (Gleason et al, 2010; Unterman, 2017).

Further, the standard argument for matching with replacement is that it reduces the probability of dropping treatment group students when you are unable to find a match. However, the number of potential comparison students is many times larger than the number of treatment students in our case.

Estimating Impacts

Our statistical approach uses a hierarchical linear model with students nested in treatment schools. Comparison students are included in clusters based on the student to whom they are matched. This is required to include a random effect at the treatment school level and estimate variation in impacts at the school level.

We model the outcome y_{is} of individual i in the matched sample for treatment school s as follows:

$$y_{is} = \alpha_0 + \beta_1 T_{is} + \sum_j \gamma_j X_{isj} + \sum_k \delta_k Z_{isk} + \sum_m \lambda_m W_{ism} + \mu_s T_{is} + \nu_s + \varepsilon_{is}$$

where T_{is} is the Treatment group indicator, X_{isj} are school characteristics measured prior to the study period for the school attended by individual i, Z_{isk} are individual baseline characteristics, and W_{ism} are indicators for year of outcome measurement. Individual characteristics include the matching variables and baseline characteristics listed

[8] We will break ties by randomly selecting one comparison student.

below. School characteristics include averages of demographic characteristics and outcome variables for a prior cohort.

The coefficient of the Treatment indicator, β_1, captures the impact of the average charter in our sample—this is the estimate we use to answer our research questions. The regression adjusted control group mean is given by the intercept term, α_0. The coefficients γ_j and δ_k capture the relationships between school and individual characteristics and outcomes. The fixed effects for year of outcome measurement, λ_m, capture differences in tests across years. The random coefficient on Treatment, μ_s, captures variability in impacts across charters in our sample. The school-level random effect, ν_s, captures contextual factors that affect outcomes for all students in the matched sample for school s, including students who do not attend. The individual-level error term, ε_{is}, captures individual-level variation in outcomes not explained by the model.

We intended to use bootstrapping to estimate the standard errors, which would have allowed us to integrate uncertainty from the matching step into the calculation. Although we did develop Stata code to accomplish this, our preliminary bootstrapped analyses took 14 hours to estimate impacts for a single study. We determined that the computational burden of bootstrapping the matching step was too great and we report the standard errors from the estimation step.

Baseline covariates

We include two types of baseline covariates in the model: school characteristics and individual student baseline characteristics. The individual baseline covariates varied across studies according to data availability, as described in Exhibit 25 below.

Exhibit 25. Individual Baseline Covariates in Lottery Study Analyses

Baseline Characteristic	High School	Middle School	Elementary School	Dual Language	Dual Language EL
Race/Ethnicity Indicators (White, Black, Hispanic, Native, Other)	✓	✓	✓	✓	✓
EL Status	✓	✓	✓	✓	
Free Reduced Price Lunch	✓	✓	✓	✓	
Baseline ELA Z Score	✓	✓	✓		
Baseline Math Z Score		✓			
Baseline Math Proficiency Level	✓				
Baseline ACCESS Z Score					✓
Baseline Reading or ACCESS Z Score				✓	

The set of school characteristics include the school level demographic and achievement measures (as listed in the experimental methods section above) for an earlier cohort. We use the most recent cohort of students that does not appear in the analysis sample. For example, if the analysis sample includes students who enter the kindergarten lottery in 2010-2013, we use the students who entered kindergarten in 2009 to define the school characteristics.

Treatment of missing data

We did not impute outcome data, limiting the sample to the individuals for whom we observe outcomes.

We used dummy-variable imputation to address missing individual-level baseline covariates. (Puma et al, 2009)

Multiple comparison corrections

To see why multiple comparison corrections are necessary, imagine that there are no true differences between the treatment and control groups and you run 20 independent hypothesis tests at a 5% level of significance. Statistically, we would expect that you reject the null hypothesis once and draw one false conclusion. Multiple comparisons corrections are designed to limit the extent to which running many tests increases the probability of spurious findings.

We committed in our pre-analysis plan to perform multiple comparison corrections for confirmatory outcomes within each study. Because the high school study has one confirmatory outcome, no multiple comparison correction is required. The elementary and middle school studies have two confirmatory outcomes (math and ELA achievement) require multiple comparison corrections.

We intended to use a multiple comparisons procedure that controlled the family-wise error rate and accounted for correlation across impacts. These tests are more powerful than simpler procedures that assume tests are independent. We planned to implement the Romano & Wolf stepwise procedure introduced in 2005 and refined in 2016 (Romano & Wolf, 2005; Romano & Wolf, 2016), using the Stata package rwolf (Clarke, 2016). However, the rwolf package is not compatible with our multi-level modeling approach.

We therefore used a simple Bonferroni correction. This approach ignores correlation across outcomes and is likely overly conservative. However, given that the QED findings that are significant prior to correction remain significant after Bonferroni correction, the loss of power does not affect the conclusions drawn from the study. We include 90% confidence intervals that reflect corrected p-values in table notes.

Charter High School Matching Study

The treatment sample for this study was selected from students who enrolled in charter high schools. The comparison sample was selected from identified schools that students most commonly attended after applying to or previously attending a treatment school. Our eligible sample is limited to students for which an 8th grade ELA test is observed (in the year or two years prior to enrolling). For our main secondary outcome, the sample is then limited to students that were enrolled in the same school in 9th and 10th grade (if the school offered 9th grade enrollment).

Secondary outcome

- ***College enrollment in the fall after on-time high school graduation.*** Enrollment is a binary outcome defined as described for the Lottery Study samples.

The outcome is well-defined for all students, including those who drop out. This analysis sample focuses on NSC outcomes which consists of students whose NM administrative data provide sufficient information to request NSC data on college enrollment—i.e., we observe name and date of birth. Students for whom we requested data and did not appear in the NSC data file summarizing college enrollment records are assumed to be not enrolled.

Sensitivity Analyses

We run three sensitivity on alternative samples to investigate the robustness of the secondary impact (Exhibit 25, analyses B, C, and D). First, we expand our sample to any student enrolled in a treatment or comparison school in the entry grade. This helps us understand the impact of our enrollment restriction for the secondary outcome sample. Next, we limit our sample to exclude students enrolled in early college high schools to investigate if our results are an artifact of dual enrollment. Lastly, we limit our sample to only include the four charter high schools in the lottery sample. This explores the whether the impact might be attributed to the additional high schools included in the matching study and not included in the lottery study. These analyses maintain the exclusion of charter schools from the comparison group.

In addition, we run four sensitivity analyses designed to determine whether the matching study approximates the lottery study (Exhibit 25, analyses E, F, G, and H). The analysis comparing one year of charter to business as usual, restricting the sample to lottery schools, (analysis H) mimics the lottery analysis as closely as possible. The additional analyses are designed to mimic the main analysis and sensitivity analyses, adding charter schools back into the comparison sample, as the lottery analysis follows applicants to whatever school they enroll in, including charter schools.

Although we originally intended to conduct similar sensitivity analyses for all the matching studies, we chose to focus that investigation on the high school sample for two reasons: (1) the High School Lottery Study had low attrition and (2) the High School Matching Study found a statistically significant impact. These two factors combine to make the comparison of the Lottery and Matching methods particularly interesting and relevant, and these factors are not in place for the other studies. Our conclusion is that the matching study does a reasonable job reproducing the lottery study findings. The impact estimate for the matching study most closely aligned with the lottery study (analysis

H) found a significant impact of 13 percentage points on college enrollment, which is of very similar magnitude to the lottery study estimate of 15 percentage points.

Exhibit 26. Charter High School Matching Study Secondary and Sensitivity Analyses

Analysis		Treatment Schools	Comparison Schools	Student Eligibility
A	2 years of Charter vs. no Charter	10 schools	No Charter	Enrolled in 9th & 10th
Sensitivity analyses exploring impact				
B	2 years of Charter vs. no Charter (No early college high school)	8 schools	No Charter	Enrolled in 9th & 10th
C	1 year of Charter vs. no Charter	10 schools	No Charter	Enrolled in entry grade
D	1 year of Charter vs. no Charter (Lottery Study schools)	Lottery Study schools	No Charter	Enrolled in entry grade
Sensitivity analyses exploring methods				
E	2 years of Charter vs. Business as usual	10 schools	All	Enrolled in 9th & 10th
F	2 years of Charter vs. Business as usual (No early college high school)	8 schools	All	Enrolled in 9th & 10th
G	1 year of Charter vs. Business as usual	10 schools	All	Enrolled in entry grade
H	1 year of Charter vs. Business as usual (Lottery Study schools)	4 schools	All	Enrolled in entry grade

Baseline Balance Before and After Matching

Exhibit 27. High School Charter Matching Study: Baseline Characteristics Before and After Student Matching

		Pre-Match Sample: Eligible Students & Schools				Matched Sample			
Baseline Variable	Variable Type	Treatment Mean	Treatment N	Comparison Mean	Comparison N	Treatment Mean	Treatment N	Comparison Mean	Comparison N
A. 2 years of Charter vs. no Charter									
White	Exact Match	40.0	785	23.7	9,777	40.1	771	40.1	771
Hispanic	Exact Match	51.6	785	63.8	9,777	51.9	771	51.9	771
Native	Exact Match	2.0	785	6.3	9,777	2.1	771	2.1	771
Other	Exact Match	6.4	785	6.3	9,777	6.0	771	6.0	771
EL Status (8th grade)	Exact Match	10.8	785	10.4	9,777	10.4	771	10.4	771
FRPL (8th grade)	Exact Match	43.6	785	58.8	9,777	43.8	771	43.8	771
ELA Test (8th grade; z-score)	Closest Match	0.501	785	0.228	9,777	0.498	771	0.486	771
Math Proficiency (8th grade)	Descriptive (Baseline Balance)	41.4	780	30.1	9,688	41.5	766	37.7	766
C. 1 year of Charter vs. no Charter									
White	Exact Match	39.2	845	22.4	11,198	39.2	817	39.2	817
Hispanic	Exact Match	52.5	845	65.5	11,198	52.9	817	52.9	817
Native	Exact Match	2.0	845	5.9	11,198	2.1	817	2.1	817
Other	Exact Match	6.3	845	6.1	11,198	5.9	817	5.9	817
EL Status (8th grade)	Exact Match	11.5	845	11.6	11,198	11.0	817	11.0	817
FRPL (8th grade)	Exact Match	44.0	845	61.5	11,198	44.6	817	44.6	817
ELA Test (8th grade; z-score)	Closest Match	0.478	845	0.159	11,198	0.471	817	0.456	817
Math Proficiency (8th grade)	Descriptive (Baseline Balance)	40.2	838	27.5	11,094	40.6	810	35.2	812
G. 1 year of Charter vs. Business as usual									
White	Exact Match	39.2	845	22.7	11,497	39.1	818	39.1	818
Hispanic	Exact Match	52.5	845	65.4	11,497	52.9	818	52.9	818
Native	Exact Match	2.0	845	5.8	11,497	2.1	818	2.1	818
Other	Exact Match	6.3	845	6.1	11,497	5.9	818	5.9	818
EL Status (8th grade)	Exact Match	11.5	845	11.7	11,497	11.0	818	11.0	818
FRPL (8th grade)	Exact Match	44.0	845	61.5	11,497	44.5	818	44.5	818
ELA Test (8th grade; z-score)	Closest Match	0.478	845	0.153	11,497	0.472	818	0.451	818
Math Proficiency (8th grade)	Descriptive (Baseline Balance)	40.2	838	27.1	11,388	40.6	811	35.8	813
H. 1 year of Charter vs. Business as usual (Lottery Study schools)									
White	Exact Match	31.8	422	21.3	8,483	31.8	421	31.8	421
Hispanic	Exact Match	60.2	422	68.0	8,483	60.3	421	60.3	421

Baseline Variable	Variable Type	Pre-Match Sample: Eligible Students & Schools				Matched Sample			
		Treatment Mean	Treatment N	Comparison Mean	Comparison N	Treatment Mean	Treatment N	Comparison Mean	Comparison N
Native	Exact Match	1.4	422	4.5	8,483	1.4	421	1.4	421
Other	Exact Match	6.6	422	6.2	8,483	6.4	421	6.4	421
EL Status (8th grade)	Exact Match	14.7	422	12.6	8,483	14.5	421	14.5	421
FRPL (8th grade)	Exact Match	49.8	422	61.8	8,483	49.9	421	49.9	421
ELA Test (8th grade; z-score)	Closest Match	0.509	422	0.118	8,483	0.511	421	0.495	421
Math Proficiency (8th grade)	Descriptive (Baseline Balance)	49.3	420	26.9	8,392	49.2	419	42.0	417

Source: NMPED Administrative Data

Findings

Exhibit 28. High School Charter Matching Study: Impact Estimates

Analysis	Outcome		Treatment Mean	Comparison Mean	Difference (Impact)	Standard Error	90% Confidence interval	Treatment N	Control N
A.	**College Enrollment**	**Secondary**	63.8	57.7	6.1*	3.2	(0.8, 11.3)	771	771
Sensitivity analyses exploring impact									
B	College Enrollment	Exploratory	62.6	57.4	5.2	3.4	(-0.4, 10.8)	725	725
C	College Enrollment	Exploratory	61.2	55.3	5.8*	3.0	(0.9, 10.8)	817	817
D	College Enrollment	Exploratory	63.5	58.3	5.1	8.3	(-8.5, 18.8)	420	420
Sensitivity analyses exploring methods									
E	College Enrollment	Exploratory	64.0	57.3	6.6**	3.1	(1.6, 11.7)	771	771
F	College Enrollment	Exploratory	63.5	56.8	6.6**	3.0	(1.6, 11.6)	725	725
G	College Enrollment	Exploratory	63.1	55.0	8.1***	3.0	(3.1, 13.0)	818	818
H	College Enrollment	Exploratory	71.7	58.4	13.3**	6.7	(2.2, 24.4)	421	421

Source: NMPED Administrative Data
Notes: Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%.

Exhibit 29. High School Charter Matching Study: Variation in Impacts Across Schools

Analysis	Outcome	Impact of Average School	Variance of Impacts Across School	Std Error of Variance	Predicted Impact of 25th %ile School	Predicted Impact of 75th %ile School
A	**College Enrollment**	6.1*	0.2	0.3	**Not Applicable**	**Not Applicable**
Explore Source of Impact						
B	College Enrollment	5.2	0.1	0.2	Not Applicable	Not Applicable
C	College Enrollment	5.8*	0.1	0.3	Not Applicable	Not Applicable
D	College Enrollment	5.1	0.0		Not Applicable	Not Applicable
Within-study Comparisons						
E	College Enrollment	6.6**	0.2	0.3	Not Applicable	Not Applicable
F	College Enrollment	6.6**	0.1	0.2	Not Applicable	Not Applicable
G	College Enrollment	8.1***	0.2	0.3	Not Applicable	Not Applicable
H	College Enrollment	13.3**	0.0	0.0	Not Applicable	Not Applicable

Source: NMPED Administrative Data
Notes: Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%. Precited impacts of 25th and 75th %ile schools are only reported for comparisons with statistically significant variation in impacts across schools. None of the comparisons show statistically significant variation in impacts.

Exhibit 30. Charter High School Matching Study: Analysis-specific Baseline Balance

Outcome	Baseline Measure	Treatment Mean	Comparison Mean	Difference (Impact)	Standard Error	90% Confidence interval	Treatment N	Control N
A. 2 years of Charter vs. no Charter								
College Enrollment	**ELA Test (8th grade; z-score)**	0.498	0.486	0.012	0.052	(-0.089, 0.114)	771	771
College Enrollment	**Math Proficiency (8th grade; %)**	41.5	37.7	3.8	2.5	(-1.1, 8.7)	766	766
B. 2 years of Charter vs. no Charter (No early college high school)								
College Enrollment	ELA Test (8th grade; z-score)	0.501	0.493	0.008	0.054	(-0.098, 0.114)	725	725
College Enrollment	Math Proficiency (8th grade)	37.8	33.9	3.8	2.5	(-1.1, 8.7)	720	722
C. 1 year of Charter vs. no Charter								
College Enrollment	ELA Test (8th grade; z-score)	0.471	0.456	0.015	0.051	(-0.084, 0.115)	817	817
College Enrollment	Math Proficiency (8th grade)	40.6	35.2	5.4**	2.4	(0.7, 10.1)	810	812
D. 1 year of Charter vs. no Charter (Lottery Study schools)								
College Enrollment	ELA Test (8th grade; z-score)	0.509	0.498	0.011	0.074	(-0.134, 0.155)	420	420
College Enrollment	Math Proficiency (8th grade)	49.3	42.3	7.0**	3.4	(0.2, 13.7)	418	416
E. 2 years of Charter vs. Business as usual								
College Enrollment	ELA Test (8th grade; z-score)	0.498	0.484	0.014	0.052	(-0.088, 0.115)	771	771
College Enrollment	Math Proficiency (8th grade; %)	41.5	37.7	3.8	2.5	(-1.1, 8.7)	766	767
F. 2 years of Charter vs. Business as usual (No early college high school)								
College Enrollment	ELA Test (8th grade; z-score)	0.501	0.491	0.010	0.054	(-0.096, 0.115)	725	725
College Enrollment	Math Proficiency (8th grade)	37.8	33.9	3.9	2.5	(-1.0, 8.8)	720	723
G. 1 year of Charter vs. Business as usual								
College Enrollment	ELA Test (8th grade; z-score)	0.472	0.452	0.021	0.051	(-0.079, 0.120)	818	818

Outcome	Baseline Measure	Treatment Mean	Comparison Mean	Difference (Impact)	Standard Error	90% Confidence interval	Treatment N	Control N
College Enrollment	Math Proficiency (8th grade)	40.6	35.8	4.8**	2.4	(0.1, 9.5)	811	813
H. 1 year of Charter vs. Business as usual (Lottery Study schools)								
College Enrollment	ELA Test (8th grade; z-score)	0.511	0.495	0.016	0.074	(-0.129, 0.161)	421	421
College Enrollment	Math Proficiency (8th grade)	49.2	42.0	7.2**	3.4	(0.5, 13.9)	419	417

Source: NMPED Administrative Data
Notes: Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%. None of the impacts are significantly different than zero.

Charter Middle School Matching Study

The treatment sample for this study was selected from students who enrolled in charter middle schools in 6th or 7th grade. The comparison sample was selected from identified schools that students most commonly attended after applying to or previously attending a treatment school. Our eligible sample is limited to students for which an 5th grade ELA test and a 5th grade math test is observed. Our sample is also limited to students for which an 8th grade ELA test and 8th grade math test can be observed.

Secondary outcomes

All outcomes in this study are constructed from NM administrative data.

- ***ELA achievement (grade 8)***, a continuous z score, constructed as described for the Lottery Study samples.
- ***Passed Algebra I PARCC test (grade 8),*** a binary indicator that takes on a value of 1 if the student took and passed one of the following PARCC tests (Algebra I, Algebra II, Integrated Math I, Integrated Math II, Integrated Math III, Geometry I) in 8th grade. Students who took the 8th grade math PARCC test (approximately Pre-Algebra level) receive a zero. Sample is limited to cohorts expected to be in 8th grade in 2014-2015 or later so that PARCC scores are available.

Baseline Balance Before and After Matching

Exhibit 31. Middle School Charter Matching Study: Baseline Characteristics Before and After Student Matching

		Pre-Match Sample: Eligible Students & School				Matched Sample			
Baseline Variable	Variable Type	Treatment Mean	Treatment N	Comparison Mean	Comparison N	Treatment Mean	Treatment N	Comparison Mean	Comparison N
White	Exact Match	32.0	1,524	21.9	28,584	32.0	1,524	32.0	1,524
Hispanic	Exact Match	64.3	1,524	69.8	28,584	64.3	1,524	64.3	1,524
Native	Exact Match	1.5	1,524	3.6	28,584	1.5	1,524	1.5	1,524
Other	Exact Match	2.2	1,524	4.7	28,584	2.2	1,524	2.2	1,524
EL Status (5th grade)	Exact Match	10.7	1,524	9.8	28,584	10.7	1,524	10.7	1,524
FRPL (5th grade)	Exact Match	54.8	1,524	67.0	28,584	54.8	1,524	54.8	1,524
ELA Test (5th grade; z-score)	Closest Match	0.328	1,524	0.111	28,584	0.328	1,524	0.329	1,524
Math Test (5th grade; z-score)	Closest Match	0.292	1,524	0.103	28,584	0.292	1,524	0.294	1,524

Source: NMPED Administrative Data

Findings

Exhibit 32. Middle School Charter Matching Study: Impact Estimates

Outcome		Treatment Mean	Comparison Mean	Difference (Impact)	Standard Error	90% Confidence interval	Treatment N	Control N
ELA Achievement (grade 8, z-score)	**Secondary**	0.381	0.272	0.109	0.069	(-0.005, 0.223)	1,524	1,524
Passed Algebra I PARCC test (grade 8, %)	**Secondary**	15.0	19.2	-4.2	3.8	(-10.5, 2.1)	1,524	1,524

Source: NMPED Administrative Data
Notes: Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%. Statistical significance for two-sided tests corrected for multiple comparisons are indicated with plus signs, as follows: $^{+++}$ = 1%; $^{++}$ = 5%; $^{+}$ = 10%. None of the impacts are significantly different than zero. After correcting for multiple comparisons, the 90% confidence interval for ELA achievement is (-0.026, 0.245) and for math achievement is (-11.7, 3.4).

Exhibit 33. Middle School Charter Matching Study: Variation in Impacts Across Schools

Outcome	Impact of Average School	Variance of Impacts Across School	Std Error of Variance	Predicted Impact of 25th %ile School	Predicted Impact of 75th %ile School
ELA Achievement (grade 8, z-score)	0.109	0.046**	0.021	0.095	0.124
Passed Algebra I PARCC test (grade 8, %)	-4.2	1.5**	0.711	-4.659	-3.700

Source: NMPED Administrative Data
Notes: Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%. Precited impacts of 25th and 75th %ile schools are only reported for comparisons with statistically significant variation in impacts across schools. None of the comparisons show statistically significant variation in impacts.

Exhibit 34. Middle School Charter Matching Study: Analysis-specific Baseline Balance

Outcome	Baseline Measure	Treatment Mean	Comparison Mean	Difference (Impact)	Standard Error	90% Confidence interval	Treatment N	Control N
ELA Achievement (grade 8, z-score)	**ELA Test (5th grade; z-score)**	0.328	0.329	-0.001	0.036	(-0.072, 0.069)	1,524	1,524
ELA Achievement (grade 8, z-score)	**Math Test (5th grade; z-score)**	0.292	0.294	-0.001	0.037	(-0.074, 0.071)	1,524	1,524
Passed Algebra I PARCC test (grade 8, %)	**ELA Test (5th grade; z-score)**	0.328	0.329	-0.001	0.036	(-0.072, 0.069)	1,524	1,524
Passed Algebra I PARCC test (grade 8, %)	**Math Test (5th grade; z-score)**	0.292	0.294	-0.001	0.037	(-0.074, 0.071)	1,524	1,524

Source: NMPED Administrative Data
Notes: Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%. None of the impacts are significantly different than zero.

Charter Elementary School Matching Study

The treatment sample for this study was selected from students who enrolled in charter middle schools in kindergarten. The comparison sample was selected from identified schools that students most commonly attended after applying to or previously attending a treatment school. Our eligible sample is limited to students for which a kindergarten DIBELS test score is observed. Our sample is also limited to students for which a 3rd grade ELA test and 3rd grade math test can be observed.

Secondary outcomes

All outcomes in this study are constructed from NM administrative data.

- ***ELA achievement (grade 3)***, a continuous z score, constructed as described for the Lottery Study samples.
- ***Math achievement (grade 3)***, a continuous z score, constructed as described for the Lottery Study samples.

Baseline Balance Before and After Matching

Exhibit 35. Elementary School Charter Matching Study: Baseline Characteristics Before and After Student Matching

		Pre-Match Sample: Eligible Students & School				Matched Sample			
Baseline Variable	Variable Type	Treatment Mean	Treatment N	Comparison Mean	Comparison N	Treatment Mean	Treatment N	Comparison Mean	Comparison N
White	Exact Match	35.4	240	25.7	4,742	35.4	240	35.4	240
Hispanic	Exact Match	57.9	240	64.7	4,742	57.9	240	57.9	240
Native	Exact Match	2.9	240	5.3	4,742	2.9	240	2.9	240
Other	Exact Match	3.8	240	4.3	4,742	3.8	240	3.8	240
EL Status (Kindergarten)	Exact Match	3.8	240	6.6	4,742	3.8	240	3.8	240
FRPL (Kindergarten)	Exact Match	60.4	240	61.8	4,742	60.4	240	60.4	240
DIBELS (Kindergarten; z-score)	Closest Match	0.126	240	0.075	4,742	0.126	240	0.133	240

Source: NMPED Administrative Data
Notes: Free or reduced price lunch data is missing for students who took the DIBELs test in the 2013-2014 school year. For those students, we use 1st grade FRPL instead.

Findings

Exhibit 36. Elementary School Charter Matching Study: Impact Estimates

Outcome		Treatment Mean	Comparison Mean	Difference (Impact)	Standard Error	90% Confidence interval	Treatment N	Control N
ELA Achievement (grade 3, z-score)	**Secondary**	0.115	0.176	-0.061	0.087	(-0.204, 0.082)	239	240
Math Achievement (grade 3, z-score)	**Secondary**	-0.035	0.168	-0.202**+	0.090	(-0.350, -0.054)	240	240

Source: NMPED Administrative Data
Notes: Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%. Statistical significance for two-sided tests corrected for multiple comparisons are indicated with plus signs, as follows: +++ = 1%; ++ = 5%; + = 10%. After correcting for multiple comparisons, the 90% confidence interval for ELA achievement is (-0.231, 0.109) and for math achievement is (-0.379, -0.025).

Exhibit 37. Elementary School Charter Matching Study: Variation in Impacts Across Schools

Outcome	Impact of Average School	Variance of Impacts Across School	Std Error of Variance	Predicted Impact of 25th %ile School	Predicted Impact of 75th %ile School
ELA Achievement (grade 3, z-score)	-0.061	0.000	0.000		
Math Achievement (grade 3, z-score)	-0.202**	0.000			

Source: NMPED Administrative Data
Notes: Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%. Precited impacts of 25th and 75th %ile schools are only reported for comparisons with statistically significant variation in impacts across schools. None of the comparisons show statistically significant variation in impacts.
The standard error of the variance was not statistically identified for the math achievement analysis.

Exhibit 38. Elementary School Charter Matching Study: Analysis-specific Baseline Balance

Outcome	Baseline Measure	Treatment Mean	Comparison Mean	Difference (Impact)	Standard Error	90% Confidence interval	Treatment N	Control N
ELA Achievement (grade 3, z-score)	ELA Test (Kindergarten; z-score)	0.126	0.139	-0.013	0.076	(-0.163, 0.137)	240	239
Math Achievement (grade 3, z-score)	ELA Test (Kindergarten; z-score)	0.126	0.133	-0.007	0.076	(-0.157, 0.143)	240	240

Source: NMPED Administrative Data
Notes: Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%. None of the impacts are significantly different than zero.

Dual Language Matching Study

This study compares students enrolled in dual language schools from 1st to 5th grade to students who were enrolled in schools that did not offer dual language from 1st to 5th grade. (We used the NMPED Bilingual Multicultural Education 2017-2018 Annual Report to identify schools that offered dual language and excluded them from the comparison school sample.) Our eligible sample is limited to students for which a 5th grade ELA test score is observed. The sample is then limited to students that were enrolled in the same school in 1st grade as they were in the 5th grade as a measure of continuous enrollment. Lastly, our overall study sample is limited to students with a valid baseline test score.

Overall DL study sample

The overall sample for the DL matching study is very small. Although the NMPED required that kindergarten and first graders take a reading test starting in the 2013-2014 school year, most dual language schools fielded the Spanish version of the test and those scores are not available in the NMPED school accountability data system. Because matching requires a test score, the sample of students never identified as English learners is limited to the students for whom we observe a first grade English language test score. We are able to expand the sample slightly by using the first grade ACCESS score, an assessment tracking progress for English learners, to match English learners. However, the analysis only includes 90 students enrolled in dual language.

English learner sample

This study focuses on understanding the effect of dual language schools on students expected to receive the full benefit of the program, those enrolled from first to fifth grade. These students are compared to students who were enrolled in a single, English language school over the same key grades. We committed to this analytic approach in our pre-analysis plan because prior studies of dual language found positive effects in 5th grade and because experts on dual language told us that they would not expect to see impacts for students who were not enrolled for the five years. However, at the analysis phase, we found that the enrollment requirement excludes most English learners who enrolled in focal schools.

As shown in Exhibit 39, current or former English learners receive a much lower dose of dual language than students never identified as Engligh learners. Among students who were enrolled in 5th grade in one of the DL treatment schools from 2014-2015 to 2018-2019, 75% of students who were never English learner were enrolled in that school for 5 years or more compared to 38% of current or former English learners. The difference in average years of enrollment in dual language between current or former English learners and students never identified as English learners is statistically significant.

Among the 441 current or former English learners who enrolled in the dual language schools over the follow up period, only 80 were enrolled in the same dual language school in first and fifth grade and were identified as English learners in kindergarten. Only 71 English learners enrolled in the same dual language school in first and fifth grade and the test scores required for inclusion in analysis. The 71 dual language English learners included in the analysis are not typical of English learners in the dual language schools. In addition, the very small sample size means that error bands will be very wide, making it difficult to detect impacts.

Exhibit 39. Dual Language Enrollment by English Learner Status

Sample	Number of Students	Number of School Years with Any Enrollment in DL School						
		1	2	3	4	5	6	7
Current or Former English Learners								
Enrolled in a DL treatment school	441	28.6	19.0	14.1	11.8	8.6	16.8	1.1
Enrolled in the DL treatment school in 5th grade	268	17.5	16.8	14.6	13.4	8.2	27.6	1.9
Identified as EL in Kindergarten and enrolled in the DL treatment school in 5th grade	169	16.6	14.2	10.1	11.2	6.5	38.5	3.0
Identified as EL in Kindergarten and enrolled in the same DL treatment school in 1st and 5th grade	80	0.0	0.0	0.0	0.0	12.5	81.3	6.3
Never identified as English Learners								
Enrolled in a DL treatment school	966	19.2	10.6	9.6	10.1	10.8	38.3	1.4
Enrolled in the DL treatment school in 5th grade	588	6.6	4.6	6.0	7.5	10.0	62.9	2.4
Enrolled in the same DL treatment school in 1st and 5th grade	441	0.0	0.0	0.0	0.0	12.9	83.9	3.2

Source: NMPED Administrative Data

Secondary outcome (All Students)

- ***ELA achievement (grade 5),*** measured as a continuous z-score, constructed as described for the Lottery Study samples.

Exploratory outcomes (Kindergarten English Learners)

- ***ELA achievement (grade 5),*** measured as a continuous z-score, constructed as described for the Lottery Study samples.

- **Redesignated Fluent English Proficient (grade 5),** measured as a binary indicator. Kindergarten EL students no longer designated as ELs in 5th grade are assumed to have been reclassified as fluent English proficient.

Baseline Balance Before and After Matching

Exhibit 40. Dual Language Matching Study: Baseline Characteristics Before and After Student Matching

		Pre-Match Sample: Eligible Students & School				Matched Sample			
Baseline Variable	Variable Type	Treatment Mean	Treatment N	Comparison Mean	Comparison N	Treatment Mean	Treatment N	Comparison Mean	Comparison N
All Students									
White	Exact Match	18.9	90	24.6	1,641	18.9	90	18.9	90
Hispanic	Exact Match	72.2	90	65.1	1,641	72.2	90	72.2	90
Native	Exact Match	0.0	90	4.2	1,641	0.0	90	0.0	90
Other	Exact Match	8.9	90	6.1	1,641	8.9	90	8.9	90
EL Status (1st grade)	Exact Match	30.0	90	16.8	1,641	30.0	90	30.0	90
FRPL (1st grade)	Exact Match	45.2	62	53.8	1,144	45.2	62	45.2	62
DIBELS or ACCESS test (1st grade; z-score)	Closest Match	0.081	90	0.139	1,641	0.081	90	0.116	90
DIBELS Test (Non-EL; 1st grade; z-score)	Descriptive (Baseline Balance)	0.104	63	0.137	1,366	0.104	63	0.102	63
ACCESS Test (EL; 1st grade; z-score)	Descriptive (Baseline Balance)	0.027	27	0.149	275	0.027	27	0.148	27
Kindergarten English Learners									
White	Exact Match	2.8	71	3.5	631	2.8	71	2.8	71
Hispanic	Exact Match	95.8	71	78.6	631	95.8	71	95.8	71
Native	Exact Match	0.0	71	6.0	631	0.0	71	0.0	71
Other	Exact Match	1.4	71	11.9	631	1.4	71	1.4	71
ACCESS Test (Kindergarten; z-score)	Closest Match	0.241	71	0.211	631	0.241	71	0.257	71

Source: NMPED Administrative Data
Notes: Free or reduced price lunch data is missing for students in the overall sample who didn't take the DIBELs test or who took the DIBELs test in the 2013-2014 school year. Students who were missing FRPL data were matched to other students who were missing FRPL data.
Free or reduced price lunch data is nearly universally missing for the kindergarten EL sample and is therefore not used to match.

Findings

Exhibit 41. Dual Language Matching Study: Impact Estimates

Outcome		Treatment Mean	Comparison Mean	Difference (Impact)	Standard Error	90% Confidence interval	Treatment N	Control N
All Students								
ELA Achievement (grade 5, z-score)	**Secondary**	0.348	0.164	0.183	0.305	(-0.321, 0.688)	90	90
Kindergarten English Learners								
ELA Achievement (grade 5, z-score)	Exploratory	-0.118	-0.304	0.187	0.142	(-0.048, 0.422)	71	71
Redesignated Fluent English Proficient (grade 5, %)	Exploratory	41.0	36.6	4.4	7.1	(-7.4, 16.1)	71	71

Source: NMPED Administrative Data
Notes: Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%. None of the impacts are significantly different than zero.

Exhibit 42. Dual Language Matching Study: Variation in Impacts Across Schools

Outcome (%)	Impact of Average School	Variance of Impacts Across School	Std Error of Variance	Predicted Impact of 25th %ile School	Predicted Impact of 75th %ile School
All Students					
ELA Achievement (grade 5, z-score)	**0.183**	**0.000**	**0.000**	**Not Applicable**	**Not Applicable**
Kindergarten English Learners					
ELA Achievement (grade 5, z-score)	0.187	0.000	0.000	Not Applicable	Not Applicable
Redesignated Fluent English Proficient (grade 5, %)	4.4	0.0		Not Applicable	Not Applicable

Source: NMPED Administrative Data
Notes: Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%. Precited impacts of 25th and 75th %ile schools are only reported for comparisons with statistically significant variation in impacts across schools. None of the comparisons show statistically significant average impacts or variation in impacts.
The standard error of the variance was not statistically identified for the Redesignated Fluent English Proficient analysis.

Exhibit 43. Elementary School Charter Matching Study: Analysis-specific Baseline Balance

Outcome	Baseline Measure	Treatment Mean	Comparison Mean	Difference (Impact)	Standard Error	90% Confidence interval	Treatment N	Control N
All Students								
ELA Achievement (grade 5, z-score)	**DIBELS or ACCESS test (1st grade; z-score)**	0.081	0.116	-0.035	0.158	(-0.348, 0.277)	90	90
Kindergarten English Learners								
ELA Achievement (grade 5, z-score)	ACCESS test (kindergarten; z-score)	0.241	0.257	-0.016	0.134	(-0.280, 0.248)	71	71
EL Status (grade 5, %)	ACCESS test (kindergarten; z-score)	0.241	0.257	-0.016	0.134	(-0.280, 0.248)	71	71

Source: NMPED Administrative Data
Notes: Statistical significance levels for two-sided tests are indicated with asterisks, as follows: *** = 1%; ** = 5%; * = 10%. None of the impacts are significantly different than zero.