Blueprint Labs

Discussion Paper #2020.10

Charter Schools and Suspensions: Evidence from Massachusetts Chapter 222

Mayara Felix

October 2020

MIT Department of Economics
77 Massachusetts Avenue, Bldg. E53-390
Cambridge, MA 02139

National Bureau of Economic Research
1050 Massachusetts Avenue, 3rd Floor
Cambridge, MA 02138

Charter Schools and Suspensions: Evidence from Massachusetts Chapter 222
Mayara Felix
Blueprint Labs Discussion Paper #2020.10
October 2020

ABSTRACT

I evaluate the impact of Massachusetts Chapter 222—a policy that limited charter schools' ability to suspend students—on student suspensions and test scores. Comparing charter attendance effects before vs. after Chapter 222, I find that Chapter 222 reduced charter suspensions by roughly 10 percentage points, but had no impact on charter learning. I then use variation in lottery offers and applicants' pre-lottery suspensions to separate the effect of suspensions from that of charter attendance on test scores. Suspensions appear to be unrelated to achievement in charters, while the causal effect of charter attendance on test scores is large and positive.

*MIT. Email: mfelix@mit.edu. I am indebted to Josh Angrist and Parag Pathak for invaluable guidance, and to Alberto Abadie, Chris Ackerman, Isaiah Andrews, Ben Olken, Sarah Cohodes, Esther Duflo, Helen Ho, Peter Hull, David Martin, Elizabeth Setren, Camille Terrier, and participants of the MIT Economics Labor Lunch and Development Tea for valuable feedback. Anran Li provided expert research assistance. School Effectiveness and Inequality Initiative (SEII) Assistant Director Eryn Heying provided invaluable administrative support. I am especially grateful to Carrie Conaway, Matt Deninger, Alison Bagg, Pierre Lucien, and staff from the Massachusetts Department of Elementary and Secondary Education for providing access to administrative student data.

1 Introduction

Student suspensions in charter schools are common yet controversial. In Massachusetts, the debate over school suspensions has centered around urban charter schools, which increase students' test scores but suspend more often than traditional public schools ("TPS") (Angrist et al., 2013). Suspensions are most prevalent in grades 5–8, most suspended students are Black or Hispanic, and suspended students are typically removed from school for a day.[1] Yet there is no evidence on whether charter suspensions harm, improve, or have no effect on student learning.

This paper leverages Massachusetts Chapter 222, a policy that limited charters' use of suspensions, to estimate the effect of suspensions on charter students' test scores. Chapter 222 was signed in August 2012 and took effect in school year 2015.[2] Under the policy, principals are required to take several steps before suspending or expelling a student, such as sending written notifications to parents and meeting with parents to discuss the circumstances that led to the suspension. Schools must also ensure that students who were excluded from school for disciplinary reasons can make academic progress during the classroom removal period, a requirement that previously applied only to students with special needs.

I analyze the effect of Chapter 222 using a Difference-in-Differences Instrumental Variables empirical strategy. Specifically, I compare the outcomes for charter vs. TPS students before and after Chapter 222 in a sample of Boston charter middle school applicants, where charter attendance is randomly assigned via lottery. I find that by the end of school year 2017 Chapter 222 reduced the causal effect of charter attendance on suspensions by 10 percentage points, nearly halving the pre-Chapter 222 gap in suspensions between charters and TPS. In contrast, the policy had no sizable or statistically significant effect on charter math test scores.

To understand how Chapter 222 reduced suspensions without affecting test scores, I use variation in lottery offers from charters of varying disciplinary environments, and heterogeneity in applicants' pre-lottery suspensions, to separately identify the effects on test scores of charter suspensions vs. those of charter attendance. Consistent with the observed impact of Chapter 222, the causal effect of suspensions on charter students' test scores is zero. Conversely, the causal effect of charter attendance on suspended students' test scores is positive, large, and similar to the effect on non-suspended students.

[1] See Appendix Figures A3 and A4 for a breakdown of suspension rates.

[2] Throughout, I refer to school years after their spring semester year (e.g., school year 2015 refers to Fall 2014 and Spring 2015).

My findings contribute to a large literature on the effects of charter attendance. Many lottery-based studies have documented large positive effects of charter attendance on test scores—see, for example, Hoxby and Murarka (2009); Dobbie and Fryer Jr (2011); Angrist et al. (2010, 2012, 2016); Abdulkadiroğlu et al. (2011); Setren (2017)—with the largest test score gains in this literature come from *No Excuses* charters (Chabrier et al., 2016). *No Excuses* charters are characterized by an "emphasis on discipline, school uniforms, cold-calling, strict adherence to school-wide standards, and the use of Teach For America alumni" (Angrist et al., 2013).[3] And while a large body of evidence shows that Boston's *No excuses* charters significantly improve learning (Angrist et al., 2016; Abdulkadiroğlu et al., 2016, 2011), they also suspend more. As OLS effects of suspensions on test scores are negative, this raises the question of whether *No excuses* Boston charters' success is because of, in spite of, or unrelated to high suspension rates. This paper addresses this knowledge gap.

2 Empirical strategy

I estimate the effect of Chapter 222 on student outcomes using a Difference-in-Differences Instrumental Variables ("DD-IV") approach. This approach compares the outcomes of charter to TPS students, before and after Chapter 222 took effect, in a sample of charter school applicants—where charter attendance is randomly assigned via lottery. The effect of Chapter 222 on student outcome Y_{it} is coefficient γ in the following second stage regression:

$$Y_{it} = \quad \alpha + \beta D_{it} + \gamma \left[D_{it} \times 1_{\{t>t^*\}} \right] + \zeta' X_i + \delta_t + \delta_{g(i,t)} + \epsilon_{it}, \tag{1}$$

where D_{it} is a dummy indicating whether charter applicant i was enrolled in a charter in school year t;[4] $1_{\{t>t^*\}}$ is a dummy indicating whether year t is after Chapter 222's effective year; δ_t and $\delta_{g(i,t)}$ are year and grade fixed effects; and X_i is a vector of applicant-level demographics and baseline grade covariates, including a fixed effect for application year and a fixed effect for the set of charter schools to which i applied (the applicant's "risk set"). Conditioning on risk sets is necessary because the probability of winning any charter lottery depends on the set of charters to which the applicant applies; controlling for student baseline covariates X_i reduces the variance of point estimates.

Since charter attendance is itself a treatment, γ can also be interpreted as Chapter 222's

[3] More specifically, these five variables are most predictive of a school self-identifying as *No Excuses*.

[4] As in Angrist et al. (2013), I define a student to be enrolled in a charter for the whole school year even if the student only attended the charter for a single day in that year.

impact on the charter attendance treatment effect. The first stage regressions for charter attendance before and after Chapter 222 are

$$D_{it} = \theta + \iota Z_i + \kappa \left[Z_{it} \times 1_{\{t>t^*\}} \right] + \nu' X_i + \lambda_t + \lambda_{g(i,t)} + \mu_{it}, \quad (2)$$

$$D_{it} \times 1_{\{t>t^*\}} = \xi + \pi Z_i + \rho \left[Z_{it} \times 1_{\{t>t^*\}} \right] + \varphi' X_i + o_t + o_{g(i,t)} + \upsilon_{it}, \quad (3)$$

where Z_i is a dummy for whether applicant i received a lottery offer from any charter; and λ and o coefficients are the same set of fixed effects as in Equation 1. The key assumption for a causal interpretation of γ is that potential outcomes of charter vs. TPS students would have followed parallel trends but for Chapter 222, or equivalently, that the charter attendance *effect* would have remained constant after Chapter 222 had the policy not taken effect.

3 Data and results

I implement the DD-IV empirical strategy from Section 2 by linking administrative data on student enrollment, demographics, test scores, and disciplinary records to the list of Boston charter middle school lottery applicants for cohort years 2005–2014. These students were in middle school grades between 2006 and 2017. Appendix Table A14 lists the charter schools and cohort years in the sample. The Data Appendix describes each of the data sources and linking procedures, which follow Setren (2017) and Angrist et al. (2016, 2013).

In order for the exercise's results to be interpreted as the casual impact of Chapter 222, the potential outcomes of treated vs. untreated students must follow parallel trends. Figure 1 presents a visual check that parallel trends in student outcomes does indeed hold prior to Chapter 222. It plots year-by-year estimates of charter attendance effects—relative to school year 2012—on a dummy for whether a student is ever suspended (in-school or out-of school). Despite year-to-year variation, the charter attendance effect prior to Chapter 222's was not statistically different from the baseline year 2012, displaying no pre-trends. However, the charter attendance effect on suspensions starts to decline in 2013, the first school year following Chapter 222's signing. On levels, Appendix Table A7 shows that up to 2012 applicants who attended charters by virtue of winning the lottery were on average 22 percentage points more likely to be suspended out-of-school than lottery-losing counterparts attending TPS. By 2017 I cannot reject that the charter attendance effect on the probability of a suspension was zero (or, alternatively, 22 percentage points lower than the 2012 estimate), suggesting that Chapter 222 closed the charter vs. TPS gap in out-of-school suspension probability

within five years of its signing.

Figure 2 replicates the exercise in Figure 1 for math test scores, also showing no pre-trends. However, in contrast to the effect of Chapter 222 on charter suspensions, Figure 2 shows that that Chapter 222 had no statistically significant impact on math test scores at charters. Instead, as shown in Appendix Table A7, the charter attendance effect on math remained steadily large and positive at 0.566 standard deviations throughout the period.[5]

Table 1 presents a formal quantification of Chapter 222's effect. Columns (1) and (2) report estimates of β and γ, respectively, from Equation 1 using Chapter 222's signature year as the key policy year ($t^* = 2012$). Column (3) replicates this exercise using the year Chapter 222 took effect as the key policy year instead ($t^* = 2014$).[6] These charter attendance effects are consistent with lottery-based charter attendance effects reported elsewhere in the literature. Attending a charter school by virtue of winning the lottery caused a 20.5% increase in a student's probability of suspension (out-of- or in-school) at any point in their middle school years relative to lottery applicants who attended TPS. Lottery-induced charter attendees also experienced substantial average gains in test scores: 0.591 standard deviations for math and 0.348 standard deviations for English.[7]

Overall, Chapter 222 reduced the probability of being suspended at a charter school by 9.5 percentage points, with no statistically significant effect on math test scores. Table 1 also shows that Chapter 222's impact on charter discipline was primarily driven a reduction in out-of-school suspensions, and that the policy's effect on the charter probability of suspension was slightly smaller (6.3 percent) when measured relative to its signing. Consistent with Figure 1,this difference shows that the effect of Chapter 222 grew over time as schools adopted the policy. Finally, note that Table 1 omits the estimated effect of Chapter 222 on English test scores because—as indicated by Appendix Figure A5—the charter attendance effect on English scores were on a steady and positive before the introduction of Chapter 222, such

[5]Appendix Figure A5 plots the equivalent estimates for English, showing that the charter attendance effect on test scores followed a steady positive trend prior to the introduction of Chapter 222. Due to pre-trends, the Difference-in-Differences estimates for the effect of Chapter 222 on English test scores are be reliable. Therefore, I focus the analysis of Chapter 222's effect on n math test scores and suspensions outcomes only, but report point estimates for all year-by-year outcomes in Appendix Table A7, along with information on first stages and sample sizes.

[6]Appendix Tables A11 and A12 show pooled and year-by-year covariate balance regression results, respectively, documenting that charter lottery offers were as good as randomly assigned. Appendix Table A13 shows no differential attrition by charter lottery offer status in the sample.

[7]As shown in Appendix Table A5, these large average gains during middle school reflect the fact charter attendance effects grow with years of charter attendance, with the effects being smallest—though already substantial—in the first year after the lottery (0.400 standard deviations for math, 0.227 standard deviations for English), and largest in the fourth year after the lottery (0.814 math, 0.716 English).

that the difference-in-differences estimate for English is not reliable.[8]

4 Mechanisms

Table 1 suggests that Boston charter schools' suspensions practices are orthogonal to their ability to deliver large test score gains, since they delivered these gains even after reducing suspensions. In this section, I investigate two mechanisms through which Chapter 222 could have reduced charter suspensions without reducing the charter attendance effect on test scores.

The first possibility is that the reduction in charter suspensions induced by Chapter 222 benefited students who would have otherwise been suspended (for example, by keeping them in class), and harmed non-suspended students (for example, by not removing distracting behavior from classrooms), such that on average the charter attendance effect on test scores remained constant. The second possibility is that suspensions had no effect on learning, meaning that they were both inconsequential to suspended students' test scores and unnecessary for non-suspended students' learning gains. While the first mechanism requires that the effect of a charter suspension on test scores be negative, the second requires that it be zero.

Differentiating between these two mechanisms therefore requires identifying the causal effect of suspensions on charter students. The key identification challenge in this analysis concerns student selection into suspension. In particular, unlike charter attendance, which is as good as randomly assigned via lottery, suspensions are not randomly assigned, and the behaviors that lead to suspensions are often a consequence of complex unobserved factors that also negatively affect students' learning, such as problems at home (Steinberg and Lacoe, 2017). [9] I next describe how I address this identification challenge by combining

[8]While the total number of suspended students is too small to allow for a breakdown of Chapter 222's impact by suspension offense type, comparing the rates of suspension in charter relative to TPS before and after Chapter 222 suggests that Chapter 222's primary incidence was on its intended offense type: non-drug, non-violent, and non-criminal offenses. Out-of-school suspensions for this offense type declined both in Boston TPS schools and in charters, though with a more pronounced decline for charters (5 percentage points) than for TPS (1 percentage point). The other offense types along with their average percent incidence are: criminal offenses, violent offenses, and bullying, harassment or property offenses. Less than one percent of students in either charters or TPS are suspended out-of-school under these offense types.

[9]As shown in Appendix Table A8, this pattern holds true among charter applicants. Suspended students who attend Boston TPS score 0.163 and 0.150 standard deviations lower in math and English, respectively, than their non-suspended peers. In charter schools, the suspended vs. non-suspended test score gap is 0.110 standard deviations for math and 0.096 for English. The key identification question is whether these gaps are in fact caused by suspensions, or by omitted factors outside of the school's control.

lottery offers with students' pre-lottery suspension records.

4.1 The effect of suspension on charter students' test scores

I start my investigation of the causal effect of suspension on charter students' test scores by pooling outcomes of applicants from all grades into a simple additive effects regression. In this regression, the effect of attending a charter and the effect of being suspended have a linear and additive effect on test scores:

$$Y_{ik} = \quad \phi + \chi_1 D_{ik} + \chi_2 S_{ik} + \omega' X_i + \varepsilon_{ik}, \tag{4}$$

where D_{ik} is a dummy indicating whether applicant i attended any charter, S_{ik} is a dummy for whether applicant i was suspended (out-of or in-school) in the kth year after the lottery, and X_i is defined as in Equation 1.

To overcome the challenge that suspensions are not randomly assigned, I leverage cross-charter variation in suspension rates to separately identify the effect of suspension from the effect of any-charter attendance on test scores. The key idea is that if a student wins a charter lottery to a school with a stricter disciplinary environment, then she is more likely to be suspended in that charter, but continues to experience the same charter attendance effect as students who attended other charters. Note that this approach implicitly assumes that — suspension decisions aside— the effect of charter attendance on test scores is homogenous across charters.

To implement this strategy, I instrument both suspensions and charter attendance using each applicant's vector of charter lottery offers (rather than with a single dummy indicating an offer from any charters). Since charters' disciplinary codes and educational philosophies are determined at the charter network level, I instrument D_{ik} and S_{ik} with a full set of charter school network dummies, each indicating whether the applicant received an offer from one of the charter schools in the respective network.[10]

Columns (1) and (2) of Table 2 present estimates of χ_1 and χ_2 from Equation 4. Column (1) shows that being suspended does not significantly impact a student's test scores, while Column (2) shows that attending a charter increases math test scores by 0.314 standard deviations (standard error 0.047), and English test scores by 0.107 (0.047) standard deviations. These findings suggest that being suspended does not meaningfully affect a student's test

[10] A total of nine networks span the fifteen charter schools for which lottery records are available, listed in Appendix Table A14. Some networks have only one charter school operating in Boston.

scores, but attending a charter does.

4.2 Robustness

A potential concern with the estimates in Columns (1) and (2) Table 2 is that the individual charter lottery offers may not provide enough variation in suspension treatment assignment, as suggested by the 9.674 first-stage F-statistic on the suspension treatment, slightly below the rule-of-thumb F-statistic of 10 for rejection of weak identification. To overcome a weak first stage, I create additional instruments for D_{ik} and S_{ik} using applicants' pre-lottery suspensions data. Specifically, I first estimate suspension propensity scores σ_i for each applicant i, and create new instruments for D_{ik} and S_{ik} by interacting the individual network lottery offers with σ_i.[11] Each applicant's σ_i is also included as a control in Equation 4 and its corresponding first stage equations,[12] such that χ_1 and χ_2 are estimated among students of similar suspension propensity.

I present results for this alternative exercise in Appendix Table 2. Columns (1)–(3) present estimates of χ_1 and χ_2 where interactions of the charter dummies with σ_i are added to the set of instruments to Equation 4. Since suspensions are rare in the pre-lottery years, most students have a very low σ_i. As a result, the average realized probability of suspension post-lottery is higher than the predicted suspension probability at low σ_i values. To address the concern that rare suspensions in the pre-lottery data might underestimate the lower values of σ_i, in Columns (4)–(6) I dichotomize the σ_i distribution with a dummy variable indicating high suspension propensity, and use that dummy variable instead of σ_i itself to generate the interactions with charter network offer dummies. The high suspension propensity dummy includes all applicants with σ_i greater than 0.2 on a scale of 0 to 1, which is the cutoff above which the post-lottery suspension probability averages fifty percent.[13]

The results in Columns (3) and (4) of Table 2 are qualitatively similar to those in Columns

[11] Setren (2017) uses a similar methodology to analyze the impact of special education re-classification in charters. I estimate the student-specific suspension propensity in two steps. First, I use the sample of non-charter applicant 5th–8th graders attending TPS schools to estimate a logistic regression of an out-of-school suspension dummy on various student demographics and disciplinary records. I then predict σ_i for each charter applicant using the applicant's demographics and baseline grade disciplinary data and the logit coefficients of each predictor from the regression estimated in the TPS sample. Finally, I include the student-specific propensity score as a control in the second stage and first stage regressions.

[12] I present the distribution of σ_i among charter applicants in Appendix Figure A6. Panel A plots this distribution separately by suspension treatment post-lottery, showing that σ_i is predictive of suspensions, while Panel B plots it separately for lottery winners and losers, showing that σ_i is balanced across charter lottery offer status.

[13] Since suspensions are rare in the pre-lottery data, $\sigma_i = 0.20$ is the 95th percentile of the σ_i distribution.

(1) and (2), but the first stage in Column (3) suggests a much stronger identification of suspension effects (with an F-statistic of 15) and near zero suspension effect point estimates of -0.047 (0.149) for math and 0.065 (0.145) for English. Finally, I now augment Equation 4 with the interaction term $D_{ik} \times S_{ik}$ in order to test whether suspensions have different effects on test scores if they happen at charters versus at TPS:

$$Y_{ik} = \quad \phi + \chi_1 D_{ik} + \chi_2 S_{ik} + \chi_3 \left(D_{ik} \times S_{ik}\right) + \omega' X_i + \varepsilon_{ik}. \tag{5}$$

With Equation 5 I can separately estimate the effect of suspension on charter students $(\chi_1 + \chi_3)$ and the effect of charter attendance on suspended students $(\chi_2 + \chi_3)$. Table 3 presents estimates of χ_1, χ_2, and χ_3 using the dummied instruments as in Columns (3)–(4) of Appendix Table 2.[14] While the first-stage F-statistics suggest weak identification of coefficients χ_2 and χ_3, I find some evidence that the net-positive effect of charter suspensions shown in Column (6) of Appendix Table 3 is driven by a large positive effect of charter attendance on suspended students (Column 5) and a zero causal effect of suspensions on charter students' test scores (Column 4). Finally, consistent with the hypothesis that the causal effect of suspensions depends on which school the suspended student attends, Column (1) suggests that the causal effect of being suspended at a TPS school on test scores is large and negative.

Taking all robustness exercises into account, the results from Table 2 stand: suspensions have no meaningful or statistically significant effect on charter students' math test scores.

5 Conclusion

Previous charter lottery studies have documented large test score gains from charter attendance, with the largest gains driven by charters who suspend more students Angrist et al. (2013); Chabrier et al. (2016). This paper leverages Boston charter middle school lotteries and Chapter 222, a Massachusetts policy aimed at reducing charter suspensions, to understand whether suspensions are a key component of charters' ability to deliver large average learning gains.

I find that Chapter 222 successfully reduced the charter attendance effect on suspensions by 10 percentage points three years after the policy took effect, nearly closing the charter to Traditional Public Schools suspensions gap. However, I find no evidence that charters'

[14] Appendix Table A10 presents robustness checks using the same instruments as in Columns (1)–(2) of Table 2.

reduction in suspensions reduced the charter attendance effect on student test scores. I then investigate the mechanisms behind Chapter 222's effects, finding suggestive evidence that the causal effect of suspensions on Boston charter students' test scores is zero, whereas the causal effect of charter attendance on suspended students' test scores is large and positive. Overall, these findings indicate that non-suspended students' large test score gains at charter schools are *not* obtained at the expense of suspended students' learning.

References

Abdulkadiroğlu, A., J. D. Angrist, S. M. Dynarski, T. J. Kane, and P. A. Pathak (2011): "Accountability and flexibility in public schools: Evidence from Boston's charters and pilots," *The Quarterly Journal of Economics*, 126, 699–748.

Abdulkadiroğlu, A., J. D. Angrist, P. D. Hull, and P. A. Pathak (2016): "Charters without lotteries: Testing takeovers in New Orleans and Boston," *The American Economic Review*, 106, 1878–1920.

Abdulkadiroğlu, A., J. D. Angrist, Y. Narita, and P. A. Pathak (2017): "Research design meets market design: Using centralized assignment for impact evaluation," *Econometrica*, 85, 1373–1432.

Angrist, J. D., S. R. Cohodes, S. M. Dynarski, P. A. Pathak, and C. R. Walters (2016): "Stand and deliver: Effects of Boston?s charter high schools on college preparation, entry, and choice," *Journal of Labor Economics*, 34, 275–318.

Angrist, J. D., S. M. Dynarski, T. J. Kane, P. A. Pathak, and C. R. Walters (2012): "Who benefits from KIPP?" *Journal of policy Analysis and Management*, 31, 837–860.

Angrist, J. D., S. M. Dynarski, T. J. Kane, P. A. Pathak, C. R. Walters, et al. (2010): "Inputs and impacts in charter schools: KIPP Lynn," *American Economic Review*, 100, 239–243.

Angrist, J. D., P. A. Pathak, and C. R. Walters (2013): "Explaining Charter School Effectiveness," *American Economic Journal: Applied Economics*, 5, 1–27.

Chabrier, J., S. Cohodes, and P. Oreopoulos (2016): "What can we learn from charter school lotteries?" *Journal of Economic Perspectives*, 30, 57–84.

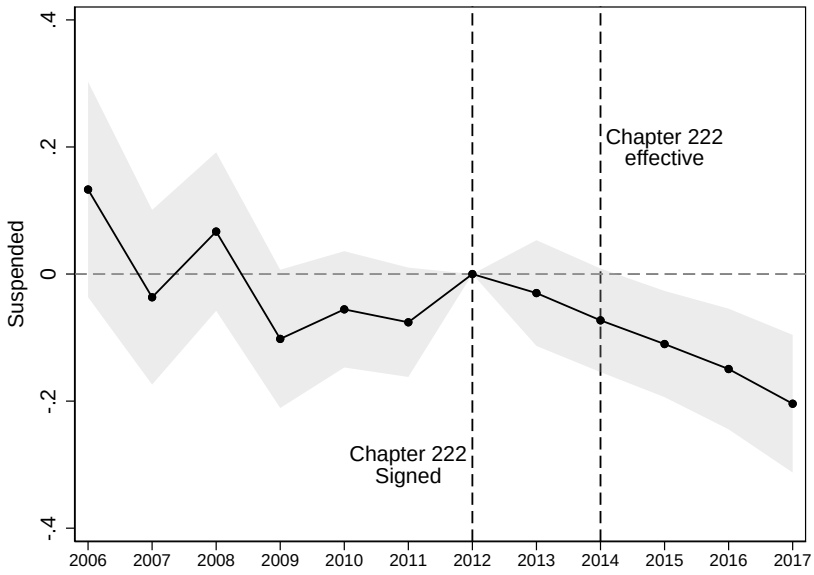
Dobbie, W. and R. G. Fryer Jr (2011): "Are high-quality schools enough to increase achievement among the poor? Evidence from the Harlem Children's Zone," *American Economic Journal: Applied Economics*, 3, 158–87.

Hoxby, C. M. and S. Murarka (2009): "Charter schools in New York City: Who enrolls and how they affect their students' achievement," Tech. rep., National Bureau of Economic Research.

Setren, E. (2017): "The Impact of Specialized Services vs. High Quality General Education on Special Education and English Language Learner Students," .

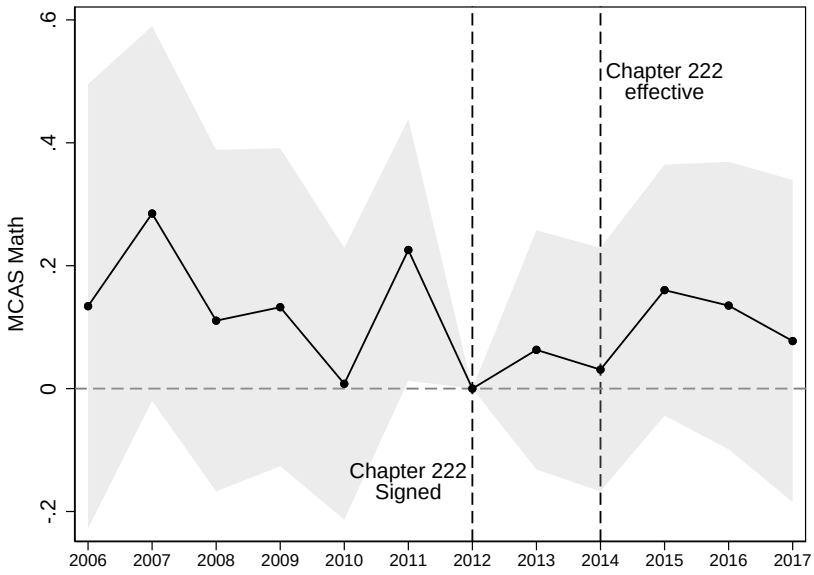
Steinberg, M. P. and J. Lacoe (2017): "What do we know about school discipline reform? Assessing the alternatives to suspensions and expulsions," *Education Next*, 17, 44–53.

Figure 1: 2SLS DD effects on probability of suspension



Note: This figure shows 2SLS difference-in-differences estimates of charter attendance on whether the student was ever suspended in the school year, relative to base year 2012. The treatment is a charter attendance dummy, interacted with year dummies, and omitting school year 2012. The instrument is a charter offer dummy interacted with the same dummies. The specification controls for applicant risk sets, grade and outcome year dummies, and baseline covariates.

Figure 2: 2SLS DD effects on MCAS math test scores



Note: See notes to Figure 2. See the Data Appendix for details on Massachusetts' MCAS tests.

Table 1: Charter/Suspension reform interactions

	Chapter 222 signed		Chapter 222 effective	
	Ever attended charter	Ever attended charter x (After 2012)	Ever attended charter	Ever attended charter x (After 2014)
	(1)	(2)	(3)	(4)
Panel A: Second stage				
Disciplinary outcomes				
Suspended	0.216***	-0.063**	0.205***	-0.095***
	(0.017)	(0.026)	(0.014)	(0.027)
Out-of-school	0.185***	-0.046*	0.182***	-0.093***
	(0.017)	(0.025)	(0.014)	(0.026)
In-school	0.072***	-0.016	0.065***	-0.007
	(0.009)	(0.014)	(0.008)	(0.015)
MCAS test scores				
MCAS Math	0.613***	-0.013	0.591***	0.072
	(0.041)	(0.062)	(0.034)	(0.065)
Panel B: First stage				
Charter offer	0.443***	-0.006***	0.399***	-0.014***
	(0.009)	(0.002)	(0.007)	(0.002)
Charter offer x (After 222)	-0.120***	0.330***	-0.091***	0.342***
	(0.012)	(0.008)	(0.012)	(0.010)
F-statistic	3,511	3,623	3,511	3,623
N	29,096		29,096	

Notes: This table displays coefficients from Two Stage Least Squares Difference-in-Differences (2SLS DD) regressions whose goal is to quantify the effect of Chapter 222 on charter attendance 2SLS treatment effects. The 2SLS DD procedure is implemented as a two-endogenous variable, two-instrument 2SLS regression, where the treatment variables are a dummy for whether the charter applicant ever attends charter, and an interaction of this dummy with whether the outcome variable year is in or after Chapter 222. Columns (1)-(2) displays results using Chapter 222's signature year (2012) to construct the interaction dummy, whereas Columns (3)-(4) use Chapter 222's effective year (2014). All regressions control for fully-saturated charter application risk sets and baseline grade covariates. Since charter applicants enter the sample in different years and at different grades, all regressions include outcome year, grade, and years-since-charter-lottery fixed effects. Sanderson-Windmeijer (2015) F-stats based on Angrist and Pischke (2009) are displayed for each the excluded instruments in the first stage regression. Test scores are standardized by grade and year to have mean zero and unit standard deviation at the state level. Robust standard errors are displayed in parentheses. *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Table 2: Charter and suspension effects identified using cross-school lottery variation and predicted suspensions

			Interacting individual charter offer instruments with			
			Applicant suspension propensity score		Dummy for suspension propensity score > 0.20	
	Suspended	Attended charter	Suspended	Attended charter	Suspended	Attended charter
	(1)	(2)	(3)	(4)	(5)	(6)
MCAS test scores						
Math	0.076	0.314***	-0.120	0.355***	-0.047	0.336***
	(0.177)	(0.047)	(0.140)	(0.042)	(0.149)	(0.043)
English	0.230	0.107**	0.061	0.151***	0.065	0.139***
	(0.174)	(0.047)	(0.144)	(0.044)	(0.145)	(0.044)
			First stage F-statistics			
F-statistic	9.674	25.613	6.264	20.885	15.065	22.404
Degrees of freedom	9	9	19	19	19	19
N			8,149		8,149	

Notes: The instruments in Columns (1)-(2) are the charter network-specific offer dummies, while those in Columns (3)-(4) are the charter network-specific offer dummies plus interactions with an applicant-specific suspension propensity score. The instruments in Columns (4)-(6) are the charter network-specific offer dummies plus interactions with a dummy indicating whether the applicant-specific suspension propensity is above 0.20 (in a scale of 0 to 1). See Table A9 and Figure A6 for details on the estimation and distribution of the student suspension propensity. All regressions control for applicant risk set dummies and all baseline grade covariates listed in Table A9. Sanderson-Windmeijer (2015) F-stats based on Angrist and Pischke (2009) are displayed for each the excluded instruments in the first stage regression. Test scores are standardized by grade and year to have mean zero and unit standard deviation at the state level. Robust standard errors are displayed in parentheses. *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Table 3: Suspension effect in charters vs. Charter effect on suspended

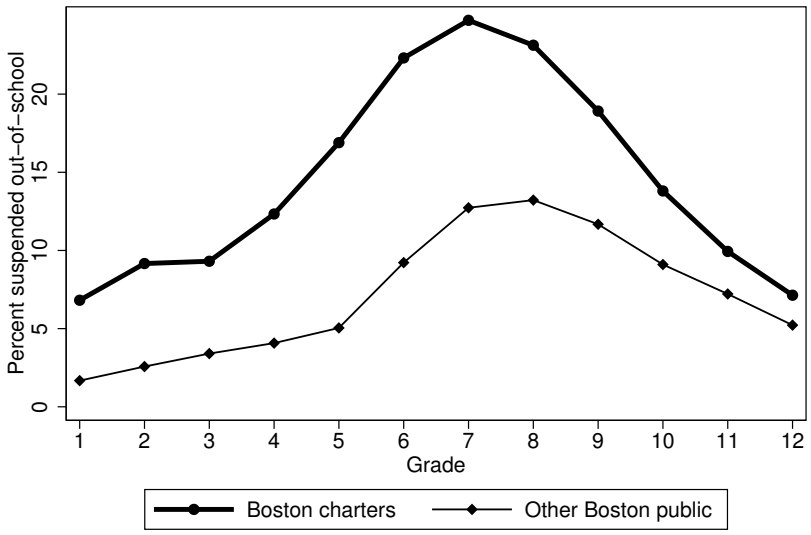
	Instruments: individual charter offers plus interactions with applicant suspension propensity score > 0.20				
	Treatments				
	Suspended	Attended charter	Attended charter x Suspended	Suspension effect in charters (1)+(3)	Charter effect on suspended (2)+(3)
	(1)	(2)	(3)	(4)	(5)
MCAS test scores					
Math	-0.629*	0.291***	0.589*	-0.040	0.879***
	(0.375)	(0.050)	(0.332)	(0.154)	(0.312)
English	-0.861**	0.057	0.967***	0.105	1.024***
	(0.345)	(0.052)	(0.318)	(0.152)	(0.295)
First stage F-statistics					
F-statistic	3.881	28.263	4.846	--	--
Degrees of freedom	18	18	18	--	--
N	8,149				

Notes: This table displays coefficients from an over-identified three-endogenous variables 2SLS regressions of charter attendance and suspensions on applicant test scores. The treatment variables are a dummy for whether the charter applicant ever attends charter, a dummy for whether the applicant is ever suspended, and the interaction of these two dummies. The instruments are charter network-specific offer dummies, and interactions of these dummies with a dummy indicating whether the applicant-specific suspension propensity score is above 0.20 (in a scale of 0 to 1). See Table A9 and Figure A6 for details on the estimation of the student suspension propensity. Columns (1)-(3) display the second state coefficients of the effect of each treatment on the respective outcome variables, while Columns (4)-(5) report net average treatment effects of interest, obtained as linear combinations of the estimated coefficients. All regressions control for applicant risk set dummies and baseline grade covariates. Sanderson-Windmeijer (2015) F-stats based on Angrist and Pischke (2009) are displayed for each the excluded instruments in the first stage regression. Test scores are standardized by grade and year to have mean zero and unit standard deviation at the state level. Robust standard errors are displayed in parentheses. *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Online Appendix

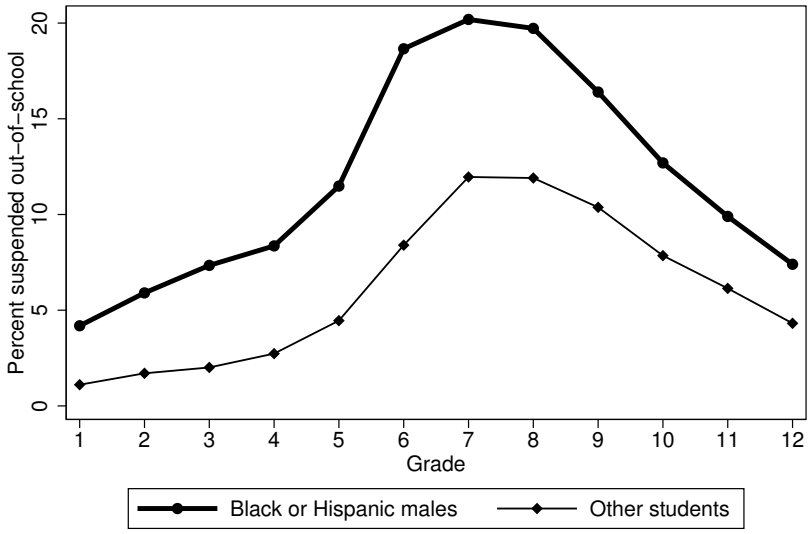
Figure A3: Percent of Boston students suspended out-of-school

Panel A: Charters vs. Other



Note: Sample includes all 149,802 students ever attending a Boston general education public school between 2006 and 2017. Excludes exam schools.

Panel B: Black or Hispanic vs. Other

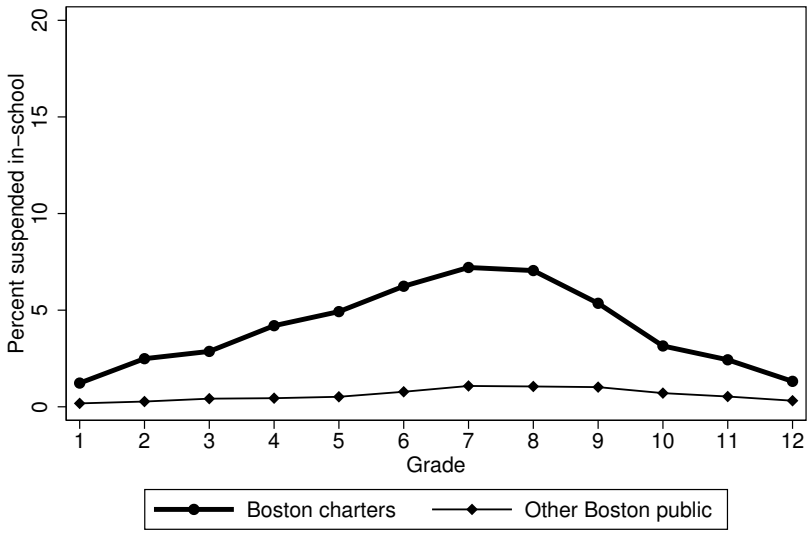


Note: Sample includes all 149,802 students ever attending a Boston general education public school between 2006 and 2017. Excludes exam schols.

Note: The sample includes all students ever attending a Boston general education public school between 2006 and 2017, excluding exam schools.

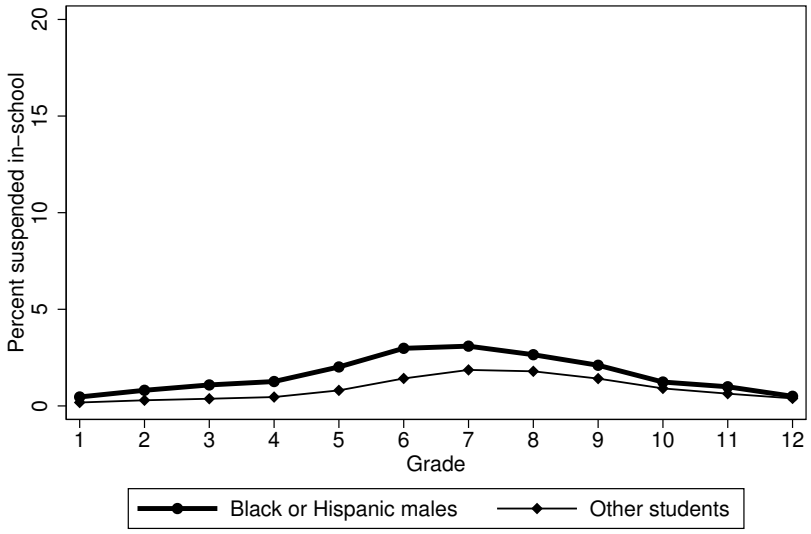
Figure A4: Percent of Boston students suspended in-school

Panel A: Charters vs. Other



Note: Sample includes all 149,802 students ever attending a Boston general education public school between 2006 and 2017. Excludes exam schools.

Panel B: Black or Hispanic vs. Other



Note: Sample includes all 149,802 students ever attending a Boston general education public school between 2006 and 2017. Excludes exam schols.

Note: The sample includes all students ever attending a Boston general education public school between 2006 and 2017, excluding exam schools.

Table A4: IV Charter attendance effect on first year post-lottery outcomes

	Lottery losers mean (1)	Charter attendance effect (2SLS) (2)
First stage		
Instrument: any charter offer		0.501***
		(0.012)
F-statistic		1,821
Discipline outcomes		
Suspended out-of-school	0.091	0.173***
		(0.018)
Suspended in-school	0.017	0.081***
		(0.010)
Days suspended out-of-school	0.296	0.824***
		(0.113)
Days suspended in-school	0.037	0.179***
		(0.030)
Expelled	0.001	0.002
		(0.002)
MCAS test scores		
Math	-0.365	0.400***
		(0.033)
English	-0.457	0.227***
		(0.034)
N	4,054	8,206

Notes: This table displays 2SLS estimates of charter attendance for Boston charter middle school applicants. The first stage estimate is the regression coefficient of the any-charter attendance dummy on an any- charter lottery offer dummy, controlling for fully-saturated charter application risk sets, and a set of baseline covariate controls. Test scores are standardized by grade and year to have mean zero and unit standard deviation at the state level. Robust standard errors are displayed in parentheses. *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Table A5: Charter attendance effect for post-lottery outcomes

		Charter attendance effect			
	Lottery losers mean	1 year after lottery	2 years after lottery	3 years after lottery	4 years after lottery
	(1)	(2)	(3)	(4)	(5)
First stage					
Instrument: any charter offer		0.501***	0.353***	0.293***	0.223***
		(0.012)	(0.013)	(0.013)	(0.014)
F-statistic		1,821	762	494	246
Discipline outcomes					
Suspended out-of-school	0.091	0.173***	0.153***	0.187***	0.050
		(0.018)	(0.027)	(0.034)	(0.043)
Days suspended out-of-school	0.296	0.824***	0.957***	0.863***	-0.633
		(0.113)	(0.204)	(0.240)	(0.715)
Suspended in-school	0.017	0.081***	0.053***	0.068***	0.034
		(0.010)	(0.016)	(0.019)	(0.023)
Days suspended in-school	0.037	0.179***	0.240***	0.198***	0.068
		(0.030)	(0.060)	(0.054)	(0.061)
Expelled	0.001	0.002	-0.002	0.000	0.000
		(0.002)	(0.003)	(0.003)	(0.006)
MCAS test scores					
Math	-0.365	0.400***	0.754***	0.761***	0.814***
		(0.033)	(0.049)	(0.061)	(0.100)
English	-0.457	0.227***	0.486***	0.511***	0.716***
		(0.034)	(0.050)	(0.063)	(0.104)
N	4,054	8,206	7,886	7,548	3,657

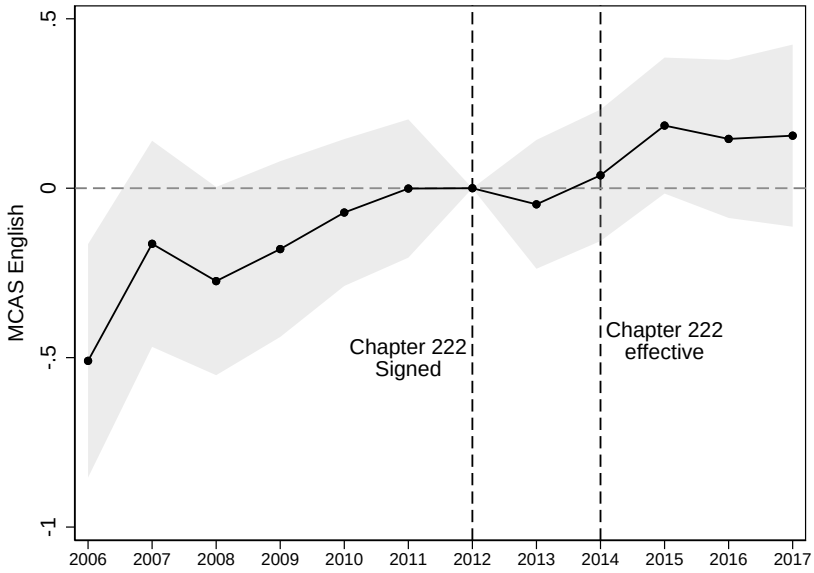
Notes: This table displays 2SLS estimates of charter attendance for Boston charter middle school applicants, separately estimated for each year since the charter lottery application. The first stage estimate is the regression coefficient of the any-charter attendance dummy on an any- charter lottery offer dummy, controlling for fully-saturated charter application risk sets, and a set of baseline covariate controls. Test scores are standardized by grade and year to have mean zero and unit standard deviation at the state level. Robust standard errors are displayed in parentheses. *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Table A6: Heterogeneity in charter attendance effect on first year post-lottery outcomes by baseline suspension

	Attended any charter	Attended any charter x Suspended at baseline
	(1)	(2)
Test scores		
MCAS Math	0.402***	-0.038
	(0.033)	(0.107)
MCAS English	0.226***	0.032
	(0.034)	(0.122)
First stage		
Excluded instruments		
Any charter offer	0.505***	-0.003*
	(0.012)	(0.001)
Any charter offer x	0.044	0.598***
Suspended at baseline	(0.045)	(0.043)
F-statistic	1,792	626
p-value	0.000	0.000
Degrees of freedom		
df1	1	1
df2	7,779	7,779
N	8,206	

Notes: This table displays 2SLS estimates of heterogeneity in charter attendance effects in the first year after lottery by applicant's baseline grade suspension status. All regressions control for fully-saturated charter application risk sets and non- disciplinary baseline covariate controls. Test scores are standardized by grade and year to have mean zero and unit standard deviation at the state level. Robust standard errors are displayed in parentheses. *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

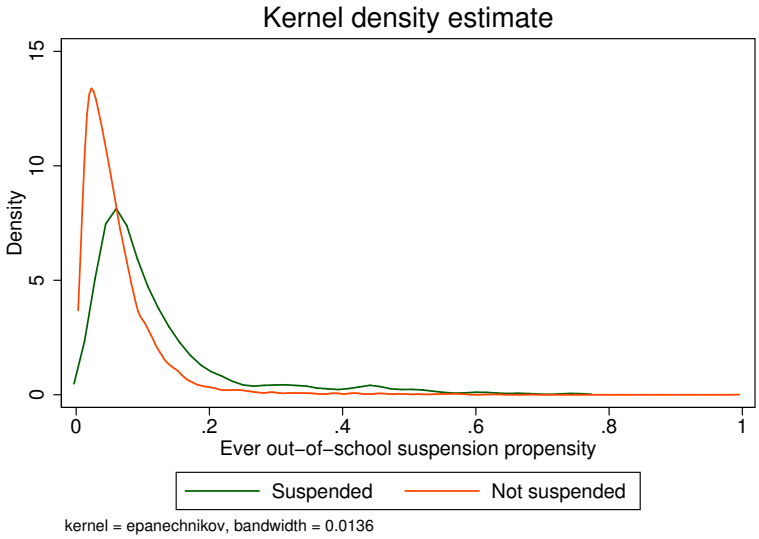
Figure A5: Year-by-year charter attendance 2SLS treatment effects on test scores



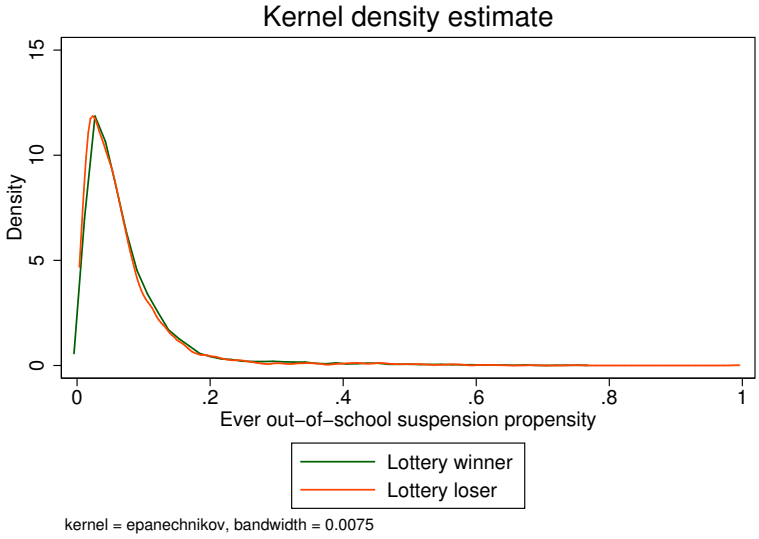
Note: See notes to Figure 1.

Figure A6: Applicant suspension propensities

Panel A: Suspended vs. Not suspended



Panel B: Charter lottery winner vs. loser



Note: This figure displays the distribution of student suspension propensity scores by suspension status and charter offer status. In Figure A4a, suspended students are those ever suspended, whether in-school or out-of-school, in the first year after the charter lottery. Applicant suspension propensities are estimated in two steps. First, a logit regression of a dummy for whether a student is ever suspended in academic year y is regressed on a rich set of predictors measured as of year y-1 on the sample of Boston students in grades 3-8 who never apply to charter schools. Table A5 displays the list of predictors along with their odds ratio coefficients. Second, the covariance structure estimated in this first step is used to predict suspension propensities in the sample of charter applicants, using each applicant's baseline grade measures as predictors for the applicant-specific suspension propensity.

Table A7: Year-by-year 2SLS estimates of charter attendance treatment effects

	Treatment: ever attended any charter											
	Before Chapter 222 Signing							After Chapter 222 Signing				
										Chapter 222 effective		
Post-lottery calendar year:	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Discipline outcomes												
Suspended	0.344***	0.181***	0.308***	0.127**	0.207***	0.176***	0.243***	0.172***	0.158***	0.101**	0.160**	0.064
	(0.080)	(0.067)	(0.061)	(0.052)	(0.041)	(0.037)	(0.034)	(0.038)	(0.040)	(0.045)	(0.064)	(0.073)
Suspended out-of-school	0.314***	0.170**	0.244***	0.050	0.167***	0.142***	0.235***	0.171***	0.167***	0.092**	0.096	0.065
	(0.077)	(0.066)	(0.057)	(0.049)	(0.039)	(0.035)	(0.033)	(0.037)	(0.039)	(0.043)	(0.061)	(0.070)
N	549	875	1,205	1,545	1,768	1,916	3,145	3,979	4,921	4,738	3,576	2,147
Test score outcomes												
MCAS Math	0.705***	0.851***	0.607***	0.571***	0.529***	0.774***	0.566***	0.691***	0.676***	0.566***	0.665***	0.588***
	(0.157)	(0.146)	(0.136)	(0.120)	(0.095)	(0.092)	(0.083)	(0.089)	(0.099)	(0.111)	(0.159)	(0.176)
MCAS English	0.038	0.393***	0.157	0.129	0.387***	0.429***	0.450***	0.421***	0.525***	0.593***	0.430***	0.621***
	(0.149)	(0.140)	(0.136)	(0.124)	(0.093)	(0.087)	(0.081)	(0.087)	(0.099)	(0.110)	(0.159)	(0.189)
N	545	855	1,170	1,460	1,714	1,863	3,055	3,866	4,736	4,400	3,374	2,058
First stage												
Instrument: any lottery offer	0.475***	0.455***	0.426***	0.437***	0.495***	0.488***	0.419***	0.369***	0.320***	0.307***	0.257***	0.289***
	(0.039)	(0.034)	(0.030)	(0.028)	(0.024)	(0.023)	(0.018)	(0.017)	(0.017)	(0.019)	(0.024)	(0.031)
F-statistic	155	183	205	259	429	454	540	453	342	241	120	92
N	545	855	1,170	1,460	1,714	1,863	3,055	3,866	4,736	4,400	3,374	2,058

Notes: This table displays 2SLS regression coefficients displayed in Figures 2 and 3, which are estimated from year-by-year regressions of the outcomes listed on the left on an ever-attended-charter dummy. The instrument in each regression is an any-charter lottery offer. All regressions control for fully-saturated charter application risk sets and baseline grade covariates. Since charter applicants enter the sample in different years and at different grades, all regressions include outcome year, grade, and years-since-charter-lottery fixed effects. Test scores are standardized by grade and year to have mean zero and unit standard deviation at the state level. Robust standard errors are displayed in parentheses. *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Table A8: OLS effect of suspensions on charter applicant test scores

	Suspension OLS in BPS		Suspension OLS in Charter	
	Not suspended mean	Suspended	Not suspended mean	Suspended
	(1)	(2)	(3)	(4)
MCAS Test Scores				
Math	-0.342	-0.163***	0.040	-0.110***
		(0.035)		(0.026)
English	-0.419	-0.150***	-0.173	-0.096***
		(0.037)		(0.027)
N		4,582		3,619

Notes: This table reports OLS estimates of the effect of being suspended on a student's test score outcomes conditional on the school type that the student attends (Charter or Boston Public Schools). Regressions control for the student's propensity to be suspended and for all baseline covariates listed in Appendix Table A9. See Table A9 and Figure A6 for details on the estimation of the student suspension propensity. The sample is applicants to charter schools offering seats for entry grades 5 or 6 in academic years 2004-2005 through 2014-2015. Test scores are standardized by grade and year to have mean zero and unit standard deviation at the state level. Robust standard errors are displayed in parentheses. *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Table A9: Predictors used in estimating charter applicant suspension propensity scores

Variables commonly used as covariate controls in charter lottery studies	Odds ratio coefficient (1)	*Additional predictors from disciplinary and enrollment recods*	Odds ratio coefficient (2)
Demographics		*Baseline grade suspension*	
Female	0.571***	Ever suspended out-of-school	4.049***
	(0.012)		(0.138)
Asian	0.648***	Ever suspended	2.326***
	(0.045)	in-school	(0.265)
Black	2.036***	Days suspended	1.098***
	(0.082)	out-of-school	(0.011)
Hispanic	1.434***	Days suspended	0.978***
	(0.061)	in-school	(0.052)
Other non-white	1.858***		
	(0.089)		
Baseline grade measures		*Baseline grade enrollment*	
Free or reduced price lunch	1.477***	Days attended school	0.994***
	(0.046)		(0.000)
English Language Learner	0.769***	Transferred to another school	1.164***
	(0.020)		(0.083)
English MCAS	0.803***	Repeated baseline grade	1.066***
	(0.012)		(0.058)
Math MCAS	0.781***	Immigrant	0.652***
	(0.013)		(0.036)
Special education	1.059***	Age	1.124***
	(0.024)		(0.016)

Note: This table reports odds ratio coefficients from a school or out-of-school) status on the listed variables plus grade fixed effects. The logistic regression is estimated on a sample of Boston students who never applied to charter schools. The sample contains students in grades 3-8 between between years 2004 and 2017. Suspension propensity scores are then predicted for charter applicants using applicants' baseline grade measures as predictors.

Table A10: Suspension effect in charters vs. Charter effect on suspended: robustness to excluded instruments

	Instruments: individual charter offers					Instruments: individual charter offers plus interactions with dummy for baseline out-of-school suspension					Instruments: individual charter offers plus interactions with applicant suspension propensity score				
	Treatments					Treatments					Treatments				
	Suspended	Attended charter	Attended charter x Suspended	Suspension effect in charters (1)+(3)	Charter effect on suspended (2)+(3)	Suspended	Attended charter	Attended charter x Suspended	Suspension effect in charters (6)+(8)	Charter effect on suspended (7)+(8)	Suspended	Attended charter	Attended charter x Suspended	Suspension effect in charters (11)+(13)	Charter effect on suspended (12)+(13)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
MCAS test scores															
Math	-3.824***	0.073	3.813***	-0.011	3.886***	-1.336***	0.239***	1.318***	-0.018	1.557***	-1.096***	0.304***	0.876***	-0.219	1.181***
	(1.239)	(0.092)	(1.175)	(0.264)	(1.115)	(0.481)	(0.054)	(0.425)	(0.172)	(0.403)	(0.412)	(0.048)	(0.340)	(0.157)	(0.325)
English	-2.890***	-0.106	3.097***	0.207	2.991***	-1.506***	-0.001	1.678***	0.172	1.677***	-1.210***	0.078	1.160***	-0.050	1.238***
	(1.037)	(0.085)	(0.995)	(0.224)	(0.937)	(0.503)	(0.057)	(0.458)	(0.172)	(0.432)	(0.423)	(0.050)	(0.354)	(0.159)	(0.337)
First stage F-statistics															
F-statistic	2.497	7.846	2.754	--	--	2.548	17.929	3.076	--	--	2.867	32.648	3.557	--	--
Degrees of freedom	8	8	8	--	--	18	18	18	--	--	18	18	18	--	--
N			8,149					8,149					8,149		

Notes: This table displays robustness to the set of excluded instruments for the estimates in Table 3. Instruments in Columns (1)-(5) are individual charter offers only; whereas Columns (6)-(10) and (11)-(15) present estimates interacting individual charter offers with a dummy indicating if the applicant was suspended out-of-school in the baseline grade, or the applicant suspension propensity score, respectively.

Table A11: Covariate balance for charter middle school lottery applicants

		Any-charter lottery losers mean (1)	Offered any charter seat (2)
Discipline baseline			
Suspended out-of-school		0.037	0.001
			(0.005)
Suspended in-school		0.003	0.000
			(0.001)
Expelled		0.000	0.000
			(0.000)
	N	9,646	
Academic achievement baseline			
MCAS Math		-0.408	-0.019
			(0.024)
MCAS English		-0.482	-0.002
			(0.025)
	N	8,906	
Time-varying demographics			
Low income		0.744	0.000
			(0.011)
Special education		0.201	-0.014
			(0.010)
Limited English Proficient		0.257	-0.003
			(0.011)
	N	9,646	
Gender and race			
Female		0.488	0.000
			(0.013)
Race			
Black		0.437	-0.013
			(0.012)
Hispanic		0.248	0.020*
			(0.011)
White		0.170	-0.005
			(0.008)
Asian		0.033	0.003
			(0.005)
	N	9,646	
Balance joint F-statistic p-value			0.456

Notes: This table displays covariate balance on baseline characteristics of charter lottery winners and losers. Column (2) displays OLS regression coefficients from regressions of each baseline characteristic on an any-charter offer dummy. All regressions control for fully-saturated charter application risk sets. Means for losers of all charter lotteries are displayed in Column (1) for reference. The joint F-statistic corresponds to the t-statistic of the any-charter offer dummy coefficient from a stacked regression of all baseline characteristics on the any-charter offer dummy. Test scores are standardized by grade and year to have mean zero and unit standard deviation at the state level. Robust standard errors at the attended school level are displayed in parentheses. *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Table A12: Year-by-year charter attendance covariate balance

	Instrument: any lottery offer											
Post-lottery calendar year:	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Baseline discipline												
Suspended out-of-school	-0.007	-0.038**	-0.009	-0.009	0.017	0.007	0.001	-0.003	-0.001	0.002	-0.004	0.006
	(0.018)	(0.017)	(0.016)	(0.014)	(0.011)	(0.009)	(0.007)	(0.007)	(0.007)	(0.008)	(0.009)	(0.013)
Suspended in-school	--	--	--	--	0.002	0.001	-0.001	-0.003	-0.002	0.001	-0.001	0.004
	--	--	--	--	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.006)
Expelled	--	--	--	--	-0.003	-0.002	-0.001	--	--	--	--	--
	--	--	--	--	(0.002)	(0.002)	(0.001)	--	--	--	--	--
N	549	875	1,205	1,545	1,768	1,916	3,145	3,979	4,921	4,738	3,576	2,147
Baseline test scores												
MCAS Math	0.081	0.005	-0.038	-0.024	-0.052	-0.035	-0.024	0.010	0.049	0.058	0.062	0.068
	(0.135)	(0.095)	(0.069)	(0.059)	(0.052)	(0.049)	(0.038)	(0.035)	(0.034)	(0.037)	(0.044)	(0.060)
MCAS English	0.051	0.028	-0.018	-0.044	-0.008	0.006	0.012	0.025	0.040	0.066*	0.074	0.060
	(0.144)	(0.100)	(0.075)	(0.062)	(0.053)	(0.050)	(0.038)	(0.036)	(0.034)	(0.037)	(0.045)	(0.060)
Baseline demographics												
Low income	0.036	0.031	0.029	0.019	0.018	0.032	-0.033**	-0.030*	-0.015	-0.007	0.009	0.012
	(0.044)	(0.037)	(0.033)	(0.030)	(0.026)	(0.024)	(0.016)	(0.016)	(0.014)	(0.016)	(0.019)	(0.025)
Special education	0.002	-0.013	-0.009	0.006	-0.035	-0.026	-0.016	-0.006	-0.013	-0.044***	-0.025	-0.045*
	(0.038)	(0.032)	(0.028)	(0.026)	(0.022)	(0.020)	(0.016)	(0.015)	(0.014)	(0.016)	(0.019)	(0.026)
Limited English Proficient	-0.015	-0.028*	-0.005	0.020	0.009	0.036*	0.014	-0.003	-0.005	-0.031*	-0.012	-0.008
	(0.018)	(0.016)	(0.018)	(0.020)	(0.018)	(0.019)	(0.017)	(0.017)	(0.017)	(0.019)	(0.022)	(0.029)
N	549	875	1,205	1,545	1,768	1,916	3,145	3,979	4,921	4,738	3,576	2,147
Gender and race												
Female	0.083*	0.053	0.006	-0.014	-0.002	-0.002	0.001	0.010	0.002	0.014	0.025	0.025
	(0.047)	(0.040)	(0.035)	(0.032)	(0.029)	(0.027)	(0.021)	(0.019)	(0.019)	(0.020)	(0.025)	(0.033)
Race												
Black	-0.036	-0.017	-0.010	-0.029	0.012	-0.014	-0.020	-0.034*	-0.017	-0.013	-0.015	-0.027
	(0.045)	(0.038)	(0.033)	(0.030)	(0.027)	(0.025)	(0.019)	(0.018)	(0.017)	(0.019)	(0.022)	(0.028)
Hispanic	0.018	0.008	-0.008	0.013	0.013	0.030	0.035**	0.025	0.015	0.006	0.011	0.024
	(0.038)	(0.031)	(0.025)	(0.025)	(0.023)	(0.021)	(0.017)	(0.016)	(0.016)	(0.018)	(0.021)	(0.028)
White	0.023	0.015	0.022	0.033	-0.017	-0.023	-0.017	0.000	-0.006	0.009	0.007	-0.006
	(0.035)	(0.031)	(0.028)	(0.025)	(0.021)	(0.020)	(0.013)	(0.012)	(0.011)	(0.012)	(0.015)	(0.022)
Asian	-0.002	-0.005	0.001	0.004	-0.009	-0.002	0.008	0.005	0.008	0.002	0.007	0.005
	(0.016)	(0.012)	(0.009)	(0.009)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.008)	(0.009)	(0.011)
N	549	875	1,205	1,545	1,768	1,916	3,145	3,979	4,921	4,738	3,576	2,147
Joint F-statistic p-value	0.455	0.176	0.769	0.669	0.607	0.556	0.320	0.420	0.823	0.194	0.740	0.505

Notes: This table displays covariate balance on baseline characteristics of charter lottery winners and losers for each outcome year. Columns (1)-(12) display OLS regression coefficients from regressions of each baseline characteristic on an any-charter offer dummy. All regressions control for fully-saturated charter application risk sets, grade, and years-since-lottery fixed effects. The joint F-statistic corresponds to the t-statistic of the any-charter offer dummy coefficient from a stacked regression of all baseline characteristics on the any-charter offer dummy. Test scores are standardized by grade and year to have mean zero and unit standard deviation at the state level. Robust standard errors at the attended school level are displayed in parentheses. *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Table A13: Charter lottery winners vs. losers: covariate balance

	Enrolled in MA Public School		Has English MCAS		Has Math MCAS	
	Any-charter lottery losers mean	Offered any charter seat	Any-charter lottery losers mean	Offered any charter seat	Any-charter lottery losers mean	Offered any charter seat
	(1)	(2)	(3)	(4)	(5)	(6)
Outcome year 1	0.950	0.014	0.896	0.013	0.906	0.014
		(0.005)		(0.007)		(0.007)
N	9,646		9,646		9,646	
Outcome year 2	0.909	0.025	0.872	0.022	0.870	0.025
		(0.007)		(0.008)		(0.008)
N	9,646		9,646		9,646	
Outcome year 3	0.885	0.015	0.842	0.011	0.839	0.015
		(0.008)		(0.009)		(0.009)
N	9,646		9,646		9,646	
Outcome year 4	0.774	-0.003	0.732	0.005	0.732	0.009
		(0.009)		(0.013)		(0.013)
N	9,646		5,385		5,385	

Notes: This table displays differential attrition between charter lottery winners and losers. Columns (2), (4), and (6) display OLS regression coefficients from regressions of dummies indicating enrollment in a MA public school, availability of English MCAS test score, and availability of math MCAS test score, respectively, on an any-charter offer dummy. Since MCAS is not administered for grade 9, differential attrition estimates for MCAS test scores in outcome year 4 excludes 6th grade applicants, for which grade 9 is the expected grade in the 4th outcome year. All regressions control for fully-saturated charter application risk sets. Means of lottery losers' attrition indicators are displayed in Columns (1), (3), and (5) for reference.

Table A14: Charter middle school lotteries: analysis sample applicant counts

	Application year										
Application year:	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
School	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Academy of the Pacific Rim Charter		139	166	292	116	172	145	222	420	467	
Boston Collegiate Charter	155	201	197	210	233	282	264	559	552	625	406
Boston Preparatory Charter		145	206	242	177	192	182	206	209	236	118
Codman											92
Brooke Charter School Roslindale			66	85	79	93					
Brooke Charter School Mattapan								182	103	273	241
Brooke Charter School East Boston									118	217	185
Excel Academy Charter					52	130	118	129	271	318	
Excel Academy Charter School - Boston II									172	235	
KIPP Academy Boston Charter School									104	132	209
MATCH Charter School					295	262	219	490	350	459	238
Uncommon Schools - Roxbury Prep	111	131	132	132	141	151	104	537	451	338	337
Uncommon Schools - Grove Hall								429	451	338	337
Uncommon Schools - Dorchester Prep									451	338	337
UP Academy Charter School of Boston								551	209	173	152

Note: This table displays the number of charter applicants in the analysis sample, by school lottery entered and by application year. Applications are submitted in indicated year for entry into grades 5 or 6 in the Fall semester of the following academic year.

Data Appendix

To estimate causal effects of suspensions and disciplinary environments on learning, I linked charter lotteries data collected by researchers at MIT's School Effectiveness and Inequality Initiative (SEII) to three administrative datasets provided by the Massachusetts Department of Elementary and Secondary Education (MA DESE): SIMS, SSDR, and MCAS.

SIMS

The SIMS dataset includes demographic information and student-level enrollment records for all MA public schools between school years 2002-3 and 2016-17. SIMS is used to compute three charter enrollment status by grade (charter treatment) and demographic controls.[15] These variables are coded as follows:

- *Charter treatment.* School codes and/or names are used to identify charters. To determine enrollment, since students may switch schools or grades in the same school year, some discretion is needed to determine in which school the student is enrolled. Since charter attendance is a treatment of interest, if in a given school year a student is enrolled for even one day at a charter school, the student is considered enrolled in a charter for that school year. Otherwise, I determine the school in which a student is enrolled based on the maximum number of days attended. This widely adopted definition of treatment is considered conservative because it counts towards treatment lower-achieving students who might leave charters mid-year.[16]

- *Demographic controls.* Dummy variables are created to indicate various demographic characteristics. While SIMS includes a wide number of interesting characteristics to be explored in further work – such as immigrant status and home language – the share students belonging to several of these characteristics is very low. Thus, I focus on more commonly explored demographic characteristics in the literature, such as gender, race, special education status, English language learner status, and low income status. Importantly, the last three characteristics may change over time. As a result, I define

[15] SIMS also includes total suspensions out-of-school, in-school, and an expulsion flag for school years 2003-04 and 2011-12. However, since the SIMS data is aggregated at the enrollment record level, incident dates are not available with which to compute all suspensions occurring prior to test-taking. As a result, I use SDDR as the primary source for data on disciplinary actions. Aggregating all incidents from SSDR at the student level gives similar figures to those reported in SIMS.

[16] Other papers implementing the same or similar strategies include Setren (2017); Abdulkadiroğlu et al. (2017, 2016); Angrist et al. (2016, 2013).

time-varying demographic controls using baseline data (that is, data prior to charter attendance) when analyzing charter and suspensions treatment effects. [17]

SSDR

The SSDR dataset includes student-level information on any drug, violent or criminal-related offenses, as well as any non-drug, non-violent or non-criminal-related offenses committed by the student on school property between school years 2003-04 and 2016-17. Incident dates, offense types, and disciplinary measures taken are available. Students are identified by the same unique identifier as in SIMS. SSDR is used to compute out-of-school and in-school suspensions, and an expulsion dummy for each student. Ever suspended dummies and total days suspended are computed. Three points must be highlighted:

- *Consider all offense types.* While rich data on offense types are available, most incidents in MA public schools entail "non-drug, non-violent or non-criminal-related" offenses only. As a result, a more detailed look into suspension effects by offense type would be limited in power, and is thus deferred to future work.

- *Timing of suspensions.* When analyzing suspensions and expulsions as outcomes, I consider incidents throughout the academic year. However, when estimating the treatment effect of suspensions on test scores and grade progression outcomes, I limit the SSDR data to incidents occurring prior to April of each school year, when the MCAS math and English test season commences.

- *Missing data.* I assume that students not cited in any SSDR incidents were not suspended. For the purposes of estimating unbiased and consistent charter attendance effects on suspensions, and suspension treatment effects on outcomes, this assumption requires no differential SSDR reporting between charters and other MA public schools. If charters are on average better reporters, charter attendance effects on suspensions will be overestimated.

While a thorough investigation of schools' reporting habits is beyond the scope of this paper, it is unlikely that differential SSDR would drive the results in this paper. If anything, since

[17] Furthermore, it is important to note that a student's classification as special education status is a function of the school in which the student is enrolled, and could therefore change if the student enrolls at a charter. In fact, Setren (2017) finds large causal effects of charter enrollment on special education declassification, as charters move special education students into more inclusive classrooms. While the study of school discipline is particularly relevant for special education populations, assessing how declassification and suspensions interact in producing aggregate charter attendance effects is beyond the scope of this project.

charters are consistently under the criticism of over-suspending students, one might expect charters to under-report rather than over-report suspensions. Moreover, higher prevalence of reported suspensions among charters is consistent throughout many US public school districts with varying degrees of data quality collection and reporting standards.

MCAS

The MCAS dataset includes annual MCAS math and English test scores for MA public school students in grades 3 through 8, and 10. Since students may retake the test, I follow the literature in considering test results for the first attempt only.[18] I then standardize the test scores for each subject by grade and year to have mean zero and unit standard deviation at the state level.

Charter lotteries

I use Boston charter middle school lottery records collected by researchers at SEII for charter seats in school years 2004-05 through 2016-17. This sample includes 12 of 17 Boston charters offering middle school grades throughout the sample period.[19] Two points on sample selection are worth emphasizing:

- *Focus on Boston.* Focusing on Boston allows me to use of multiple charter lottery offers as instruments for suspensions and charter attendance in investigating the mechanisms behind Chapter 222's effect.

- *Focus on middle school.* I focus on lotteries for middle school entry (grades 5 and 6) for three reasons. First, as I show in Appendix Figures A3-A4, suspensions are primarily a middle school phenomenon in Massachusetts. Second, test scores are available for grades 3-8, allowing for analysis of estimation of test score treatment effects for 1 to 4 years following charter treatment, which is not possible for high school applicants. Finally, middle school applicants have 2-3 baseline grades with test score and discipline histories with which suspension propensities can be computed.

Lotteries take place in the Spring semester for entrance in the following Fall. Charters typically make initial offers and include several other students on a waitlist. When students

[18] In school years 2014 and 2015, Massachusetts experimented switching the standardized test to PARCC exams instead of MCAS. I use the MCAS-corresponding scores provided by MA DESE in the PARCC test score datasets for all PARCC scores.

[19] These figures exclude five charters that specialize in alternative and special education, for which there are no oversubscribed lotteries.

initially offered seats decline attendance, offers are made to waitlisted students. For the purposes of this paper, an applicant is considered a lottery winner if he or she receives either an initial or an off-waitlist offer. (Angrist et al., 2016) presents charter attendance effects on test scores for initially and waitlisted applicants separately.

Importantly, some lottery applicants may be guaranteed a seat at the charter if she/he either has a sibling in the school or fills any special school priorities. These applicants are excluded from analyses as they are not subject to randomization.

Linking datasets

Lottery records and administrative datasets contain identifiable information, such as names and dates of birth, and are thus stored in a restricted access facility at the National Bureau of Economic Research, in accordance with this project's Memorandum of Understanding with MA DESE. Once lottery records are matched to SIMS on names and date of birth, identifiable information are discarded from analyses files. Unique identifiers, available in all administrative datasets, are used to construct a panel dataset tracking applicants across time. This panel dataset includes demographic controls, baseline variables, treatment variables, and outcome variables.