BHI Policy Study

Why Massachusetts Should Double the Number of Charter Schools

Jonathan Haughton, PhD
Paul Bachman, MSIE
David G. Tuerck, PhD
Eli Cutler, BA

THE BEACON HILL INSTITUTE AT SUFFOLK UNIVERSITY
8 Ashburton Place
Boston, MA 02108
Tel: 617-573-8750, Fax: 617-994-4279
Email bhi@beaconhill.org, Web www.beaconhill.org
MARCH 2011

TABLE OF CONTENTS

TABLE OF TABLES

Executive Summary

Do charter schools outperform traditional schools? The question is hard to answer because a simple comparison of raw test scores for charter and traditional schools does not control for factors that distinguish charter school students from traditional school students, notably the fact that charter schools have a much higher representation of minority and low-income students.

As important as it is for Massachusetts and other states that have made a substantial commitment to charter school education, we do not know of any study that has attempted to answer this question directly, which is to say, by measuring the performance of all charter schools in Massachusetts against similar traditional schools. An earlier study, which was sponsored by the Boston Foundation, compared the performance, as measured by Massachusetts Comprehensive Assessment System (MCAS) scores, of students who attend charter schools only in Boston with students who do not.[1] One question asked by this study was whether students who are randomly selected for admission to charter schools perform better than students who apply but are not selected for admission. This goes to the policy question of whether it would improve student performance to provide additional seats in existing charter schools so that the students denied admission to those schools would have the opportunity to attend.

Here we ask a different question. Our comparison is between schools, not students. Specifically, we compare the performance of charter schools to similar traditional schools (matched according to their "propensity" to attract similar students) to determine whether, and for what grades and subjects, charter schools perform better than the traditional schools.

Both questions are important – whether to provide additional seats or to open new schools. Interestingly, it turns out that our results are broadly similar to those reached in the Boston Foundation study. That is, charter schools outperform traditional schools at the eighth and tenth grade levels, though not the fourth grade level.

[1] Abdulkadiroglu, et al., "Informing the Debate: Comparing Boston's charter, pilot and traditional schools," Boston: The Boston Foundation (January, 2009).

Specifically, when we match charter schools having high concentrations of minority and low-income students with traditional schools having similar concentrations of minority and low-income students, the charter schools outperform the traditional schools at the eighth and tenth grade levels.

Compared to similar traditional schools, the MCAS scores for charter schools show an improvement of

- 6.2% to 9.9% improvement over the scores for traditional schools in eighth grade English;
- 9.8% to 23.0% improvement over the scores for traditional schools in eighth grade mathematics;
- 25.7% improvement over the scores for traditional schools in tenth grade mathematics; and
- 8.2% to 13.7% improvement over the scores for traditional schools in tenth grade English.[2]

These data help explain why there are almost twice as many students who want to get into charter schools as there are slots to accommodate them. They make a compelling case for doubling the number of charter schools as quickly as possible.

[2] Performance at the tenth grade level depends on whether we consider the presence of special education students as a factor in establishing the "propensity" between charter and traditional schools.

Introduction

Charter schools are a recent development in the history of United States education reform. The US Charter Schools website describes charter schools as:

> "nonsectarian public schools of choice that operate with freedom from many of the regulations that apply to traditional public schools. … Charter schools are accountable to their sponsor – usually a state or local school board – to produce positive academic results and adhere to the charter contract. The basic concept of charter schools is that they exercise increased autonomy in return for this accountability. They are accountable for both academic results and fiscal practices to several groups: the sponsor that grants them, the parents who chose them, and the public that funds them."[3]

The first charter school law in the United States was enacted in Minnesota in 1991. There are now more than 5,400 charter schools enrolling over 1.7 million pupils in forty states.[4]

In 1993 the 178th General Court of the Commonwealth of Massachusetts amended the laws governing public schools to include a provision for the establishment and growth of charter schools in Massachusetts. The law allowed for a maximum of 120 charter schools to be operating in Massachusetts at any given time. It further mandated that "no public school district's total charter school tuition payment to commonwealth charter schools shall exceed 9% of said district's net school spending," save for a negligible sibling exception rule.[5] In 2003 the law was amended to "clarify the process for granting, amending, renewing and revoking charters."[6] The law was amended most recently in 2010 to address the strong demand for charter schools in struggling districts, as evidenced by the near-ubiquity of waiting lists for charter schools in these regions.[7]

[3] "US Charter Schools", http://www.uscharterschools.org/pub/uscs_docs/o/index.htm (accessed February 1, 2011).
[4] Center for Education Reform, "Charter Connection," http://www.edreform.com/Issues/Charter_Connection/ (accessed March 8, 2011).
[5] "Part I, Title XII, Chapter 71, Section 89," 187th General Court of the Commonwealth of Massachusetts, http://www.malegislature.gov/Laws/GeneralLaws/PartI/TitleXII/Chapter71/Section89 (accessed March 7, 2011).
[6] Massachusetts Department of Elementary and Secondary Education , "Board in Brief: Tuesday, June 24, 2003," http://www.doe.mass.edu/boe/bib/03/0624.html (accessed March 7, 2011).
[7] Department of Elementary and Secondary Education, "Massachusetts Charter School Enrollment" http://www.doe.mass.edu/charter/enrollment/ (accessed March 7, 2011).

A total of 26,708 students were on waiting lists for charter schools as of April 2010, which means there was almost one student on the waiting list for every student already enrolled.

The recent amendment raised the 9% cap on commonwealth charter school spending to 18% for districts in the lowest 10% of all statewide student test scores[8] and was instrumental in securing federal "Race to the Top" funding offered by the Obama Administration. By one measure, the laws and rules governing charter schools in Massachusetts put it among the top three states in the country (Ziebarth 2011).

Charter schools operate independently of any public school committee and are managed by a board of trustees. Groups eligible to submit a charter school application include (but are not limited to) two or more certified teachers, any non-profit business or corporate entity, and ten or more parents. Boards may favor "schools that have demonstrated broad community support, an innovative educational plan and a demonstrated commitment to assisting the district in which it is located in bringing about educational change."[9] Charters are granted and renewed for five year periods, although they may be revoked at the Board of Elementary and Secondary Education's discretion. The Massachusetts Department of Elementary & Secondary Education explains: "The renewal of a public school charter is contingent on affirmative evidence in three areas: faithfulness to the terms of the charter, academic program success and organizational viability."[10]

By November 2010, there were over 28,000 pupils in 63 charter schools in Massachusetts.[11] A quarter of the charter schools are in the city of Boston. Since 1994, the state has received 392 applications to establish charter schools, and granted charters to 77 of these. Six of these charters were revoked or not renewed, and eight were surrendered for other reasons.

[8] "Part I, Title XII, Chapter 71, Section 89," 187th General Court of the Commonwealth of Massachusetts, http://www.malegislature.gov/Laws/GeneralLaws/PartI/TitleXII/Chapter71/Section89 (accessed March 7, 2011).
[9] Ibid.
[10] "Massachusetts Primer on Special Education and Charter Schools," Massachusetts Department of Elementary & Secondary Education, February 2009.
[11] Commonwealth of Massachusetts Department of Elementary and Secondary Education, *Massachusetts Charter Schools Fact Sheet*, Version updated to November 4, 2010.

Do Charter Schools Perform Better?

This report investigates two fundamental questions: Have charter schools proven to be an effective alternative to traditional public schools for educating children in Massachusetts? If so, then to what factors should we attribute their success?

Evidence from charter schools in New York (Hoxby et al. 2009), and Boston (Abdulkadiroglu et al. 2009) suggests that charter schools do boost student performance relative to the alternative of traditional public schools. Both of these excellent studies take advantage of the fact that more people want to attend charter schools in New York and Boston than there are available places; in such cases, students are chosen by lottery. The idea is then to compare the performance of those who were (randomly) accepted into charter schools with those who were (randomly) turned away. This automatically controls for many potentially unobservable— but important— variables, such as student or parental motivation.

Our analysis is done at the level of an individual grade (e.g. fourth graders in school X, or eighth graders in school Y), for which we have exam results. We combine this information with socioeconomic and financial information that is available at the level of individual schools and school districts. The measures of performance on which we focus are the scores on the Massachusetts Comprehensive Assessment System (MCAS) exams in mathematics, and English Language Arts, in 2007 and 2008, for grades four, eight, and ten.

Other conceivably useful performance measures might include graduation rates, college admission rates, dropout rates, and aptitude test scores, but adequate data are not available for such measures, or their relevance is limited; for example, while high school graduation rates may be interesting, the proportion of children who move up from, say, grade four to grade five is a lot less revealing.

Every student in a publicly-funded school in Massachusetts is obliged to take the MCAS exams. The underlying numerical grades, which we use in our analysis and run from 0 (bad) to 100 (good), are then sorted into four categories: above proficient, proficient, needs improvement, or fail.

Table 1 shows that the average Composite Performance Index (CPI) score in mathematics in 2007-2008 was 70.2 for charter schools and 78.5 in non-charter schools

among 4th graders.[12] Table 1 also shows that test results for eighth grade mathematics are very similar. It reports that the average score in English was 91.1 for charter schools and 90.0 in non-charter schools among 10th graders. In most cases non-charter school exam scores are higher, or only slightly below, those of charter schools.

Table 1: Raw MCAS CPI Scores of Non-Charter Public Schools and Charter Schools, 2007-2008 averages

	Mathematics		English Language Arts	
Test results 4th Grade	Non-charter	Charter	Non-charter	Charter
Performance score	78.5	70.2	80.3	71.5
% above proficient	20.1	13.0	9.4	4.1
% proficient	30.1	25.0	44.4	33.2
% need improvement	38.3	45.2	35.9	48.7
% failing	11.6	16.9	10.4	13.9
Test results 8th Grade				
Performance score	72.4	72.3	90.3	91.5
% above proficient	19.1	16.9	12.9	10.1
% proficient	29.8	31.3	64.0	68.6
% need improvement	28.6	31.1	17.5	17.9
% failing	22.6	20.8	5.7	3.5
Test results 10th Grade				
Performance score	87.0	86.6	90.0	91.1
% above proficient	43.8	39.6	23.5	18.6
% proficient	28.4	31.9	50.7	57.2
% need improvement	20.2	21.6	21.7	22.0
% failing	7.7	6.9	4.1	2.3

Notes and sources: Based on Tables A1, A2, and A3. The performance varies from 0 (all fail) to 100 (all are "above proficient").

However, these simple comparisons of MCAS scores between charter and non-charter schools are misleading, because they do not measure the difference in performance between the two groups. This is because charter schools may serve a different socio-demographic constituency — for instance, they may serve relatively more poor students — and so might be expected to show lower performance levels. This is relevant, because it is well established that exam performance is linked to socio-economic status.

[12] Instead of raw test scores, we use the Composite Performance Index (CPI). The CPI is a 100-point index that combines the scores of students who take standard MCAS tests (the Proficiency Index) with the scores of those who take the MCAS-Alternate Assessment (MCAS-Alt) and is a measure of the extent to which students are progressing toward proficiency in ELA and mathematics, respectively. We divide the CPI by 100 to derive an index between 0 and 1.

Table 2: Characteristics of Charter and Non-Charter Public Schools, 2007-2008 averages

Pupil Characteristics	4th Grade		8th Grade		10th Grade	
Percentage of Grade Enrollment:	Non-Charter	Charter	Non-Charter	Charter	Non-Charter	Charter
Male	50.3	46.5	50.7	47.9	50.2	45.1
Native American*	0.3	0.4	0.3	0.3	0.3	0.3
Asian*	4.9	5.4	4.6	2.9	4.6	2.1
African American*	7.2	22.4	7.7	24.9	7.7	31.9
Hispanic*	12.8	24.8	12.3	18.7	12.0	17.8
Caucasian	74.8	47.2	75.2	53.4	75.5	48.1
With subsidized lunch*	28.8	44.5	29.4	40.2	24.5	42.8
First language not English*	14.1	17.1	14.4	12.0	14.0	10.9
Limited English proficiency*	7.0	5.6	4.4	2.8	3.3	0.7
Special education student*	15.7	10.5	16.9	13.4	15.1	12.3
District EQV per capita ($'000)*	163.8	88.1	162.1	96.5	163.3	103.1
School Information						
Enrolment in 4th grade*	98.6	87.4	228.9	71.6	317.9	58.6
Attendance rate, school	95.7	95.4	95.3	95.4	93.1	93.6
In-school suspension rate	0.5	4.5	4.4	5.2	6.4	3.1
Out-of-school suspension rate	1.6	6.2	6.8	11.5	10.4	11.8
Truancy rate	1.5	4.3	1.7	0.6	2.2	0.8
Repetition rate	1.2	3.7	1.2	5.0	3.9	8.5
Current spending/pupil, $	9,345.5	9,233.5	9,425.5	9,428.0	9,984.0	9,852.0
Share of spending on admin.	3.5	10.7	3.4	11.4	3.6	12.1
Teacher Information						
Full time Equivalent in school*	34.1	57.6	54.0	40.8		
Percentage licensed*	98.1	73.1	95.9	61.7	95.2	63.3
Percentage highly qualified	97.6	87.9	94.3	87.6	94.1	89.6
Pupil-teacher ratio	14.3	14.0	13.3	12.6	13.8	12.5
Average teacher salary, $/yr	50,123.5	42,904.0	50,324.5	42,582.5	50,502.0	44,269.0

Source: Tables A1-A3.

Before going further, we therefore need to establish whether charter schools look different from non-charter schools. The answer is that they do, and this shows clearly in Table 2. Charter schools serve poorer areas. There are two main measures of this: upwards of 40% of pupils in charter schools received free or subsidized lunches, compared to about 25-30% in non-charter schools. The second measure shows the property valuation per capita; it averaged less than $100,000 in towns served by charter schools, compared to more than $150,000 in towns served by non-charter public schools.

The demographic character of charter schools is also different. About one in four of their students is African-American, which is three times the proportion found in non-charter schools; they also have about one-and-a-half times the proportion of Hispanic students as non-charters. Although charter schools are about the same size as non-charter schools at the 4th grade level, they are much smaller at the 8th grade and 10th grade levels: the average 10th grade in a charter school has about 60 pupils, compared to over 310 in its non-charter counterparts.

Charter schools have been criticized for not shouldering their fair share of students who face language barriers or learning disabilities.[13] The numbers in Table 2 support this view, at least to some extent. The proportion of charter students whose first language is not English is higher than in public schools at the fourth grade level, but not at the eighth or tenth grades. One interpretation is that charter schools are not attracting non-native speakers, or that non-native speakers are dropping out; another view is that charters have begun to make a greater effort to attract non-native speakers, but the effects are only showing up at the fourth grade level for now.

Charter schools have lower proportions of students with limited English proficiency than public schools. This is more difficult to interpret; some might argue that charters are deterring students with poor English language skills, while others might argue that charters are doing a better job at bringing poor speakers up to speed and pulling them out of the ranks of those with limited English proficiency. Our data do not allow us to choose between these two interpretations.

At every grade level, the proportion of special education students in charter schools is between three and five percentage points below the proportion in public schools. Here too, the interpretation is far from clear. The procedures for determining whether a student has special needs begin at the school level with interactions between the parents, specialist personnel, and teachers. It is certainly possible that charter schools do not welcome special education students; it is also possible that charter schools may do a better job of mainstreaming them into the classroom.

Charter schools in Massachusetts are required to comply with all the same regulations concerning the provision of special education as are other public schools, except that they do not have to provide out-of-district placements to "students who may require a full time day or residential program of special education," which is typically funded

[13] See for instance James Vaznis, "Charter schools lag in serving the neediest," *Boston Globe*, August 12, 2009.

separately.[14] Specifically, charter schools "may not discriminate in admissions against students on the basis of disability, special need or prior academic achievement."[15]

The substantial differences in student composition between charter and non-charter schools need to be taken into account when making comparisons. To illustrate this point, we compare the scores of charter with non-charter schools within the City of Boston (see Tables A1-A3). In 10 of the 12 comparisons, Boston charters outperform their peers in the city, and the gap is strong and persistent in the higher grades. In other words, when we come closer to comparing "like-with-like" schools, the performance of charter schools looks relatively better.

A more satisfactory approach is to match each charter school with one or more "similar" non-charter schools, where similarity is based on socio-demographic characteristics. One of the most common ways to do this is with propensity score matching, which is discussed more fully in the Appendix. The first step is to estimate a model of "participation" – in other words, to estimate an equation where the dependent variable is 1 if the school is a charter, and 0 otherwise, and the independent variables include such things as the district equalized valuation ratio per capita, the share of African-American children in the school, the proportion of students that receive free or subsidized lunches, the share of students with limited English proficiency, and so on. This model is used to predict whether a school is likely to be a charter or not – its "propensity score" – and is refined to ensure balance.

Then, each charter school is matched with the non-charter that has the closest propensity score (nearest neighbor matching), or a weighted average of those with close propensity scores (kernel matching), or those in the same school district that have not-too-different propensity scores (stratification matching). For each charter, we compute the difference in exam scores between the charter and the match-mate(s), and take the overall average; this is the "average treatment effect of the treated" (ATT), which we explain the Appendix below. In making comparisons, we confine ourselves to the "region of common support", where we are likely to find propensity scores of both charters and non-charters, so that the matches are indeed plausible. This is because it would not be meaningful to compare a charter school with a non-charter school that has wildly different characteristics.

[14]Massachusetts Department of Elementary and Secondary Education, "Massachusetts Primer on Special Education and Charter Schools," February 2009.
[15]*Ibid.*

The results of this procedure are shown in Table 3. Consider, for example, the results for grade eight for mathematics in 2008. The average score for charter schools was 72.4, compared to 73.3 for non-charters, a difference of -0.9. But when we compare charter schools just with peer non-charter schools that have similar socio-demographic characteristics it turns out that the MCAS scores are significantly higher at the charter schools – by somewhere between 7.2 and 13.2 points (out of 100), depending on the matching method used. The t-statistics are high enough to indicate that these results are statistically significant, in the sense that they are indeed greater than zero and are not the product of pure chance.

The results in Table 3 are at the heart of our analysis. They show that at the fourth grade level, charter schools did not perform very differently from non-charter public schools in 2007 and 2008, at least as measured by MCAS scores in mathematics and in English Language Arts. The t-statistics accompanying the measured differences in exam scores are low, indicating that the reported results may reflect chance rather than a strong underlying process.

By the eighth grade, the exam scores for charter schools appear to be solidly higher for charter than for comparable non-charter schools. Our results are entirely consistent with the findings by Abdulkadiroglu et al. (2009) for Boston schools. The variables used in the propensity score equations are noted with an asterisk in Table 2. Among the more important are the following:

- the proportion of students who come from different ethnic backgrounds;
- the proportion of students who are poor (as measured by whether they get free lunches);
- the proportion of students defined as having special education needs,
- the proportion of students who are have no difficulty communicating in English,
- the extent to which the catchment area of the school is a low-income area (as measured by property values per capita in the school district).

Table 3: Estimated Differences in CPI Scores, Propensity Score Matching model

	Mathematics		English Language Arts	
	2007	2008	2007	2008
Grade 4				
ATT, nearest neighbor	**5.7**	**1.6**	**13.1**	**-3.4**
t-statistic	*0.46*	*0.32*	*1.11*	*-0.83*
ATT, stratified by district	**3.9**	**-2.2**	**1.4**	**-2.9**
t-statistic (bootstrap)	*0.35*	*-0.58*	*0.14*	*-0.88*
Memo items:				
Raw score, non-charters	78.0	79.0	82.1	78.5
Raw score, charters	69.3	71.1	74.2	68.7
Difference	-8.7	-7.9	-7.9	-9.8
Grade 8				
ATT, nearest neighbor	**7.1***	**7.2***	**2.8**	**5.6***
t-statistic	*1.46*	*1.36*	*1.06*	*1.70*
ATT, stratified by district	**16.4***	**13.2***	**8.2***	**8.9***
t-statistic (bootstrap)	*2.78*	*2.68*	*2.30*	*2.38*
Memo items:				
Raw score, non-charters	71.4	73.3	90.4	90.1
Raw score, charters	72.2	72.4	92.0	90.9
Difference	0.8	-0.9	1.6	0.8
Grade 10				
ATT, nearest neighbor	**-11.8**	**2.1**	**-5.0**	**7.5***
t-statistic	*-0.72*	*0.21*	*-0.37*	*1.60*
ATT, stratified by district	**9.8**	**11.7**	**10.3**	**8.3***
t-statistic (bootstrap)	*0.97*	*1.01*	*1.13*	*1.62*
Grade 10, excl. special edn.				
ATT, nearest neighbor	**-9.69**	**22.6***	**-3.6**	**12.1***
t-statistic	*-0.69*	*2.62*	*-0.30*	*2.00*
ATT, stratified by district	**13.5**	**11.7**	**12.2***	**10.1***
t-statistic (bootstrap)	*1.19*	*1.01*	*1.38*	*1.78*
Memo items:				
Raw score, non-charters	86.2	87.8	88.9	91.1
Raw score, charters	86.6	86.6	90.0	92.1
Difference	0.4	-1.2	1.1	1.0

*Note: Independent variables in the propensity score equations include all those items marked with * in Table 2. *statistically significant at the 10% level. ATT stands for "Average Treatment Effect of the Treated."*

When comparisons are made on a district-wide basis ("stratification matching"), the charter school advantage in exam results is maintained, and at the tenth grade level even widened. However, the results at the 10th grade level are somewhat sensitive to the method of matching that is used; with nearest neighbor matching, charter schools

had no particular advantage in 2007, although they did do better than non-charters in 2008.

The differences between the upper and lower panels for Grade 10 in Table 3 reflect different approaches to handling the variable that measures the proportion of special needs students; when the variable is included, the charter school edge is smaller; but if the number of special needs students is endogenous, in the sense that charter schools may have found a way to educate these students without special classes, then it is not appropriate to include this variable in the propensity score equation, and the estimates in the lower panel of that part of Table 3 are the appropriate ones.

It is legitimate to ask how much difference it makes that the 2007 score in eighth grade mathematics for charter schools was, on average, 7.1 points better than the same score for "the nearest neighbor" non-charter schools. One way to address this question is to ask how much an improvement this represents over the raw score for all non-charter schools, which was 71.4 in this instance. By this measure, charter schools raised performance in eighth grade mathematics by 9.9%. If we consider only the statistically significant results on Table 3, we can compute similar percentage improvements for each subject, for eighth and tenth grades, and so we obtain a range of estimated improvement generated by charter schools as follows:

- 6.2% to 9.9% improvement over the scores for traditional schools in eighth grade English;
- 9.8% to 23.0% improvement over the scores for traditional schools in eighth grade mathematics;
- 25.7% improvement over the scores for traditional schools in tenth grade mathematics; and
- 8.2% to 13.7% improvement over the scores for traditional schools in tenth grade English.[16]

Another technique for measuring the effect of charter schools on student performance is regression analysis. The dependent variable is the MCAS score – for instance, for fourth graders in mathematics for 2007 – and the unit of observation is the class. The

[16] Performance at the tenth grade level depends on whether we consider the presence of special education students as a factor in establishing the "propensity" between charter and traditional schools.

independent variables include the same socio-economic variables employed in the propensity score matching analysis reported in Table 3; they are indicated with an asterisk (*) in Table 2.

Table 4: Estimated Effects of Charter Schools in CPI Scores

	Mathematics		English Language Arts	
	2007	2008	2007	2008
Grade 4				
Regression, basic	**-2.68**	**-1.87**	**-0.43**	**-3.46***
p-value / N	*0.23 / 993*	*0.31 / 984*	*0.85 / 993*	*0.03 / 984*
Regress, incl. special edn.	**-3.60**	**-2.64**	**-1.69**	**-4.02***
p-value / N	*0.11 / 993*	*0.17 / 984*	*0.48 / 993*	*0.01 / 984*
Memo: raw difference	**-8.7**	**-7.9**	**-7.9**	**-9.8**
Grade 8				
Regression, basic	**8.63***	**5.73***	**5.29***	**4.71***
p-value / N	*0.00 / 459*	*0.03 / 462*	*0.00 / 460*	*0.00 / 462*
Regress, incl. special edn.	**6.93***	**3.41**	**2.85***	**1.92**
p-value / N	*0.01 / 459*	*0.15 / 462*	*0.04 / 460*	*0.11 / 462*
Memo: raw difference	**0.8**	**-0.9**	**1.6**	**0.8**
Grade 10				
Regression, basic	**9.59***	**4.19**	**7.86***	**4.45***
p-value / N	*0.00 / 340*	*0.15 / 345*	*0.00 / 340*	*0.02 / 348*
Regress, incl. special edn.	**6.39***	**0.89**	**5.28***	**1.55**
p-value / N	*0.01 / 340*	*0.75 / 345*	*0.00 / 340*	*0.36 / 348*
Memo: raw difference	**0.4**	**-1.2**	**1.1**	**1.0**

*Source: As for Tables A1-A3. *statistically significant at the 10% level.*

The results of the regression analysis are shown in Table 4, where we report the effects of the charter school on exam results (in bold face), along with p-values (which are statistically significant if the values are close to zero) and the sample size. Two sets of results are shown: those with and without the variable that measures the proportion of special education children in the school.

The findings confirm our earlier analysis: charter schools are not associated with higher exam scores at the fourth grade level, but they do seem to boost grades at the eighth grade, and maintain this advantage through the tenth grade.

These results do not entirely solve the problem of sample selection bias. If the students, and parents, who want to go to charter schools are more dynamic or more motivated than their peers, then the higher exam scores at charter schools may be due to the higher quality of student, rather than superior educational delivery. The inclusion of a variety of socio-economic variables in the regressions helps attenuate this problem, but

does not guarantee that it has been reduced to insignificance. The propensity score matching is somewhat less subject to this problem, to the extent that it is able to match charter schools with comparable peer institutions more successfully.

The observation that charter schools are not associated with higher exam scores in fourth grade, but are in eighth and tenth grade, strongly suggests that something more than student sorting based on relevant unobserved characteristics is at work here; if charter school students are particularly capable or motivated, why do they not perform better at the fourth grade level? A more plausible interpretation is that charter school students do not enter with any special motivation or ability, but do better in later grades as the beneficial effects of the charter school education grow stronger the longer students are in the system.

Why Do Charter Schools Perform Better?

Our study demonstrates that charter schools, after controlling for socio-economic differences in the student population, better prepare 8th and 10th grade students for MCAS testing than do traditional public schools. We offer several suggestions that may help explain this result.

One determinant of the superior performance of charter schools over traditional public schools is likely the autonomy that charter schools enjoy in setting their own curricula. Traditionally, all students in a district are taught the same material using the same methods. The freedom granted to charter schools encourages teachers to experiment and innovate with teaching strategies that would remain underused in traditional public schools because of bureaucratic rigidity. Not all children benefit uniformly from a particular educational regime; charter schools allow teachers the freedom to cater to a variety of student needs that traditional public schools cannot meet. If legislators wish to promote a learning environment responsive to student diversity, it makes sense to decentralize the education structure and give authority to teachers and parents.

Another important factor driving charter schools' success is the increased accountability they face compared to non-charter public schools. Charter schools are judged largely by output-based metrics (such as standardized test scores), and their continued existence depends upon their ability to produce satisfactory results; non-charter public

schools do not face nearly the same degree of scrutiny. Charter schools can have their charters revoked if they fail to meet their sponsors' expectations, which creates a strong incentive for their administration and faculty to thoroughly prepare students for the standardized tests. This is not an idle threat; six of the 77 charters granted in Massachusetts since 1994 have been revoked or not renewed.

A few other points are relevant to the discussion. First, charters spend almost as much per pupil (on current spending) as do non-charter schools. They do not represent cut-price education. Second, charter schools devote almost three times as much to administration as do non-charter schools; this may reflect their smaller average size, or the need to do in-house some of the tasks that school districts undertake at the district level. Third, the pupil-teacher ratio is somewhat lower at charter schools, particularly at the higher grade levels, but the difference between charter and non-charter schools is relatively modest.

Charter schools currently pay lower salaries than non-charters; it is not clear whether this is because of greater bargaining ability, or because charter schools have a work force that may be younger, less highly credentialed, and perhaps more idealistic. In non-charter public schools 95% of teachers are licensed, and almost all of the "core" teachers are "highly-qualified." On the other hand, only between two-thirds and three-quarters of teachers in charter schools are fully licensed, and almost one in six of the core teachers are not considered to be highly qualified.

Why does this ostensible quality gap in teacher competency not manifest itself in student test scores? One possibility is that the state licensing procedure does a poor job of stressing critical attributes; or perhaps the cost and time commitment required to become fully licensed discourages some capable teachers from undertaking the procedure. The most significant factor is likely the freedom from bureaucratic oversight, which allows charter school teachers more leeway in designing courses and effectively educating students.

Charter schools also use in-school and out-of-school suspensions more frequently than non-charter schools, particularly in the 4th grade, but also in the 8th grade. They are more likely to hold students back to repeat a year (the "retention rate"). One interpretation is that charter schools set more stringent rules and set higher expectations in the early years, and the payoff comes in the form of higher academic performance in

the later years. Another possibility for why 4th grade exam scores did not reveal statistically significant differences is that the material taught may be basic enough that the advantages of charter schools are a non-factor at such an elementary level.

There are other possible explanations – perhaps a longer school year, or relatively more time devoted to English, or academically-oriented mission statements (Hoxby et al. 2009, p. viii) – that may be relevant, but we do not have enough information to be able to say much about these issues.

Conclusion

Our study demonstrates that charter schools outperform non-charter public schools on the 8th and 10th grade MCAS exams by statistically significant margins after controlling for socio-economic differences among the student populations. We offer several explanations to account for these findings, paramount among them being the freedom charter schools enjoy from bureaucratic restrictions. This freedom allows charter school teachers to innovate with methods denied to their non-charter counterparts. Charter schools appear to practice tough love as well. Although the data required to conduct more comprehensive analysis is currently lacking, our research indicates that charter schools are a viable alternative to traditional public schools. If policymakers desire improved student test scores without greater costs, they would be well served to lift the restrictions obstructing charter school growth and amend legislation governing traditional public schools to incorporate similar elements of freedom and accountability.

Appendix

Propensity Score Matching

If households are not assigned randomly to an intervention—such as food stamps, vaccinations, or schooling—then those who benefit are unlikely to be typical of the eligible population. There are two reasons for this. First, there may be non-random program placement, of which the researcher may or may not be aware; for instance, charter schools may be more likely to have been established in poor school districts. This is the problem of *unobserved area heterogeneity*. Second, there may be self-selection into program participation; for instance, more dynamic individuals may be the first to sign up, or program benefits may flow to those who are politically well-connected, or people may move to school districts that now have a charter school. Such effects are often hard to detect, and give rise to the problem of *unobserved individual and household heterogeneity*.

The presence of these unobservables immediately creates the problem of *selection bias*. The basic idea behind quasi-experimental designs is to construct statistical models of selection—including propensity score matching—that permit one to compare program participants and nonparticipants (the *comparison* group) holding the selection processes constant. [17]

To see why these problems arise, suppose we are interested in determining whether a charter school scheme, initiated in time period 0, raises the educational performance of individual i in time period 1. An appealing approach would be to collect data on the outcome indicator (performance, given by Y_{i1}), and on individual and household characteristics (X_{i1}), for a sample of individuals that do, and do not, participate in the scheme. Then we could estimate an impact equation of the form—

$$Y_{i1} = X_{i1}\beta + P_{i1}\delta + \varepsilon_{i1}^{Y}, \qquad \text{(A.1)}$$

where P_{i1} is a dummy variable that is set equal to 1 if the individual i participates in the scheme and to 0 otherwise. At first sight, it would appear that the value of the estimated

[17] This appendix draws heavily on Haughton and Khandker (2009), chapter 13.

coefficient δ would measure the impact of the charter school program on academic performance.[18]

Unfortunately, this is unlikely to be the case, because program participation is often related to the other individual, household, and community variables, some of which may not be observable. For instance, those who send their children to a charter school may be better educated, or younger, or be more motivated. With enough information it may be possible to control for many of these variables, including education and age, but it is never possible to control for all the relevant effects. For example, the degree of individual motivation is unobservable; but a more motivated individual is more likely to participate in the program (a higher P_{i1}) and to benefit more from it (a higher ε^{y}_{i1}). This creates a correlation between P_{i1} and ε^{y}_{i1} and so leads to a biased estimate of δ. As a practical matter, unobservables are always present in such circumstances, and thus this selection bias (which may also be thought of as a form of omitted variable bias) is present as well.

The path to a solution requires us to envisage a separate program *participation equation* of the form—

$$P_{ij} = Z_{i1}\gamma + \varepsilon^{P}_{i1}, \qquad \text{(A.2)}$$

where the Z variables may be the same as the X variables, or include additional variables. If one can identify a set of variables that affect only participation, equation (A.2), and not the household outcome, equation (A.1)—generally a difficult task—then it may be possible to arrive at a satisfactory estimate of δ, the impact of program participation on the outcome of interest.

One of the more popular approaches is to undertake a matching comparison. For this one needs data for a substantial number of nonparticipants (i.e. non-charter schools) as well as for the participants (charter schools). The basic idea is to match each participant with a similar nonparticipant (or a small "comparison group") and then to compare the outcomes between them.

The most common procedure starts by pooling the two samples (that is, the participants and nonparticipants, or in the jargon of matching, the treatment and comparison

[18]Equation (A.1) as shown here is linear, but other forms, including logit and probit specifications, are typically used in practice.

groups) and estimating a logit model of program participation as a function of all the variables that might influence participation—equation (A.2). Ironically, one does not want the equation to fit too well, because that would make it difficult to identify nonparticipants who are otherwise similar to participants.

The next step is to generate the *propensity score*, which is the predicted probability of participation, given the observed characteristics *Z*. Some of the individuals in the comparison group may have propensity scores that are far outside the range of the treatment group—outside the range of "common support"—and these cases may need to be dropped.

Next, for each person in the treatment group, we have to find the member of the comparison group with the closest propensity score ("nearest neighbor" matching), or a group of neighbors, putting the highest weight on those who are most similar ("kernel density matching"), or making the comparison with those in the same district ("stratification matching"). We then compute the difference between the outcome indicator for the charter school, and the mean of the outcome indicators for the nearest neighbors. The mean of these differences, over all the members of the treatment group, gives a measure of the overall impact of the program, as measured by the average treatment effect of the treated (ATT).

When the correlates of participation in the project are observable, this approach works well, but it is less satisfactory if unobservable differences are important—for instance, if those who sign their children up for charter schools are the more dynamic households. The procedure falls short in this case because the differences between the treatment and comparison groups cannot be entirely attributed to whether or not they participated in the program; some of the difference may be due to (possibly unobserved) differences in the inherent characteristics of individuals or households in the two groups.

Table A1. Performance of Charter and Non-Charter Schools in Grade 4, for 2007 and 2008

	Mathematics				English Language Arts			
	2007		2008		2007		2008	
Test results	Non-charter	Charter	Non-charter	Charter	Non-charter	Charter	Non-charter	Charter
Performance score	78.0	69.3	79.0	71.1	82.1	74.2	78.5	68.7
% above proficient	19.4	13.1	20.8	12.8	10.5	4.6	8.3	3.6
% proficient	30.1	22.8	30.1	27.1	47.0	37.4	41.8	28.9
% need improvement	39.0	47.5	37.5	42.8	33.5	46.4	38.2	50.9
% failing	11.5	16.5	11.7	17.2	9.0	11.4	11.7	16.4
Boston: Performance score	63.3	62.9	67.1	64.2	65.1	70.1	62.8	65.3
Number of schools	972	24	960	24	972	24	960	24
Number taking exam	94.5	84.3	96.1	87.9	94.5	84.4	96.0	87.9
Background information on pupils								
4th grade % male	50.3	46.7	50.3	46.2				
4th grade % Native American	0.3	0.3	0.3	0.4				
4th grade % Asian	4.8	6.6	5.0	4.1				
4th grade % African American	7.2	22.1	7.2	22.6				
4th grade % Hispanic	12.5	24.0	13.1	25.5				
4th grade % Caucasian	75.2	47.0	74.4	47.4				
4th grade % subsidized lunch	28.5	45.7	29.0	43.3				
First language is not English	13.9	17.2	14.3	17.0				
Limited English proficiency	6.8	5.2	7.2	5.9				
Special education student	15.5	10.6	15.8	10.4				
District EQV/capita ($'000)	160.5	85.0	167.0	91.2				
Information on school								
Enrolment in 4th grade	97.8	85.8	99.4	89.0				
FTE teachers in school	35.1	57.5	33.1	57.6				
% of teachers licensed	97.8	74.2	98.4	71.9				
% core teachers high-qualified	97.3	87.9	97.9	87.9				
Pupil-teacher Ratio	13.9	13.6	14.7	14.4				
Attendance rate, school	95.7	95.2	95.6	95.6				
In-school suspension rate	0.5	4.7	0.5	4.3				
Out-of-school suspension rate	1.5	5.6	1.7	6.8				
Truancy rate	1.5	5.6	1.5	3.0				
Repetition rate	1.2	3.7						
Current spending/pupil, $	9,220	9,341	9,471	9,126				
Average teacher salary, $/yr	48,272	40,036	51,975	45,772				
Share of spending on admin	3.5	10.5	3.5	10.8				

Table A2. Performance of Charter and Non-Charter Schools in Grade 8, for 2007 and 2008

	Mathematics				English Language Arts			
	2007		2008		2007		2008	
Test results	Non-charter	Charter	Non-charter	Charter	Non-charter	Charter	Non-charter	Charter
Performance score	71.4	72.2	73.3	72.4	90.4	92.0	90.1	90.9
% above proficient	18.0	17.6	20.1	16.1	12.8	9.9	12.9	10.2
% proficient	28.7	30.3	30.9	32.3	64.0	69.4	64.0	67.8
% need improvement	30.3	30.8	26.9	31.4	17.8	17.7	17.1	18.1
% failing	23.0	21.3	22.2	20.3	5.3	3.0	6.1	4.0
Boston: Performance score	57.8	86.8	60.8	84.8	80.3	95.6	80.1	95.2
Number of schools	420	41	419	43	421	41	419	43
Number taking exam	222.1	69.3	217.6	71.5	222.7	69.4	217.7	71.7
Background information on pupils								
8th grade % male	50.8	47.0	50.6	48.8				
8th grade % Native American	0.3	0.2	0.3	0.3				
8th grade % Asian	4.4	2.6	4.7	3.1				
8th grade % African American	7.9	25.0	7.5	24.7				
8th grade % Hispanic	12.2	17.4	12.4	20.0				
8th grade % Caucasian	75.2	54.8	75.1	52.0				
8th grade % subsidized lunch	29.3	40.4	29.4	40.0				
First language is not English	14.4	12.1	14.4	11.9				
Limited English proficiency	4.3	2.7	4.4	2.9				
Special education student	16.7	13.5	17.0	13.3				
District EQV/capita ($'000)	158.9	96.0	165.2	97.0				
Information on school								
Enrolment in 8th grade	231.9	69.2	225.9	73.9				
FTE teachers in school	54.8	39.9	53.2	41.6				
% of teachers licensed	95.2	66.5	96.5	56.8				
% core teachers high-qualified	93.8	85.0	94.8	90.2				
Pupil-teacher Ratio	13.2	12.3	13.4	12.9				
Attendance rate, school	95.2	95.3	95.3	95.4				
In-school suspension rate	4.3	5.3	4.4	5.0				
Out-of-school suspension rate	6.9	10.8	6.6	12.1				
Truancy rate	1.7	0.5	1.7	0.7				
Repetition rate	1.2	5.0						
Current spending/pupil, $	9,308	9,328	9,543	9,528				
Average teacher salary, $/yr	48,454	39,687	52,195	45,478				
Share of spending on admin	3.4	11.2	3.4	11.5				

Table A3. Performance of Charter and Non-charter Schools in Grade 10 for 2007 and 2008

	Mathematics				English Language Arts			
	2007		2008		2007		2008	
Test results	Non-charter	Charter	Non-charter	Charter	Non-charter	Charter	Non-charter	Charter
Performance score	86.2	86.6	87.8	86.6	88.9	90.0	91.1	92.1
% above proficient	43.0	38.3	44.6	40.9	22.6	16.0	24.4	21.1
% proficient	27.8	32.3	29.0	31.5	49.6	56.8	51.8	57.5
% need improvement	21.3	23.3	19.0	19.9	23.0	24.8	20.4	19.2
% failing	8.0	6.0	7.4	7.8	4.8	2.3	3.4	2.2
Boston: Performance score	78.6	84.7	80.7	84.3	79.2	85.3	83.0	89.1
Number of schools	315	25	321	24	315	25	323	25
Number taking exam	291.7	51.7	284.0	57.7	295.7	52.9	285.6	57.1
Background information on pupils								
8th grade % male	50.2	46.6	50.1	43.5				
8th grade % Native American	0.3	0.4	0.3	0.2				
8th grade % Asian	4.6	2.4	4.6	1.7				
8th grade % African American	7.7	31.1	7.7	32.6				
8th grade % Hispanic	11.9	16.8	12.0	18.7				
8th grade % Caucasian	75.6	49.3	75.4	46.9				
8th grade % subsidized lunch	24.1	42.7	24.8	42.8				
First language is not English	14.0	10.0	13.6	11.8				
Limited English proficiency	3.3	0.5	3.2	0.8				
Special education student	15.0	12.3	15.1	12.3				
District EQV/capita ($'000)	160.4	97.3	166.2	108.9				
Information on school								
Enrolment in 10th grade	322.9	56.3	312.8	60.9				
FTE teachers in school								
% of teachers licensed	95.0	67.8	95.4	58.7				
% core teachers high-qualified	94.0	84.9	94.1	94.3				
Pupil-teacher Ratio	13.7	12.2	13.8	12.7				
Attendance rate, school	93.0	93.8	93.2	93.3				
In-school suspension rate	6.2	4.1	6.6	2.1				
Out-of-school suspension rate	10.5	12.2	10.2	11.3				
Truancy rate	2.1	0.1	2.2	1.4				
Repetition rate	3.9	8.5						
Current spending/pupil, $	9,838	9,806	10,130	9,898				
Average teacher salary, $/yr	48,537	40,501	52,467	48,037				
Share of spending on admin	3.6	12.3	3.6	11.9				

References

Abdulkadiroglu, A., Angrist, J., Cohodes, S., Dynarski, S., Fullerton, J., Kane, T., Pathak, P. (2009). "Informing the Debate. Comparing Boston's charter, pilot and traditional schools." The Boston Foundation. (January).

Abdulkadiroglu, A., Angrist, J., Dynarski, S., Fullerton, J., Kane, T., Pathak, P. (2009). "Accountability and flexibility in public schools: evidence from Boston's charters and pilots." *NBER Working Paper* 15549. (November).

Bingham, R., Heywood, J., White, S. (1991) "Evaluating schools and teachers based on student performance. Testing an Alternative Methodology." *Evaluation Review* 15 (2): 191-218.

Dee, T., Levin, J. (2004) "The Fate of New Funding: Evidence from Massachusetts' Education Finance Reforms," *Educational Evaluation and Policy Analysis* 26 (3): 199-215.

Downes, Thomas, Jeffrey Zabel, and Dana Ansel (2009). *Incomplete Grade: Massachusetts Education Reform at 15*, MassInc, Boston MA.

Hanushek, E. (1986). "The economics of schooling: Production and efficiency in public schools," *Journal of Economic Literature* 24 (3): 1141-1177.

Haughton, Jonathan and Shahidur Khandker (2009). *Handbook on Poverty and Inequality*, World Bank, Washington DC.

Hoxby, Caroline, Sonali Murarka, and Jenny Kang (2009). How New York City's Charter Schools Affect Achievement. New York City Charter Schools Evaluation Project, Stanford CA.

Jaggia, S., Kelly-Hawke, A. (1999). "An Analysis of the factors that influence student performance: a fresh approach to an old debate." *Contemporary Economic Policy*, 17 (2): 189-198.

Therriault, Susan Bowles, Allison Gruner Gandhi, Julia Casasanto, and Samantha Carney (2010). Out of the Debate and Into the Schools: Comparing Practices and Strategies in Traditional, Pilot and Charter Schools in the City of Boston. Boston Foundation, Boston.

Ziebarth, Todd (2011). Measuring Up to the Model: A Ranking of State Charter School Laws (2nd edition), National Alliance for Public Charter Schools, Washington DC.

The Beacon Hill Institute Policy Study Team

David G. Tuerck is Executive Director of the Beacon Hill Institute for Public Policy Research at Suffolk University where he also serves as Chairman and Professor of Economics. He holds a Ph.D. in Economics from the University of Virginia and has written extensively on issues of taxation and public economics.

Paul Bachman is Director of Research at BHI. He manages the institute's research projects, including the STAMP model and conducts research on other projects at the BHI. Mr. Bachman has authored research papers on state and national tax policy and on state labor policy. He also produces the institute's state revenue forecasts for the Massachusetts legislature. He holds a Master of Science in International Economics from Suffolk University.

Jonathan Haughton is a Professor of Economics at Suffolk University. A specialist in the areas of economic development, international trade, and taxation, he has published widely, and lectured, taught, or conducted research in over 20 countries. He received his B.A. (Mod.) in Economics from Trinity College, Dublin and his Ph.D. in Economics from Harvard University.

Eli Cutler is a Ph.D. candidate at Suffolk University and a research assistant at the Beacon Hill Institute. Mr. Cutler received his B.A. in Economics from the University of Connecticut.

The authors would like to thank Frank Conte, BHI Director of Communications for editorial assistance and Michael Head for technical assistance.

The Beacon Hill Institute at Suffolk University in Boston focuses on federal, state and local economic policies as they affect citizens and businesses. The Institute conducts research and educational programs to provide timely, concise and readable analyses that help voters, policymakers and opinion leaders understand today's leading public policy issues.

THE BEACON HILL INSTITUTE
FOR PUBLIC POLICY RESEARCH
Suffolk University
8 Ashburton Place
Boston, MA 02108
Phone: 617-573-8750 Fax: 617-994-4279
bhi@beaconhill.org
http://www.beaconhill.org