

Education Working Paper Archive

Apples to Apples:

An Evaluation of Charter Schools Serving General Student Populations

April 4, 2006

Jay P. Greene

Senior Fellow, Manhattan Institute for Policy Research

Greg Forster

Senior Research Associate, Manhattan Institute for Policy Research

Marcus A. Winters

Research Associate, Manhattan Institute for Policy Research

Executive Summary

Charter schools—public schools that are exempt from many of the procedural regulations that apply to regular public schools—are a widespread but poorly-studied form of education reform. With nearly 2,700 charter schools now educating more than 684,000 children nationwide, policymakers and parents need to know how the education charter schools provide compares to that provided by regular public schools.

Assessing the academic performance of charter schools is difficult, because many charter schools are targeted toward specific populations such as at-risk students, disabled students, and juvenile delinquents. This makes it very challenging for researchers to draw a fair comparison—comparing targeted charter schools to regular public schools is like comparing apples and zebras. As a result, there are very few reliable research findings on the academic quality of charter schools as compared to regular public schools.

This is the first national empirical study of charter schools that compares apples to apples—that is, test scores at charter schools and regular public schools serving similar student populations. By comparing “untargeted” charter schools serving the general population to their closest neighboring regular public schools, we can draw a fair comparison and get an accurate picture of how well charter schools are performing.

Measuring test score improvements in eleven states over a one-year period, this study finds that charter schools serving the general student population outperformed nearby regular public schools on math tests by 0.08 standard deviations, equivalent to a benefit of 3 percentile points for a student starting at the 50th percentile. These charter schools also outperformed nearby regular public schools on reading tests by 0.04 standard deviations, equal to a benefit of 2 percentile points for a student starting at the 50th percentile.

The study’s strongest results came in Florida and Texas. In Texas, charter schools achieved year-to-year math score improvements 0.18 standard deviations higher than those of comparable regular public schools, and reading score improvements 0.19 standard deviations higher. These benefits are equivalent to 7 and 8 percentile points, respectively, from the 50th percentile. Florida charter schools achieved year-to-year math and reading score improvements that were each 0.15 standard deviations greater than those of nearby regular public schools, equivalent to a gain of 6 percentile points for a student starting at the 50th percentile.

Acknowledgements

The authors would like to thank all of the charter school and state education officials who helped provide data for this study.

Introduction

Charter schools—public schools that are exempt from many of the procedural regulations that apply to regular public schools—are a widespread but poorly-studied form of education reform. Because charter schools provoke less political resistance than other kinds of school choice, they have spread across the country rapidly; there are now nearly 2,700 charter schools spread across 39 states and the District of Columbia and educating more than 684,000 students.[1] With the size and rapid growth of the charter school movement, policymakers and researchers need to know whether the education provided by charter schools is superior to that available in traditional public schools.

But assessing the academic performance of charter schools is difficult, because many charter schools serve specifically targeted populations such as at-risk students, disabled students, and juvenile delinquents. This makes it very difficult for researchers to draw a fair comparison between charter schools and regular public schools; simply comparing the performance of regular public school students to the performance of charter school students is like comparing apples and zebras. As a result, there are very few reliable research findings on the academic quality of charter schools as compared to regular public schools.

A comparison between charter schools and regular public schools that all serve general student populations, however, would be fair—a comparison of apples to apples. Using this approach, we found that untargeted charter schools—that is, those that serve general student populations—perform moderately better than neighboring regular public schools. Our overall national analysis found that over a one-year time period untargeted charter schools outperformed nearby public schools on math tests by 0.08 standard deviations, which is the equivalent of a benefit of 3 percentile points for a student starting at the 50th percentile. We also found that untargeted charter schools outperformed nearby public schools on reading tests by 0.04 standard deviations, equal to a benefit of 2 percentile points from the 50th percentile. Both these results were statistically significant at a high level. In individual states, we found inconclusive results in Arizona, California, and North Carolina, and statistically significant positive results in Florida and Texas. Some states may have had inconclusive results because of small sample sizes. Our strongest results, in Texas, found that charter schools achieved year-to-year score improvements of 0.18 standard deviations relative to comparable public schools on standardized math tests, and 0.19 standard deviations on reading tests. These benefits are equivalent to 7 and 8 percentile points, respectively, from the 50th percentile.

Previous Research

Charter schools are public schools that operate outside the normal bureaucratic and regulatory school system. They are created by special agreement between the school itself and an authorizing agency. State laws differ regarding what institutions can authorize charter schools; in some states only local school boards can do so, while in other states the state department of education or state universities can also serve as authorizing agencies. In all states with charter school laws, however, the agreement

between the school and the authorizer, called the school's "charter," lays out the terms by which the school is to be run. Charter schools are schools of choice; students may choose to go to a charter school instead of the regular public school to which they are assigned. In exchange for freedom from many of the usual rules and regulations public schools must follow (for example, rules covering the hiring and firing of teachers), charter schools can be held directly accountable for their performance by their authorizers.

Despite the rapid growth of charter schools, relatively little is known about whether charter schools perform better than regular public schools. One of the most important difficulties in studying charter schools is that many of them are targeted specifically at particular student populations, and thus serve dramatically different kinds of students than regular public schools do. Although most states require charter schools to have open enrollment policies, charter schools can still target specific populations by describing themselves as schools for a particular kind of student or by otherwise encouraging a certain kind of student to apply for admission. A study by the U.S. Department of Education found that one quarter of charter schools are targeted toward certain populations (see Nelson, et. al. 2000). Many people familiar with charter schools would find this a very conservative estimate; as we will see, this impression is confirmed by our findings in this study.

Targeted charter schools usually serve populations that are disadvantaged. Gronberg and Jansen find that Texas charter schools serve more black students, more non-Anglo students, and more at-risk students than regular public schools (see Gronberg and Jansen 2001). Fuller and his co-authors attempt to downplay the differences between charter school student populations and regular public school student populations by noting that the percentage of students in the federal free and reduced lunch program is similar at both kinds of schools. But the free and reduced lunch program serves a very broad population—it includes about 40% of all students. Because free and reduced lunch status does not differentiate between lower middle class students and students living in extreme poverty, it does not accurately identify the pool of students who are so poor that they are truly "at risk" educationally. Fuller and his co-authors themselves find that 44% of charter schools are identified as having a "special program focus" or as being an "alternative school" (see Fuller, et. al. 2003).

Because special-focus or alternative schools tend to target students with educational disadvantages, students at those schools typically do more poorly in school and perform worse on tests than the general student population. Common examples include schools for at-risk youth, disabled students, dropouts, girls who are pregnant or have children, and juvenile delinquents. Some charter schools are also targeted at gifted or college-bound students, but these are much less common.

The targeting of charter schools to disadvantaged populations is so common that many people have come to believe, incorrectly, that all charter schools serve disadvantaged students. Even some of the charter school staff members we spoke to in our phone survey (see below) shared this belief. One, when asked if her school was targeted to a particular student population such as at-risk youth, replied, "well, that's

what a charter school is.”[2] Another told us, “that’s why charter schools were created in Texas, to serve at-risk kids.”[3]

One reason many charter schools serve targeted populations is that the procedures by which new charter schools are created often encourage such targeting. These procedures vary considerably from state to state. In many cases, local school boards are given veto power over the creation of new charter schools within their districts, and even where they do not have veto power they can have other forms of influence over the process. This tends to discourage the creation of charter schools that might effectively compete with regular public schools for students who are easier to educate, since school boards have no incentive to approve of the creation of such schools and every incentive to resist them. By the same token, the charter school approval process in most states tends to encourage the creation of charter schools targeted towards those students that regular public schools find less desirable because they are more likely to cause trouble, require extra academic help, or in some other way be difficult and expensive to educate.

Of course it’s to be expected that charter schools, like many school reform efforts, would emphasize helping disadvantaged students who need it the most. However, the existence of so many charter schools that serve targeted student populations is a serious obstacle to researching charter schools’ academic performance. If charter schools have smaller year-to-year improvements in test scores than regular public schools, this does not necessarily tell us anything about charter schools’ performance. It may be that charter schools do a better job of teaching, but their performance on tests is less impressive because their students are more likely to face unusually severe obstacles to learning over which the school has no control. Without more information about the student populations, we can’t know.

The lack of a valid comparison between charter school students and regular public school students is almost never adequately addressed in research on charter schools. The problem is often acknowledged, but few studies attempt to overcome it, and those that do make the attempt seldom have sufficient information to do a good job of it. Most of the existing research on charter schools is therefore of little real value in evaluating the performance of charter schools.

This point has begun to get some attention in the media. When charter school opponents in Illinois pointed out that charter schools were doing no better than regular public schools on state tests, Newsweek’s Jonathan Alter wrote that “the comparisons are meaningless” because “many of those charter schools are for troubled kids.” Alter called this kind of false comparison a “smear” perpetrated by “mindless boards of education and reactionary teachers’ unions” afraid of competition from charter schools (Alter 2002).

But commentators like Alter still make up only a small minority of the public discussion of charter schools. Far more typical is Francis X. Clines, who pointed out in a recent New York Times op-ed article that nearly two-thirds of Texas schools rated as low-performing under initial implementation of the federal No Child Left Behind law are “laissez-faire charter schools” (Clines 2003). Clines makes no mention of the enormous

number of Texas charter schools that are targeted at at-risk youth and other educationally disadvantaged populations. His readers are left with the clear implication that charter schools are failing to provide an adequate education.

The root of the problem for performing good research on the academic effect of charter schools is that states do not make available sufficient information on the recruitment and admissions practices of charter schools, nor on the demographic profile of students that charter schools ultimately attract. Information on recruitment and admissions practices is supposed to be included in each school's charter, of course, but it isn't feasible for researchers to separately obtain and analyze the charter of every charter school in a state, much less in a broad sample of states. Researchers must therefore rely on information collected and made available by state education agencies. In some states, no information on charter school population targeting is made available; other states report some information, but not enough. For example, Arizona reports which schools choose to identify themselves on state paperwork as being targeted to specific populations, but targeted schools are not required to identify themselves in this manner, and even among those that do, many provide vague or incomplete information. As for demographic information, the only data typically available on both charter and regular public schools is student race and ethnicity. Information on free and reduced lunch status is usually available for regular public schools, but it is not available for most charter schools, and in any event the free and reduced lunch program is far too broadly inclusive to be useful in measuring truly at-risk student populations. With very different student populations being served in charter and regular public schools, and with virtually no demographic data that could be used to control for those differences statistically, it is extremely difficult to conduct an appropriate evaluation of charter schools' academic quality.

This problem has left researchers struggling for other ways to study charter schools. Of course, one way to solve the problem is simply to study something other than test scores, something for which student background differences are less likely to be relevant. For example, Vanourek, Manno, Finn, and Bierlein studied satisfaction levels by surveying students, parents, and teachers in charter schools on such matters as what they liked most and least about their school, how their school compares to public schools they have previously been associated with, satisfaction with academic performance, and the reasons they choose charter schools. The survey found high satisfaction with charter schools, which is hardly surprising since those who are dissatisfied with charter schools are likely to leave them, and thus not be included in a survey. However, the survey is useful in that it documents other reasons besides academics that people seek out charter schools, such as better communication and responsiveness from school staff (see Vanourek, et. al. 1998).

Other researchers have studied other kinds of outcomes for charter schools. Fuller and his co-authors find that charter schools have less money, fewer credentialed teachers, and fewer special-education students than regular public schools; that brand new charter schools and charter schools managed by private firms have fewer credentialed teachers; and that charter schools with more black students are less well-funded (see Fuller, et. al.

2003). A U.S. Department of Education phone survey of school district leaders in districts where charter schools are located found that nearly half report a budget impact from student transfers to charter schools, nearly half report that their districts have become more customer-service oriented, and most report that their districts have made educational changes in response to charter schools. In states where school districts are the only charter authorizers, or where districts are paid money when they lose students to charter schools, districts are less likely to report budget impacts or changes to education and customer service in response to charter schools. All districts where enrollment was declining prior to the opening of charter schools reported a budget impact, including layoffs, and all reported an increased focus on customer service as well as educational changes (see Ericson, et. al. 2001).

A number of researchers have approached charter schools through case studies. Wells and her co-authors performed case studies of 17 charter schools in California; they argue that charter schools are “not held accountable” for performance by their authorizers, are not very innovative in their practices, and are not more cost-efficient than regular public schools. They also argue that the perception by regular public schools that charter schools have unfair advantages “inhibits competition” (see Wells, et. al. 1998). However, two U.S. Department of Education case studies that look at much broader samples of charter schools find more encouraging results. One study included field visits to 91 schools and concluded that charter schools are in fact held accountable for their performance, but not always by the same methods. The study also found that resource limitations are the biggest obstacle facing charter schools (see Nelson 2000). The other study included 150 schools and 60 authorizing agencies, and found that charters learn quickly that the best way to satisfy their various constituents (parents, donors, community leaders, district officials, etc.) is to focus on quality instruction. It also found that new types of charter authorizers (state boards of education, universities, etc.) learn more quickly than do local school districts to break habits of accountability based on process compliance rather than on performance and outcomes (see Hill, et. al. 2001).

In the end, though, there really is no substitute for knowing whether charter schools can make the grade academically. Some studies have compared charter school test scores to regular public school test scores without accounting for population differences at all. A recent study of California charter schools by the Charter Schools Development Center found that while charter schools overall had test scores below the average for regular public schools, charter schools more than five years old (a total of 80 schools) had scores above the average for regular public schools (see Agostini 2003). However, in the absence of information on student populations, it is not clear how much this descriptive statistic really tells us.

Horn and Miron compare Michigan charter schools to their host school districts. Comparing individual schools to whole school districts is a very imperfect comparison, given the enormous demographic diversity that is possible within a school district. This problem is somewhat mitigated because Horn and Miron look at year-to-year changes in test scores rather than the level of test scores in any one isolated year; year-to-year changes in test scores are less influenced by demographic factors. But the most severely

at-risk students may still show worse year-to-year changes in performance than other students, so uncontrolled demographic differences are still a problem for Horn and Miron's study.

Horn and Miron measure the difference between charter schools and host districts in two ways: the average score for all charter schools compared to the average score for all host districts, and the number of charter schools that outperform their host districts compared to the number of host districts that outperform their charter schools. The second method gives us a limited amount of insight into the problem of targeted charter schools: if charter schools are divided between well-performing schools serving general student populations and low-performing schools for at-risk populations, the average test scores might conceal this pattern, but looking at the number of schools that outperform their districts and vice versa might reveal it. Horn and Miron find that districts outperform charters in the aggregate, but their data tables reveal that in most grade levels and test subjects, more charter schools outperform their host districts than vice versa (see Horn and Miron 2000). Clearly, a minority of charter schools that may well be serving disadvantaged student populations dragged down the average charter school performance. Thus, Horn and Miron's study is seriously compromised by the incomparability of targeted charter schools and regular public schools.

In studying California charter schools, Raymond also compares them to regular public schools in their host districts (as well as to the overall performance of regular public schools across the state). Raymond is able to cope with differences in student background better than Horn and Miron, because her study collects school-level data on poverty, race, parents' education, limited-English status, and testing inclusion. Using these data to provide demographic controls, Raymond finds that charter schools have test score improvements that are generally better than those of regular public schools, though the differences are only sometimes statistically significant (see Raymond 2003).

One way for researchers to overcome the problem of incomparability between charter school populations and regular public school populations is to obtain specific demographic information on every individual student in the study. Obtaining a sufficient amount of such information is rarely feasible, but when it can be done researchers can draw valid comparisons between charter schools and regular public schools. Using individual-level data on students in Arizona, Solmon, Paark, and Garcia were able to control for race, primary language, gifted or disability status, absenteeism, and prior year test scores. Although this does not include information on income and family status, and thus does not fully tell us whether students are "at risk" or not, the controls for race, language, disability, and absenteeism should do a decent job of allowing meaningful population comparisons. They found that students who were in their first year in a charter school had significantly lower reading scores than comparable students in regular public schools, but students who were in their second and third consecutive years in a charter school had significantly higher reading scores than comparable regular public school students. For math scores they found that charter school students had lower scores in their first year and higher scores in their second year, but they found no statistically significant difference in their third year in a charter school (Solmon, Paark, and Garcia 2001).

Another approach, useful when multiple years of individual student test score data are available but individual demographic data are not, is to use what statisticians call a “fixed-effects” model. This method keeps track of each student’s performance separately, comparing a student’s performance in one year to his performance in the next year, the year after that, and so on. By comparing each student only with himself, this method removes much of the influence of differences in student populations. However, the tradeoff for this advantage is that more years of data are needed to make the comparisons meaningful. This is a particularly serious problem when studying charter schools, which in most cases have not existed for more than a few years. In addition, disadvantaged students may well exhibit slower year-to-year progress in addition to starting with lower test scores.

Gronberg and Jansen use this method to study charter school students in Texas over three years. They conclude that students in charter schools officially designated as serving at-risk students saw larger test score gains than students in regular public schools. (During their study period, Texas had two categories of charter schools, one category of schools that admitted exclusively at-risk students and another category that contained schools with various admission policies; since then, all Texas charter schools have been converted to open-enrollment policies.) They found that the reverse was true for students in other charter schools, although this finding is somewhat suspect because there were fewer schools of this type. More importantly, both these conclusions are tainted by the availability of only three years of data. Using other methods that do not address the problem of student population differences, they find that students tend to score worse after their first year in a charter school, but their performance compared to students in regular public schools recovers in subsequent years, and that charter schools are more cost-efficient than regular public schools (see Gronberg and Jansen 2001).

It should be obvious that there is very little we can say with confidence about the academic performance of charter schools based on this research. Even those few studies that make some contribution to our understanding are typically confined to only one state. A study that includes only one state is always of limited application to other states, and in the case of charter schools it may well have no application at all, given the large differences in the design and functioning of charter school systems in different states. We have been unable to find any previous multi-state study of charter school test scores that addresses the problem of student population differences between charter schools and regular public schools.

Method

We included in our study eleven states—Arizona, California, Florida, Texas, Michigan, Wisconsin, Ohio, Colorado, North Carolina, Minnesota, and Pennsylvania—that had large enough numbers of charter schools to give us a good chance of yielding sufficiently large samples after targeted schools were eliminated. We also included four other states—Illinois, New York, New Jersey, and Indiana—that had fewer charter schools but were of particular interest. Four of the states we examined—Illinois, New

York, Indiana, and Wisconsin—were ultimately excluded from our analysis for lack of sufficient data.[4] This left us with eleven states included in our analysis.

We obtained a complete list of charter schools from each state, along with any available information on targeting of student populations in these schools. We then eliminated from these lists any schools that were known to be targeted to particular student populations, along with any schools opened too recently to have test score data available for our study. (We could only examine schools that had two years' worth of test scores.) This gave us the list for our telephone survey. Calling every school on the list, we collected some basic information about each school (grades served, number of students, etc.) and then asked whether the school was targeted to a particular student population such as at-risk youth, dropouts, gifted students, or special education students.

We considered a school targeted to an educationally advantaged or disadvantaged population if it gave that population preferential treatment in admissions or if it made specific efforts to recruit that population, such as describing itself as a school for that type of student. We did not count a school as targeted if it had a distinctive curriculum or pedagogy, such as the Montessori system or computer-based instruction, so long as it made no attempt to recruit or attract educationally advantaged or disadvantaged populations in particular. We also did not count a school as targeted based solely on the local population in its location. For example, many charter schools told us that they had large numbers of at-risk students; in such cases, we asked whether this was because the school targeted at-risk students or because the school just happened to get a lot of at-risk students as a result of its location. If the latter was the case, the school was not counted as targeted because it was serving the general student population in its area. Educationally disadvantaged student populations included groups such as at-risk students, pregnant or young-mother students, dropout recovery students, or disabled students, while educationally advantaged students included groups such as gifted or college-bound students.

We excluded some charter schools from our study because they function in a way that does not allow them to be studied appropriately by our method. For example, we excluded “cyber” charter schools that provide instruction over the Internet, because the lack of a central campus to which students must travel leaves us without a real physical location for the school. This makes it impossible to compare that school with the closest regular public school (see below). We also excluded “conversion charters”—regular public schools that converted into charter schools—if they retained their assigned neighborhoods after the conversion. Students in those neighborhoods are automatically assigned to the converted charter school unless they specifically request to be transferred to a non-charter school. We found 235 such schools in California, and a handful in several other states. We excluded these schools because they are not operating as schools of choice like other charter schools.

Once we had completed our phone survey, we had a list of charter schools in each state that we had confirmed were serving general student populations. We then obtained from each state a complete list of regular public schools. Using a street atlas computer

program, we mapped out the location of every regular public school and every charter school in each state. For each test score we had from a given charter school, we found the closest regular public school for which we had test scores from the same grade. We ignored regular public schools known to be targeted to particular populations, such as magnet schools and schools for juvenile delinquents, as well as regular public schools for which we had no test scores comparable to the test scores we had for our charter schools.

Comparing charter school test scores to the test scores of their closest regular public school produces a much more fair comparison than, for example, comparing charter schools to their school districts. A given school district will normally contain a wide variety of student populations. In many cases, the same school district might contain both a depressed urban area and a booming suburban area. If it is the case, as we have good reason to believe, that charter schools are more likely to be located in the poorest parts of their districts, then comparing each charter school to its district will not be a fair comparison, as the charter schools are serving a much more disadvantaged population. However, the closest regular public school to a given charter school can reasonably be expected to serve a similar population. To be sure, their student populations will not be exactly the same, but they will be much more similar to one another than any other available comparison. In this way, we used geography to control for demography.

Rather than look at each school's test score levels in a single year, we looked at year-to-year test score changes. Specifically, we looked at the change in each test score for the two most recent years. For example, if the most recent available year for 3rd grade math scores was 2002, we looked at the change in 3rd grade math scores between 2001 and 2002. We used each school's average scale score or percentile rank, except in Michigan and Ohio where only the percentage of students at each school who passed the test was available. In Michigan, test scores from the first and second years included in our study were not directly comparable because of changes in the testing system, so we standardized the test scores to make them comparable.

Looking at year-to-year score changes rather than single-year score levels allows us to further filter out some of the influence of student and family background factors, focusing instead on the contribution each school makes to learning. It is well established that students from more advantaged backgrounds, who have had more opportunities to learn and face fewer obstacles to learning, score better on standardized tests. However, the year-to-year improvement (or lack thereof) that each student makes is more directly influenced by the quality of the school. This is sometimes called "value added" analysis, since it is a more direct measurement of the educational value added to a student by his school rather than the baseline from which the student started. Since the goal of our study is to overcome the problem of differences in student populations, it is particularly important that we remove the influence of student background characteristics to the extent that we can. Excluding targeted charter schools that are unlikely to have a counterpart in the regular public school system is by far the most important way in which we do this; comparing each charter school to its closest regular public school is another. Using year-to-year score changes instead of single-year score levels further refines our research design.

In each state, we ran a regression analysis on year-to-year score changes for each test subject (math, reading, language, or science). This analysis controlled for whether the school was a charter school, in order to determine what effect charter schools have on test scores as compared to their closest regular public schools. We also included a control variable for race (specifically, the percentage of the student body that is white, according to data from the National Center for Education Statistics) and a control variable for each tested grade level.[5]

We report results in each of the states where we found enough untargeted charter schools to be confident that our results were meaningful. The results reported for each state have been standardized to allow for accurate state-to-state comparisons (Table containing these results is located in the Appendix). In the table, the first column gives the standardized coefficient representing the effect that untargeted charter schools had on test scores. A positive number means charter schools had greater year-to-year test score improvement than their closest regular public schools, while a negative number means charter schools had smaller year-to-year test score improvement. Results that are statistically significant at a high level ($p < 0.05$), meaning we can have very high confidence they reflect the real effect of untargeted charter schools, are marked with double asterisks, while results that are statistically significant according to a lower standard ($p < 0.1$) are marked with single asterisks. Results not marked with asterisks are not statistically significant.

The figures in the first column represent the change in test score as measured in standard deviations. For example, untargeted charter schools in Texas made 0.18 standard deviations greater progress on TAAS math scores than neighboring public schools. The figures in the second column translate these numbers into the percentile point change this would represent for a student who started at the 50th percentile—that is, exactly in the middle. Returning to our example, an improvement in TAAS math scores of 0.18 standard deviations would be equivalent to an improvement of 7 percentile points for a student who started at the 50th percentile. In other words, it would be the equivalent of a student who originally scored at the 50th percentile improving his test score to the point where he scored at the 57th percentile. Students starting either above or below the 50th percentile would see a correspondingly smaller change in percentile point terms from the same standard deviation change in test scores.

We also combined the data from all eleven states and performed a national analysis. In addition to controlling for race and grade level, the national analysis controlled for each state included in the study. This analysis gives us a picture of the effect that untargeted charter schools have on test scores across the country.

Results

Populations Served by Charter Schools

One thing we wanted from our study was simply to get a rough picture of how frequently charter schools were targeted to educationally advantaged or disadvantaged student populations. We do not have a perfectly representative picture here, since we do not have data on schools that did not respond to our telephone survey, or that were not called because they were too new to have test scores available. In particular, some states provided a lot of up-front information on which schools were targeted, while others provided little information, and still others no information at all. This means that our telephone survey method was not precisely the same in every state; in some states we began by excluding large numbers of schools from our survey, in others fewer schools, and in some states we had to call each and every charter school in the state. However, by looking at those schools for which we do have data, we can get at least a rough picture of what populations are being served by charter schools.

Results for targeting of charter schools to particular populations are presented in Table 1. We found very large variations from state to state in the targeting of charter schools. In Michigan, 95% of charter schools for which we have data were untargeted, whereas in Wisconsin 12.0% of charter schools for which we have data were untargeted. Other states ran the gamut between these extremes. There was also variation in the presence of charter schools targeted at advantaged populations; in Arizona 9.2% of schools for which we have data were “targeted high,” while in four states we found no such schools whatsoever. In California we found a very large number of so-called “conversion” schools that had to be excluded from our study, and in Pennsylvania we found a significant number of “cyber” schools. Those excluded schools are represented in the “other” column.

National Charter School Effect

The results of our national analysis were drawn from the combined data for all eleven states. They represent the effect that all the untargeted charter schools in our study had on test scores when compared to the performance of their closest regular public schools.

These results showed a positive effect from charter schools and were statistically significant, but the size of the effect was modest. Untargeted charter schools made math test score improvements that were 0.08 standard deviations greater than those of neighboring public schools during a one year period. For a student starting at the 50th percentile, this would amount to a gain of 3 percentile points, to the 53rd percentile. Reading test score results showed 0.04 standard deviations greater improvement in untargeted charter schools than in their closest regular public schools over the course of a year, a benefit that would raise a 50th-percentile student 2 percentile points to the 52nd percentile.

Because these results are statistically significant, we can be very confident that the charter schools in our study did have a positive effect on test scores. However, the small size of this effect should caution us against too much enthusiasm regarding the benefits of charter schools. The results of our study strongly support the conclusion that

untargeted charter schools are somewhat better than regular public schools serving similar populations, but not a great deal better.

Arizona Charter School Effect

Arizona gave two statewide standardized tests. The Stanford Achievement Test, Ninth Edition (SAT-9), a nationally normed test used in many states, was given in grades 2-9 (the reading test was given in grades 1-9). Arizona's Instrument to Measure Standards (AIMS), a state test geared to Arizona's school curriculum, was given in grades 3, 5, 8, and 10-12.

Our Arizona results found weak and mixed effects from charter schools, and none that was statistically significant. We found small positive effects on the SAT-9 and small negative effects on the AIMS, but because none of these results was statistically significant we cannot be confident that either of these effects was really caused by charter schools. On the SAT-9 we found that charter schools had made progress of 0.03 standard deviations higher in math scores and 0.02 standard deviations higher in reading scores than that of regular public schools in one year's time, each of which is equivalent to 1 percentile point of improvement for a student starting at the 50th percentile. They also had made progress in language scores over the course of one year that was 0.01 standard deviations greater than that of regular public schools, equivalent to less than 1 percentile point of improvement. On the AIMS we found that charter schools had made less improvement in math scores by 0.06 standard deviations, equivalent to a drop of 2 percentile points for a student starting at the 50th percentile, and less progress in reading scores by 0.03 standard deviations, equivalent to a drop of 1 percentile point from the 50th percentile.

Given that the direction of the untargeted charter school effect (positive or negative) was mixed for different tests, the magnitude was always small, and none of the results was statistically significant, it is difficult to reach any conclusions about Arizona charter schools. We cannot be confident that they have any effect at all on test scores.

California Charter School Effect

California gave the SAT-9 in grades 2-11. Our California results found effects that were somewhat larger than those in Arizona, and in this case all the effects were positive. However, our results still were not statistically significant. We found that untargeted charter schools had math test score gains over a one-year period higher than those of nearby regular public schools by 0.06 standard deviations, and reading test score gains higher by 0.05 standard deviations, each of which is equivalent to a 2 percentile point gain for a student starting at the 50th percentile. Untargeted charter schools also had language test score gains 0.07 standard deviations higher than those of regular public schools, equivalent to a 3 percentile point gain from the 50th percentile. None of these results was statistically significant, so although these results are more positive than our Arizona results they still do not justify strong conclusions about the effect of untargeted charter schools in California.

Florida Charter School Effect

Florida gives two statewide standardized tests in grades 3-10. They are the SAT-9 and Florida's Comprehensive Assessment Test (FCAT), a state test geared to Florida's school curriculum.

Our Florida results yielded stronger positive effects from untargeted charter schools than our Arizona and California results, and two of these findings were statistically significant. We found that charter schools achieved year-to-year SAT-9 math score gains and FCAT reading score gains that were each 0.15 standard deviations greater than those of nearby regular public schools, equivalent to a gain of 6 percentile points for a student starting at the 50th percentile. Both these results reached a high level of statistical significance, so we can be very confident that the charter school effects were positive. We also found that charter schools had greater annual improvement on SAT-9 reading scores and FCAT math scores by 0.08 and 0.12 standard deviations, respectively, equal to gains of 3 and 5 percentile points from the 50th percentile. However, these two findings were not statistically significant.

These findings allow us to reach a firm conclusion regarding Florida charter schools. We can be very confident that charter schools that serve the general population of students experienced greater gains in SAT-9 math and FCAT reading scores than did neighboring regular public schools. While still moderate, these gains were larger than those in Arizona or California.

North Carolina Charter School Effect

North Carolina gives a state-designed "end of grade" exam in grades 3-8. Our North Carolina results were the weakest of the five states for which we report separate results. We found that untargeted charter schools had year-to-year math score gains that were 0.02 standard deviations higher than those of regular public schools, equal to a 1 percentile point gain for a student starting at the 50th percentile, and reading score gains that were 0.01 standard deviations higher, equal to less than a 1 percentile point gain. Neither of these findings was statistically significant, so we cannot be confident that untargeted charter schools in North Carolina had any effect on test scores.

Texas Charter School Effect

Texas gave the Texas Assessment of Academic Skills in grades 3-10. Our Texas results were the strongest in our study. We found that untargeted charter schools in Texas achieved math score improvements over a one year period that were 0.18 percentile points greater than those of neighboring public schools, equivalent to a gain of 7 percentile points for a student starting at the 50th percentile. We found that Texas charter schools produced a reading score benefit of 0.19 percentile points, equivalent to a gain of 8 percentile points from the 50th percentile. Both these findings were statistically significant at a high level. These findings suggest that Texas charter schools that are not

targeted toward educationally advantaged or disadvantaged populations are able to make significantly greater annual progress on standardized test scores than their regular public school counterparts.

Conclusion

We found positive effects from charter schools serving general populations, but for the most part these effects were modest in size. This suggests that while charter schools benefit from their partial independence from the regular school system, they still face significant obstacles to more effective reform.

There are several possible explanations for the finding that charter schools outperformed their regular public school counterparts. One possibility is that charter schools benefit from the freedom they enjoy from many state regulations. With less of a regulatory burden, charter schools may be able to focus more of their energies on assisting students and enjoy greater flexibility in meeting student needs. The better academic performance of charter schools could also be a result of their being schools of choice. By allowing families to send their children to schools of their own choosing, charter schools may permit more efficient matching of student needs with school capabilities. Given that not every child learns in the same way, charter schools help families find a school that works well for them and their children. This “choice factor” may also produce better academic performance in charter schools because of the incentives that charter schools face when attempting to attract and retain students. Unlike public schools, to which students are assigned based on residence, charter schools have to earn their students. This makes it more likely that they will better serve those students’ needs.

One possible explanation for the modest size of positive charter school effects is that charter schools typically receive less money than regular public schools. The per-pupil allocation charter schools receive for operating expenses is often somewhat less than regular public schools receive, and charter schools almost always receive significantly less money than regular public schools—or no money at all—to cover capital costs such as construction and maintenance. However, it is unlikely that this is the most important explanation, because our results do not show a clear relationship between funding of charter schools in each state and charter schools’ effects on test scores. For example, while the state with our weakest test score results, North Carolina, provides no capital funding at all for charter schools, the same is true of Texas, the state with our strongest results.[6]

Another possible explanation is that charter schools, though exempt from many regulations, are still public schools subject to significant regulatory burdens. These will vary considerably from state to state, of course. In many states, charter schools are not automatically exempted from all relevant state regulations, but instead must apply for the specific exemptions they want. Different states give charter schools different degrees of latitude in curriculum, hiring, firing, organizational structure, and budgeting. A recent article in the *Los Angeles Times* detailed the significant regulatory burdens on charter

schools in California (see Mathews 2003). Also, depending on its structure, the charter granting process may discourage bolder innovations in favor of schools that will look more or less like the status quo. Furthermore, charter schools are subject to the oversight powers of their charter authorizing agencies. In places where these powers are generously defined and aggressively wielded, the real independence of charter schools may be limited. If the regulatory burden imposed on charter schools is not all that different in practice from the burden on regular public schools, it would not be surprising to find only modest differences in how these different types of schools perform academically.

One other possible factor is the requirement in many states that charter school students take standardized tests specifically geared to state curricula. Part of the possible benefit that charter schools might provide is through curriculum innovation; if schools are judged by student performance on a test designed for the state's curriculum, they will not be as free to experiment with curriculum changes. However, this point would not apply to nationally norm-referenced tests such as the SAT-9, which are broader measures of student performance.

Finally, the most important factor limiting the performance of charter schools may well be their newness. Charter schools have not been around for a very long time, and in most states they are overwhelmingly brand new schools rather than converted schools. It takes a while for a new school to get its house in order—to figure out what works and what doesn't. It may be that as charter schools get older, their performance compared to regular public schools may become more impressive.

While this study's research design does a better job than previous designs at screening out the effects of charter school targeting and drawing fair comparisons between charter and regular public schools, it is still possible to do better. A random assignment study of charter and regular school students would go even further toward establishing a fair comparison. Random assignment, which compares the performance of two randomly-selected groups that are the same in almost every way except for the factor being studied, is the "gold standard" for social science research. A study comparing students who did and did not get admitted to charter schools through a random lottery would provide the best possible comparison between charter and regular public schools. Every effort should be made to design and implement a random assignment study of charter school performance to increase the confidence with which we can draw conclusions about their academic effectiveness.

Endnotes

1. Data obtained from the U.S. Charter Schools website, sponsored by the U.S. Department of Education (www.uscharterschools.org).
2. Phone interview with a charter school staff member (January 16, 2003, 1:35 pm).
3. Phone interview with a charter school staff member (January 28, 2003, 10:43 am).
4. Indiana was excluded because all 13 of its charter schools are too new to have test score data available. New York was excluded because the state was only able to provide test score data in the two most recent years for five schools, too few to produce meaningful results. We had already conducted our phone survey in New York when we discovered this, so we do have information on the targeting of New York charter

schools. This information is presented in Table 1. After excluding targeted schools, Illinois and Wisconsin lacked a sufficient number of remaining charter schools to warrant further test score analysis.

5. Data in our regression analysis were weighted based on each school's enrollment.

6. North Carolina funding information obtained via interview with Otho Tucker of the state's education department (April 15, 2003, 3:10 pm); Texas funding information obtained from the state's website (<http://www.tea.state.tx.us/charter/faqfunding.htm>).

References

- Michael Agostini, "Veteran Charter Schools Outperform Non-charters on API," Charter Schools Development Center, 2003. http://www.cacharterschools.org/pdf_files/veteran_API.pdf
- Jonathan Alter, "Attack of 'The Blob'," Newsweek Web, November 27, 2002.
- Francis X. Clines, "Re-educating the Voters About Texas' Schools," New York Times, June 3, 2003.
- John Ericson and Debra Silverman, "Challenge and Opportunity: The Impact of Charter Schools on School Districts," U.S. Department of Education, June 2001. http://www.ed.gov/pubs/chartimpact/district_impact.doc
- Bruce Fuller, Marytza Gawlik, Emlei Kuboyama Gonzales, and Sandra Park, "Charter Schools and Inequality," Policy Analysis for California Education, April 2003. <http://pace.berkeley.edu/Chartersummary.pdf>
- Timothy J. Gronberg and Dennis W. Jansen, "Navigating Newly Chartered Waters," Texas Public Policy Foundation, April 2001. http://www.tppf.org/education/nncw/tppf_nncw.pdf
- Paul Hill, et. al., "A Study of Charter School Accountability," U.S. Department of Education, June 2001. <http://www.ed.gov/pubs/chartacct/>
- Jerry Horn and Gary Miron, "An Evaluation of the Michigan Charter School Initiative," The Evaluation Center, Western Michigan University, July 2000. <http://www.wmich.edu/evalctr/charter/michigan/>
- Jay Mathews, "Charter Schools Choke on Rulebook," Los Angeles Times, June 1, 2003.
- Beryl Nelson, et. al., "The State of Charter Schools 2000," U.S. Department of Education, January 2000. <http://www.ed.gov/PDFDocs/4yrrpt.pdf>
- Margaret Raymond, "The Performance of California Charter Schools," Hoover Institution, May 2003. <http://credo.stanford.edu/Performance%20of%20California%20Charter%20School.FINAL.complete.pdf>
- Lewis Solomon, Kern Paark, and David Garcia, "Does Charter School Attendance Improve Test Scores?" Goldwater Institute, March 2001. <http://www.goldwaterinstitute.org/pdf/materials/111.pdf>
- Gregg Vanourek, Bruno V. Manno, Chester E. Finn, Jr., and Louann A. Bierlein, "Charter Schools as Seen by Students, Teachers, and Parents," in Learning from School Choice, Paul E. Peterson and Bryan C. Hassel, eds., Brookings Institution, 1998.
- Amy Stuart Wells, et. al., "Beyond the Rhetoric of Charter School Reform," UCLA Charter School Study, 1998. <http://www.gseis.ucla.edu/docs/charter.PDF>

Tables:

Table 1: Populations Served by Charter Schools

	Untargeted	Targeted Low	Targeted High	Other
Arizona	47.5%	42.3%	9.2%	0.9%
California	18.6%	3.6%	1.0%	76.8%
Colorado	67.3%	25.0%	5.8%	1.9%
Florida	57.6%	37.4%	0.7%	4.3%
Illinois	25.0%	66.7%	8.3%	0.0%
Michigan	95.0%	2.8%	0.7%	1.4%
Minnesota	38.8%	57.1%	4.1%	0.0%
New Jersey	74.2%	19.4%	6.5%	0.0%
New York	78.3%	13.0%	0.0%	8.7%
North Carolina	84.9%	12.3%	2.7%	0.0%
Ohio	65.2%	33.3%	0.0%	1.4%
Pennsylvania	69.1%	18.2%	0.0%	12.7%
Texas	58.6%	41.4%	0.0%	0.0%
Wisconsin	12.0%	78.7%	4.0%	5.3%

Table 2: Charter School Effects

	Charter Effect	Change from 50th Percentile	Number of Cases
National			
Math	+0.08**	+3 percentile points	2,126
Reading	+0.04**	+2 percentile points	2,388
Arizona			
SAT-9 Math	+0.03	+1 percentile point	641
SAT-9 Reading	+0.02	+1 percentile point	643
SAT-9 Language	+0.01	Less than 1 percentile point	637
AIMS Math	-0.06	-2 percentile points	316
AIMS Reading	-0.03	-1 percentile point	305
California			
SAT-9 Math	+0.06	+2 percentile points	175
SAT-9 Reading	+0.05	+2 percentile points	174
SAT-9 Language	+0.07	+3 percentile points	171
Florida			
SAT-9 Math	+0.15**	+6 percentile points	293
SAT-9 Reading	+0.08	+3 percentile points	293
FCAT Math	+0.12	+5 percentile points	286
FCAT Reading	+0.15**	+6 percentile points	286
North Carolina			
End of Grade Math	+0.02	+1 percentile point	368
End of Grade Reading	+0.01	Less than 1 percentile point	364
Texas			
TAAS Math	+0.18**	+7 percentile points	319
TAAS Reading	+0.19**	+8 percentile points	319

** = statistically significant at $p < .05$; * = statistically significant at $p < .10$