Policy Analysis

No. 677 June 6, 2011

The Other Lottery

Are Philanthropists Backing the Best Charter Schools?

by Andrew J. Coulson

Executive Summary

The central problem confronting education systems around the world is not that we lack models of excellence; it is our inability to routinely replicate those models. In other fields, we take for granted an endless cycle of innovation and productivity growth that continually makes products and services better, more affordable, or both. That cycle has not manifested itself in education. Brilliant teachers and high-performing schools can be found in every state and nation, but, like floating candles, they flicker in isolation, failing to touch off a larger blaze.

Over the past decade, one of the most prominent strategies for overcoming this problem has been for philanthropists to partner with the best charter schools in an attempt to bring them to scale. The present study seeks an empirical answer to the question: Is that strategy working—are the highest-performing charters attracting the most funding?

We search for the answer in California, the state with the largest number of charter schools and the largest number of charter school networks (groups of two or more schools following the same pedagogical model or founded, overseen, or operated by the same person or group). Specifically, we compare the amount of philanthropic funding received by these networks with their performance on state-administered and Advanced Placement tests. The results are discouraging. There is effectively no correlation between grant funding and charter network performance, after controlling for individual student characteristics and peer effects, and addressing the problem of selection bias.

For example, the three highest-performing charter school networks perform dramatically above the level of conventional public schools on the California Standards Tests, but rank 21st, 27th, and 39th in terms of the grant funding they have received, out of 68 charter networks. The AP results are worse; the correlations between charter networks' AP performance and their grant funding are negative, though negligible in magnitude.

The top-performing charter networks, like top-ranked American Indian charter schools, play a transformative role in children's educational and career prospects, and lay bare the failure of our conventional educational arrangements to fulfill each child's potential. We should indeed strive to preserve and replicate them. But philanthropy has not proven to be a reliable, systematic mechanism for accomplishing that goal in California, which has enjoyed more earnest and extensive efforts in this regard than perhaps any other state.

Andrew Coulson is director of the Cato Institute's Center for Educational Freedom and author of Market Education: The Unknown History *(1999).*

CATO INSTITUTE

Are philanthropists reliably targeting their investments at these high-performing networks? The present study aims to address that question.

Introduction

When Ambassador Walter Annenberg launched his $500 million education challenge in December 1993, President Bill Clinton made an incisive observation:

> [The] people in this room who have devoted their lives to education are constantly plagued by the fact that nearly every problem has been solved by somebody somewhere, and yet we can't seem to replicate it everywhere else. . . . That is the central challenge of this age in education.

Sadly, the consistent replication of educational excellence remains elusive nearly two decades later. Nevertheless, efforts continue to find a way of achieving that goal. One of the most high-profile strategies in recent years has been for major philanthropists and foundations to fund the expansion of charter school networks (groups of two or more charter schools founded, overseen, or operated by the same organization or individual). The most common goal of such charitable giving is to raise the academic achievement of low-income and minority students, who typically lag behind their peers across subjects.

Though research on the performance of charter schools as a whole has been mixed (see below), there is growing evidence that at least some networks are significantly more academically effective than traditional public schools. But are philanthropists reliably targeting their investments at these high-performing networks? The present study aims to address that question by analyzing the link between academic performance and philanthropic funding levels among California's charter school networks.

The Charter School Research to Date

In December 2008, Julian Betts and Emily Tang reviewed the charter school research and found modestly positive, though mixed, student achievement effects.[1] There were many statistically insignificant findings, quite a few small positive findings, but also several small negative findings.

Betts and Tang also reported that the median effect sizes attributable to attending charter schools were small, usually less than 0.1 of one standard deviation (see the sidebar for an explanation of standardized effect sizes). Even when looking only at the quarter of studies reporting the largest effect sizes, the largest median effect size across school levels and subjects was roughly 0.2 standard deviations.

Among the studies conducted since the Betts and Tang review was published, Hoxby, Murarka, and Kang's randomized lottery experiment of New York City charter schools is the most methodologically compelling.[2] Charter schools that are over-subscribed must generally accept students via random drawings. This makes it possible to compare the performance of students who were randomly accepted into the charter school with that of students who were randomly rejected and so attended traditional public schools. This is the kind of "randomized controlled trial" that is used in drug testing and is considered the gold standard of empirical research.[3]

What Hoxby and her colleagues found is that attending a charter school has an average annual effect of 0.12 standard deviations in mathematics and 0.09 standard deviations in English Language Arts (ELA). Added up over the entire K-8 grade range under investigation, this amounted to a large cumulative effect in mathematics and a moderately large effect in ELA.

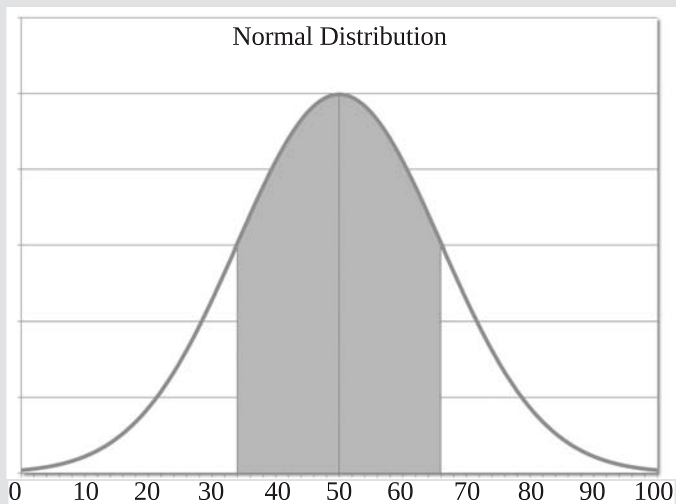
A subsequent analysis of NYC charter schools by Stanford's Center for Research on Education Outcomes (CREDO) also found positive effects. (It should be noted that the CREDO team did not use a randomized lottery model.) When the CREDO team performed a multi-state study covering 70 percent of all students in the charter sector,

Standardized Effect Sizes

Many different tests are used to measure academic outcomes, making direct comparisons difficult. SAT scores, for instance, are reported on an 800-point scale, scores on the National Assessment of Educational Progress are reported on a 500-point scale, and many state and district level tests are reported on a 100-point scale. Even when the scales are the same, test difficulty can vary so much that a 5-point increase on one test might represent a bigger gain in real achievement than a 10-point increase on another. To allow scores to be compared across tests, social scientists "standardize" them by subtracting a test's mean (average) score from each individual score, and then dividing the result by the standard deviation of the scores on that test. The result is an effect size measured in standard deviations.[1]

The standard deviation is an indication of how varied scores are. For instance, a test with very similar questions that are all at the same difficulty level will have a relatively low standard deviation, because most students will get all the questions correct (or all incorrect), while a test that explores many different areas of a subject at different levels of difficulty will likely have a larger standard deviation.

Test scores typically follow a normal distribution (see figure at right), in which the bulk of students are clustered around the mean (the middle of the bell curve) and many fewer have scores that are very high or very low (in the tails of the curve). More precisely, when test results are normally distributed, about two thirds of the students will score within one standard deviation of the mean (the shaded area in the figure).



For example, consider a challenging test on which the overall mean score is 50 out of 100, and the standard deviation is 16. If a study finds that charter school students score 54 on that test even after taking into account nonschool factors (such as their parents' level of education), then the standardized effect size of attending those charter schools would be equal to (54−50) / 16, or 0.25 standard deviations. By convention, effect sizes around 0.2 standard deviations are considered "small," those around 0.5 are considered "moderate," and those near or above 0.8 are considered "large."[2]

---

[1] This produces standardized test scores for each test that have a common mean of zero and a common standard deviation of 1. Since a normal distribution with a mean of zero and a standard deviation of 1 is also called a z-distribution, these standardized scores are often referred to as "z-scores." Once we have z-scores for each subject and test, we can take their mean (average) to obtain an overall score for each charter network. For an example of the use of z-scores to compare educational outcomes across subjects and grades, see Christina Clark Tuttle, Tara Anderson, and Steven Glazerman, "ABCTE Teachers in Florida and Their Effect on Student Performance Final Report," Mathematica Policy Research, Inc., September 4, 2009, http://www.mathematica-mpr.com/publications/pdfs/education/ABCTE_FL_Teachers.pdf.

[2] Jacob Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. (Hillsdale, NJ: Lawrence Erlbaum Associates, 1988), p. 26.

however, they concluded that charter schools slightly underperformed compared to traditional district schools.[4]

Caroline Hoxby has faulted this multi-state study for suffering from a methodological error that biased the effects of charter schools toward zero.[5] This precipitated a rebuttal from the CREDO researchers defending their findings,[6] though Hoxby remains unconvinced. Assuming, for the sake of argument, that the Hoxby critique is correct, one thing still seems clear: the underlying multi-state data analyzed by the CREDO team do not seem to show a large, consistent positive effect for charter schools. If they did, that effect likely would have overcome the bias described by Hoxby, though at a reduced level, because the CREDO team applied essentially the same methods to the NYC data, where, as already noted, they *did* find a positive overall impact.

To sum up, it is not entirely clear whether overall charter school performance is slightly better, slightly worse, or equivalent to that of traditional district schools, though the bulk of studies seems to suggest it is slightly better. Four of the five randomized assignment studies also indicate a positive charter effect, though they have so far only been conducted in a few parts of the country.[7]

What we can say with greater confidence is that some charters perform significantly better and others perform significantly worse than traditional district public schools. From a policy standpoint, this begs the question: Is there a mechanism for routinely scaling up the top performers so that they command a progressively larger share of the marketplace, crowding out the lower performers?

Is there a mechanism for routinely scaling up the top performers so that they command a progressively larger share of the marketplace, crowding out the lower performers?

There has been very little empirical research on this question to date. Hanushek et al. found that parents in Texas were somewhat more likely to pull their children out of poor-performing than high-performing charters, but could not say with certainty what long-term effect this might have since they did not have data on the propensity of families to choose high-performing over low-performing charters when selecting a *new* school.[8] It is possible, in other words, that parents might leave one bad charter school for another, which would not lead to improved overall quality of the sector.

Ongoing efforts by philanthropists to spur the growth of high-performing charter schools are thus of keen interest. The remainder of this paper offers suggestive evidence on their degree of success using data from California, the state that has both the largest total number of charter schools and the largest number of charter school networks.

The Data Used in this Study

This study examines California's charter school "networks," which we define as any group of two or more charters that share a common management organization, common founder, or common pedagogical model.[9]

Using electronic databases of charitable giving,[10] we compiled records of philanthropic grants made to these charter networks, or to any of the California-based schools belonging to them, within the past eight years.[11] These results were then supplemented with Internet searches for any grants that were too recent to be included in the electronic databases or that might have been missed by their data collection procedures. Finally, we searched California's Fair Political Practices grant database for donations made to charter schools.[12] While it is possible that errors or omissions exist in these grant data, we have no reason to suspect that any such "noise" could be of sufficient magnitude to affect the overall results of this study.

Next, we obtained academic achievement data from the California Standardized Testing and Reporting (STAR) 2010 research files, the only source covering all public (including charter) schools in the state. From these we extracted the average scores on the California Standards Tests, broken down by student socioeconomic status (SES), subject, and grade.[13] The eight SES categories covered in this study include both "low-income" and "not low-income" Asian, black, Hispanic, and white students.

Though earlier-year STAR data are also available, California does not track the academic growth of individual students, precluding a study of student gains over time.

Because achievement is known to be affected by the SES of students' peers, we also obtained a measure of schoolwide family income (the percent of the school's student body eligible for free or reduced price meals under the federal government's lunch program).

In addition to the California Standards Tests, we also report results on Advanced Placement tests as a measure of high-end performance at the high school level. Though detailed SES breakdowns are not available for the AP tests, we do have the total number of passing scores earned by black and Hispanic students, which at least helps to control for the confounding effects of minority status. By dividing that number by black and Hispanic enrollment, we obtain a minority student AP pass rate for each charter network with at least one high school. This measure is of particular interest given the explicit goal of many philanthropists to boost black and Hispanic student achievement.

Caution is required in using AP results, however. Several well-known national high school ranking systems now weigh the total number of AP tests taken or passed, and this has created an incentive for schools to boost those statistics. One increasingly popular means of doing so has been to encourage native Spanish speakers to enroll in the AP class in Spanish as a foreign language. University of Texas economist Kristin Klopfenstein notes that Spanish-speaking students are being told "go take AP Spanish language and get easy AP credits, because it looks good." The same phenomenon has been observed, to a lesser extent, for Asian students.[14] This, of course, renders overall AP test results useless as a measure of school quality, because the success of native speakers on the foreign language tests has little to do with their schools' performance. To overcome this problem, we exclude AP foreign language tests when computing our AP performance metric.

To gain further insight into school performance, we also separately report the number of AP passing scores in mathematics and science alone, per black and Hispanic student. Math and science are areas generally considered to be affected most strongly by school instruction as opposed to home environment, and they include some of the most challenging AP tests.

The methodology used to analyze all of the above data is described in Appendix B.

Findings

Regression results for charter network performance on the California Standards Tests, broken down by each of the eight SES subgroups, can be found in Appendix C. The average of those effects is presented in Table 1, sorted from best to worst, along with the amount of grant funding each network has received. As a point of comparison, the scores for two of the nation's most elite and academically selective public schools, Lowell High in San Francisco and Gretchen Whitney High outside of Los Angeles, have also been included. The average result for charter schools not belonging to a multi-school network are reported under the heading "Other Charter."

For perspective, recall the effect size categorization presented earlier in this paper. An effect around 0.5 SD is considered "moderate" and one near 0.8 SD is considered "large." The American Indian charter school network is an astonishing 4 SD above the statewide public school mean. For further perspective, it is worth noting that low-income black and Hispanic students attending American Indian charter schools outperform middle- and upper-income white and Asian students attending conventional public schools in most subjects.

Three things become obvious in reviewing these academic performance results:

1. Charter schools that are not part of a

The American Indian charter school network is an astonishing 4 SD above the statewide public school mean.

Table 1
Charter Network Performance on the CST (Relative to District Public Schools) and Grant Funding

Charter Network	Average Effect	Rank	Total Grants ($)	Grant Rank
American Indian Public Charters	4.21	1	1,229,000	21
Oakland Charter Academies	3.76	2	660,000	27
Wilder’s Foundation	2.48	3	200,000	39
Rocketship Education	2.46	4	11,682,500	10
Whitney High (selective)	1.98	-	-	-
Camino Nuevo	1.78	5	1,532,050	20
Lowell High (selective)	1.68	-	-	-
East Oakland Leadership	1.58	6	235,000	37
Knowledge Is Power Program (KIPP)	1.44	7	16,821,926	7
Celerity Education Group	1.43	8	475,000	31
Synergy Academies	1.22	9	240,000	36
Environmental Charter Schools	1.19	10	498,000	29
St. Hope Public Schools	1.09	11	528,252	28
Crescendo Schools	0.91	12	$0	61
Partnerships to Uplift Communities	0.86	13	5,240,636	12
Sherman Thomas	0.84	14	0	61
Alliance College-Ready	0.83	15	19,065,677	4
Today’s Fresh Start	0.81	16	10,000	49
Charter Academy of the Redwoods	0.80	17	0	61
Bright Star	0.67	18	405,000	33
The Accelerated School	0.63	19	17,222,725	6
Larchmont Charter School	0.56	20	784,500	25
Lighthouse	0.56	21	361,500	34
Leadership Public Schools	0.52	22	3,215,450	14
Aspire	0.41	23	36,299,474	1
American Heritage Education	0.41	24	329,000	35
Value Schools	0.40	25	874,200	24
Nova Academy	0.37	26	0	61
New Designs Charter School	0.37	27	101,346	40
The Learner-Centered School	0.34	28	0	61
EJE Academies Charter School	0.33	29	5,704	50
Watts Learning Center	0.29	30	2,810,500	15
Jerry Brown’s Charter Schools	0.25	31	13,220,808	8
The Classical Academies	0.22	32	40,000	42
Inner City Education Foundation	0.21	33	18,405,092	5

Charter Network	Average Effect	Rank	Total Grants ($)	Grant Rank
Albert Einstein Academies	0.16	34	468,750	32
King-Chavez Public Schools	0.15	35	781,000	26
Wiseburn 21st Century	0.15	36	10,000	49
Literacy First Charter Schools	0.12	37	0	61
Magnolia Schools	0.10	38	960,000	22
Downtown College Preparatory	0.08	39	2,059,230	18
Education for Change	0.05	40	2,587,000	16
Semillas Community Schools	0.04	41	0	61
Green Dot Public Schools	0.04	42	32,701,166	2
Other Charter	-0.04	43	21,579,186	3
California Virtual Ed Partners	-0.21	44	0	61
Envision Schools	-0.29	45	10,979,500	11
Connections Academy	-0.31	46	0	61
High Tech High	-0.32	47	12,214,951	9
Rocklin Academy Charter Schools	-0.34	48	0	61
Para Los Ninos	-0.38	49	2,546,134	17
Community Learning Center	-0.50	50	941,292	23
Century Community	-0.50	51	210,000	38
Mare Island Technology Academy	-0.70	52	0	61
University Charter Schools, CSU	-0.70	53	20,000	46
CiviCorps Schools	-0.76	54	53,900	41
California Virtual Academy	-0.77	55	0	61
California Montessori Project	-0.87	56	37,800	43
Golden Valley Charter Schools	-0.94	57	490,000	30
Western Sierra Charter Schools	-0.94	58	0	61
Big Picture Learning	-1.09	59	3,291,393	13
Gateway Community Charters	-1.12	60	0	61
Agape Corporation	-1.20	61	5,000	51
New City Public Schools	-1.22	62	1,625,000	19
New Jerusalem Charter Schools	-1.26	63	0	61
Tracy Learning Center	-1.27	64	0	61
Aveson Charter Schools	-1.32	65	30,000	44
Innovative Education Management	-1.43	66	4,124	52
Santa Barbara Charter School	-1.64	67	14,950	47
Escuela Popular del Pueblo	-2.19	68	20,000	46

The correlation between the amount of grant funding charter networks receive and the length of their names is −0.23, *twice as strong as the link between funding and performance.*

network (see the "Other Charter" row in Table 1), which constitute the vast majority of charter schools in the state, perform at roughly the same level as district public schools (consistent with the findings of earlier research).

2. There is wide variation in performance among California's charter school networks.
3. The best charter school networks are far ahead of the statewide average of conventional public schools.

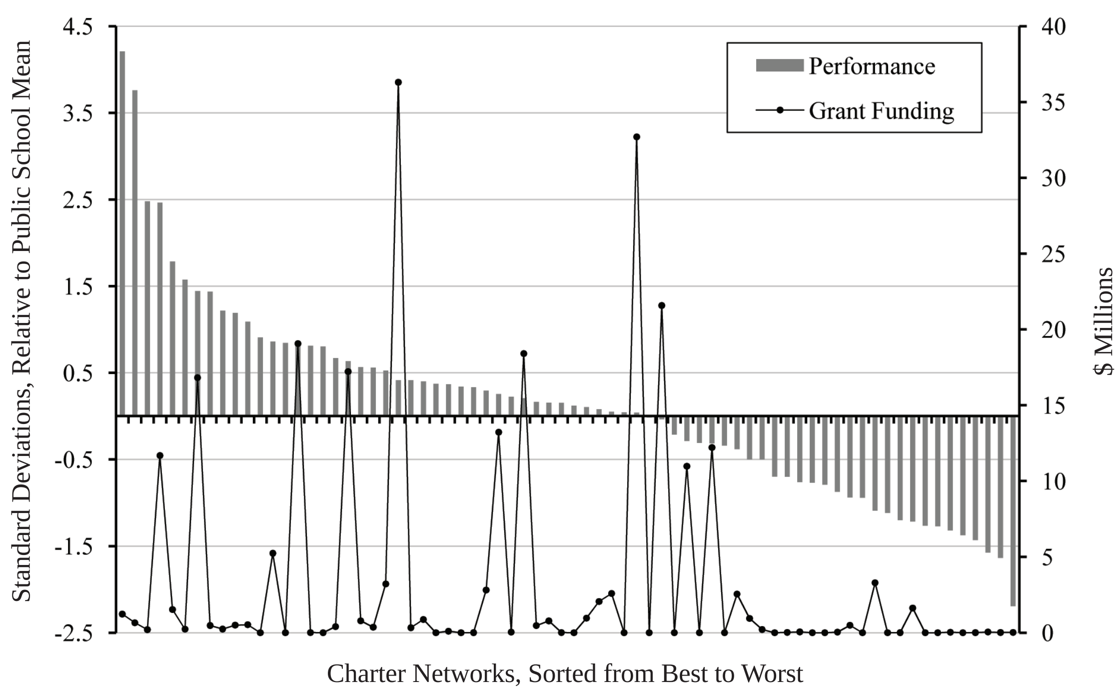
To address the central question of this study, we can contrast charter networks' academic performance with the amount of philanthropic funding they have received over the past eight years.[15] The first sign of a problem is that the three highest-performing charter networks are ranked 21st, 27th, and 39th in terms of the grant funding they have received, out of 68 networks (see the "Grant Rank" column of Table 1).

A more comprehensive quantitative assessment of the relationship can be obtained by calculating the correlation coefficient (Pearson's *r*) between the "Average Effect" and "Total Grants" columns in Table 1. That value is 0.11. Correlations range from −1 (perfect negative correlation), to zero (no correlation), to +1 (perfect positive correlation). Correlation values below 0.2 are typically considered negligible.[16]

It is also helpful to keep in mind that spurious weak correlations are to be expected as a result of random chance. For instance, the correlation between the amount of grant funding charter networks receive and the length of their names is −0.23, *twice as strong as the link between funding and performance* (although still very weak in absolute terms).

Perhaps the best way of appreciating the disconnect between charter network performance and grant funding is to superimpose the two data series on the same chart, as has been done in Figure 1.

Figure 1
Charter Network Performance on the CST Relative to District Public Schools, and Grant Funding (in $Millions)



The bars represent the academic effect sizes (left-hand scale) associated with each of the charter school networks (column two of Table 1, excluding Lowell and Whitney). Each bar is vertically aligned with a dot that represents the grant funding received by that charter network, in millions of dollars (right hand scale). If there were a strong positive correlation between the two, the dots would follow the height of the bars. In practice, the two data series seem to have nothing to do with one another.

Note that it is common in statistical analyses to be concerned with "outliers"—individual observations that are far outside the range in which the other observations are clustered. Outliers can be a product of measurement error, and, where there is reason to believe this is the case, they are generally dropped from the dataset so they do not skew the results. In the present study, however, we have no evidence that the top performing networks, or the best-funded networks, have greater measurement error than the others. And, even when they are dropped, the correlation between performance and grant funding remains negligible (ranging from 0.14 to 0.17).

Appendix D presents a series of tests to determine if selection bias can explain these results. The evidence suggests that it cannot.

Appendix E investigates whether charter network performance approaches the public school average as enrollment rises—in other words, whether scaling-up charter networks inevitably makes them mediocre. The evidence indicates conclusively that, in California at least, it does not.

AP Test Results

Table 2 presents the number of passing AP scores per student, for black and Hispanic students. The results are for the year 2010 and include only charter networks with at least one high school enrolling black or Hispanic students. Lowell and Whitney selective public high schools are again included for comparison.

These results are even more striking than those for CST scores. The correlations between AP performance and charter network grant funding are negative, though negligible, ranging from −0.01 to −0.05. Of the 48 charter networks that enroll black or Hispanic high school students, only 20 have black or Hispanic students who passed an AP test (excluding foreign language tests, as discussed in the "Data" section). In comparison, the statewide public school average number of AP passes per student, for black and Hispanic students, is 0.046.

The gap between the best and the rest is also considerably larger. Overall, the American Indian charter schools network has more than four times as many passing AP scores per black and Hispanic student as its closest charter network competitor (Knowledge Is Power Program). In mathematics and the sciences, American Indian has more than 20 times as many passing scores per black/Hispanic student as its closest charter competitors, and six to nine times as many such scores as even the selective Lowell and Whitney high schools.

It is important to note in reviewing these AP results that it was not possible to adjust them for either the individual family income of the students or the peer effects associated with low schoolwide income. Since most charter networks are majority low-income (including the highest performers), this unavoidable omission would likely have little impact on the correlations between AP performance and grant funding, although it would have some impact on the precise rankings of the charter networks. For this reason, the percentage of students qualifying for free or reduced price lunches at each of the charter networks is also included in the table.

Figure 2 illustrates the disconnect between charter network funding and minority student AP pass rates. As with Figure 1, Lowell and Whitney are omitted from this chart.

Although participation in the AP program is voluntary, and not every high-quality

The correlations between AP performance and charter network grant funding are negative, though negligible, ranging from −0.01 to −0.05.

Table 2
Passing AP Scores per Black and Hispanic Student, and Grants Received

Charter Network	Passes / Student (excluding foreign language)	Passes / Student (math and science)	% Free or Reduced Lunch*	Grant Rank	Total Grants ($)
American Indian Public Charters	1.00	0.52	94	21	1,229,000
Whitney High (selective)	0.46	0.06	15	-	-
Lowell High (selective)	0.42	0.08	36	-	-
Knowledge Is Power Program (KIPP)	0.23	0.00	68	7	16,821,926
Charter Academy of the Redwoods	0.14	0.00	63	61	0
Oakland Charter Academies	0.12	0.00	88	27	660,000
Magnolia Schools	0.09	0.02	82	22	960,000
Alliance College-Ready	0.05	0.00	93	4	19,065,677
Camino Nuevo	0.05	0.01	97	20	1,532,050
St. Hope Public Schools	0.04	0.00	71	28	528,252
Inner City Education Foundation	0.03	0.00	44	5	18,405,092
Downtown College Preparatory	0.03	0.00	83	18	2,059,230
The Accelerated School	0.02	0.02	79	6	17,222,725
Bright Star	0.02	0.00	88	33	405,000
Green Dot Public Schools	0.02	0.00	90	2	32,701,166
American Heritage Education	0.02	0.00	14	35	329,000
Jerry Brown's Charter Schools	0.01	0.00	56	8	13,220,808
Nova Academy	0.01	0.00	90	61	0
Partnerships to Uplift Communities	0.01	0.00	82	12	5,240,636
Leadership Public Schools	0.01	0.00	80	14	3,215,450
Value Schools	0.01	0.00	99	24	874,200
New Designs Charter School	0.01	0.00	84	40	101,346
East Oakland Leadership	0.00	0.00	92	37	235,000
Environmental Charter Schools	0.00	0.00	79	29	498,000
Sherman Thomas	0.00	0.00	-	61	0
Lighthouse	0.00	0.00	86	34	361,500
Aspire	0.00	0.00	74	1	36,299,474
The Classical Academies	0.00	0.00	14	42	40,000
Wiseburn 21st Century	0.00	0.00	-	49	10,000
Literacy First Charter Schools	0.00	0.00	16	61	0
Semillas Community Schools	0.00	0.00	87	61	0
California Virtual Ed Partners	0.00	0.00	-	61	0
Envision Schools	0.00	0.00	59	11	10,979,500
Connections Academy	0.00	0.00	38	61	0

Charter Network	Passes / Student (excluding foreign language)	Passes / Student (math and science)	% Free or Reduced Lunch*	Grant Rank	Total Grants ($)
High Tech High	0.00	0.00	37	9	12,214,951
Community Learning Center	0.00	0.00	11	23	941,292
Mare Island Technology Academy	0.00	0.00	52	61	0
CiviCorps Schools	0.00	0.00	-	41	53,900
California Virtual Academy	0.00	0.00	26	61	0
Golden Valley Charter Schools	0.00	0.00	-	30	490,000
Western Sierra Charter Schools	0.00	0.00	-	61	0
Big Picture Learning	0.00	0.00	91	13	3,291,393
Gateway Community Charters	0.00	0.00	62	61	0
Agape Corporation	0.00	0.00	96	51	5,000
New City Public Schools	0.00	0.00	96	19	1,625,000
New Jerusalem Charter Schools	0.00	0.00	39	61	0
Tracy Learning Center	0.00	0.00	7	61	0
Aveson Charter Schools	0.00	0.00	27	44	30,000
Innovative Education Management	0.00	0.00	34	52	4,124
Escuela Popular del Pueblo	0.00	0.00	92	46	20,000

*Percentage of students eligible for free or reduced price lunches, calculated as an enrollment-weighted average across high schools in the given network enrolling black or Hispanic students.

school participates, participation is nevertheless sufficiently widespread to offer suggestive evidence about high-end academic performance at the high school level. Based on that evidence, there appears to be a great chasm between what is academically possible for minority students and what is currently being achieved by virtually all of California's charter school networks and traditional district schools.

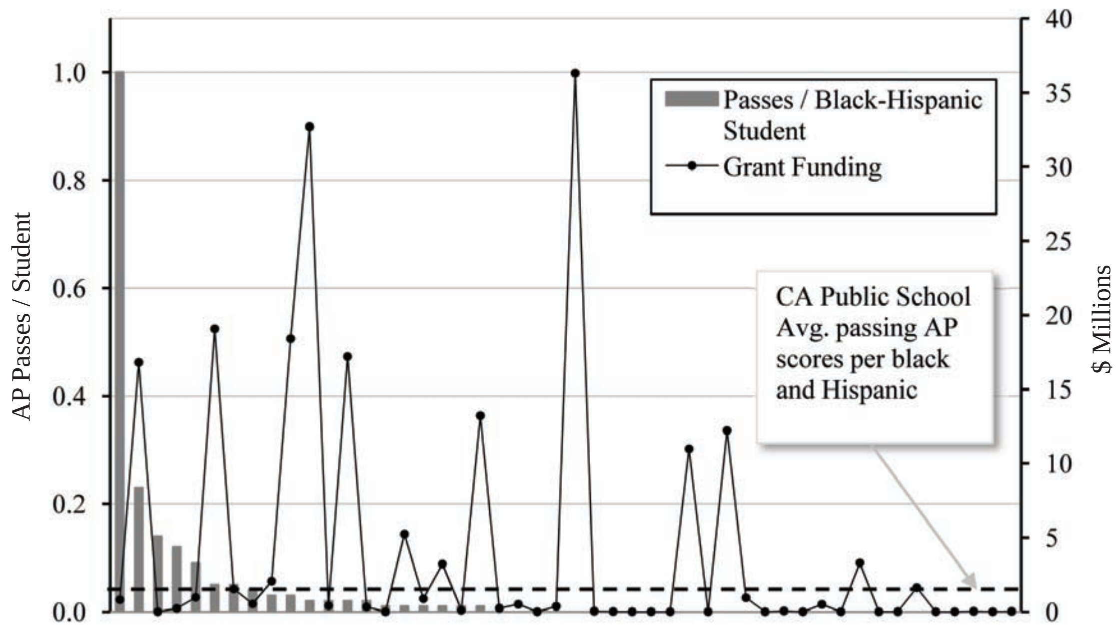
Conclusions

Philanthropists have shown great generosity to charter schools in recent times, donating roughly $250 million to California charters alone over the past eight years. Regrettably, this generosity appears to have been disconnected from the academic performance of charter school networks. It is as if the same random lottery that determines admission to oversubscribed charter schools were being used to allocate grants. Our nation's search for a system that will reliably scale up the best schools and crowd out the poor-performing ones remains unrequited.

That does not mean philanthropists should stop funding all charter networks. The best networks not only transform the educational and career prospects of the children they serve, they play an indispensable policy role, dramatically illustrating the failure of our current educational arrangements to fulfill every child's potential. We should indeed strive to sustain and replicate these beacons of excellence.

It is as if the same random lottery that determines admission to oversubscribed charter schools were being used to allocate grants.

Figure 2
Passing AP Scores per Black and Hispanic Student (Excluding Foreign Language), and Grants Received



But we can no longer assume that philanthropy is a reliable mechanism for doing so in a systematic way. If education is ever to enjoy the automatic replication of excellence that we take for granted in every other field, a proven system for achieving that goal must be identified and popularized. Until then, millions of children will continue to have their hopes and dreams dashed by schools that do not begin to realize their full potential.

A good place to start in identifying such a system would be to ask why excellence routinely scales up in most fields but is elusive in education. Why does the world beat a path to the door of whoever builds a better cell phone or sells a better cup of coffee but not to those who find a better way to teach math or science? What distinguishes education from other fields, structurally and economically?

Is it purely a coincidence that in those places where education operates within the free enterprise system it enjoys the replication of success typical of that system?

Another starting point is to look at education systems around the world in search of places where the best schools and the best teachers *do* reach large and growing audiences. Why has the for-profit Kumon network of tutoring schools grown to serve 4 million students in 42 countries, while the nonprofit KIPP, one of the fastest growing charter school networks in America, serves fewer than 30,000? Is it purely a coincidence that in those places where education operates within the free enterprise system it enjoys the replication of success typical of that system?

When we answer these questions, and act accordingly, education will join the ranks of fields in which excellence regularly catches on like wildfire. Until we do, it will remain a floating candle in a sea of darkness, isolated and transitory.

Appendix A: Selection Bias

This section reviews estimates of selection bias reported in comparisons between charter schools and public schools as well as between private schools and public schools, after controls for student SES (and sometimes peer effects). We are being very conservative by including private vs. public studies because selection effects should be larger between sectors than between public charter and public district schools, since the need to pay tuition poses an additional barrier to private school consumption (and hence may be associated with greater selectivity) and because private schools are generally free to consider the prior academic performance of their applicants, whereas charter schools are not.

First, empirical estimates of selection bias effects vary not only in magnitude but in *direction*. In other words, some studies find that selection bias works *against* charter and private schools and in favor of traditional public schools. For instance, while reviewing the research in 1998, University of Chicago economist Derek Neal noted that studies had been mixed in their estimates of selection bias effects for Catholic school attendance, with most showing weak positive or negative effects, and none showing large positive effects (i.e., none showing a large selection bias in favor of Catholic schools).[17] The same remains true today with respect to selection into public charter schools or private schools, as the literature review below demonstrates.

The well-known 1981 study by Coleman, Hoffer, and Kilgore (CHK) comparing public and Catholic schools found a 0.15 to 0.2 standard deviation advantage to Catholic school attendance after SES controls (depending on subject) but failed to control for selectivity.[18] A 1985 reanalysis by Willms specifically aimed to control for selection bias and found a Catholic school effect between 0 and 0.1 (depending on subject), suggesting a rough upper bound on the effect of selectivity of 0.2 SD.[19]

A separate reanalysis of the 1981 CHK study, also aiming to control for selection bias, echoed Willms's results, once again indicating an upper bound on selection bias of roughly 0.2 SD.[20]

Christopher Jepsen (2003) found evidence of negative selection into Catholic schools, but because he was not able to offer a precise estimate of its magnitude, the conservative approach for our purposes is to treat his results as indicating an upper bound of 0 on the effect of selectivity.[21]

In 1996, Dan Goldhaber found that controlling for selection bias had no impact on sectoral achievement coefficients when comparing public to private schools overall, or when comparing two subsets of private schools (Catholic and "elite") to public schools, a result that corresponds to a selection bias effect size of zero.[22]

Hanushek et al. (2006), who compared charter schools to district public schools, found that their model addressing selection effects lowered the charter school effect by between 0.03 and 0.15 SD.[23]

A 1996 study by Adam Gamoran comparing chosen public "magnet" schools, Catholic schools, and secular private schools to district public schools found a range of selectivity effects for the chosen sector ranging from −0.09 SD to 0.2 SD, consistent with the findings of the other studies discussed above.[24]

In their 1997 study comparing achievement in religious and secular private schools to that in district public schools, Figlio and Stone report selection bias effects equivalent to between −0.06 SD and 0.3 SD.[25] A recent Vanderbilt University study also finds a maximum estimated selection effect of 0.3 SD.[26]

Last year, researchers from Mathematica, Vanderbilt, and Florida State University were unable to reject the possibility that there was no selection bias at work in the decision to choose charter over district public schools. They further concluded that if selection bias was at work it was negative: after SES controls, the students who chose to enroll in charter high schools were (if anything) dis-

Some studies find that selection bias works *against* charter and private schools.

advantaged relative to those who chose traditional district schools.[27] Their results thus place an upper bound on the selection bias effect of zero.

Two other studies that discuss private school selection bias, one by Carbonaro and Covay and another by Altonji, Elder, and Taber, do not offer specific estimates for the selection effect. Nevertheless, they present the results of several tests indicating that, to the extent it might be present, the selection effect is unlikely to be of a magnitude outside that of the studies discussed above.[28]

Based on this literature, the largest selection bias effect favoring charter or private schools is 0.3 SD.

Further evidence that selection effects cannot plausibly explain away the performance of the top charter school networks comes from the fact that some of the largest and most consistent positive effect sizes for charters have been found in randomized lottery experiments that greatly reduce selection bias as a concern (because selection into the school is done at random from the population of applicants). As Nicotera, Mendiburo, and Berends observed in late 2009:

> Currently there are four studies of charter school student achievement that have used the lottery-in/lottery-out research design. Hoxby & Rockhoff (2004) examined achievement effects of students in nine Chicago charter schools; Hoxby et al. (2007, 2009) are conducting an ongoing study of New York City's charter schools; Abdulkadiroglu et al. (2009) studied the effect of charter schools in Boston; and McClure et al. (2005) examined one school in California, which limits its usefulness for generalization. Results from these charter school studies have been overwhelmingly positive. In Chicago, charter students in kindergarten through fifth grade, students improved 6 to 7 percentile points in math and 5 to 6 percentile points in reading. In New York City, charter school students had higher achievement in math and reading in all grade levels compared with their counterparts who lost the lottery. And in Boston, students who attend middle and high school charter schools outperform students in the traditional public schools.[29]

There is still the concern with lottery experiments that applicants to oversubscribed charter schools might differ in unmeasured respects from those who did not apply to them, and therefore these results might not generalize to the population as a whole. But, at least within the applicant population, selection bias is eliminated as a factor. Furthermore, some of the lottery-studied charter networks that show gains not explainable by selection effects also operate in California and significantly *underperform* the top-scoring California charter networks (making it even less likely that those top-scoring networks are largely dependent for their advantages on selection bias).

Some of the lottery-studied charter networks that show gains not explainable by selection effects also operate in California and significantly *underperform* the top-scoring California charter networks.

Appendix B: Methodology

Model

Since California Standards Test score averages are broken out by eight different SES subgroups (Asian, black, Hispanic, and white students from low-income and non-low-income families), and since the performance of charter networks with each of these different subgroups is of significant policy interest, we perform eight separate regressions, one for each subgroup.

Within each SES category, we have a hierarchical dataset composed of standardized average class test scores within charter networks. And since we have data for the entire population of California charter school networks, a fixed-effects hierarchical model is appropriate. This model can be written as follows:

$$z\_Score_{ij} = \alpha + \beta_j \times NetworkDummy_j + \gamma \times SchoolwideIncome_i + \varepsilon_{ij},$$

where z_Score_{ij} is the mean standardized score for class i in charter network j, α is a constant, $NetworkDummy_j$ is an array of dummy variables corresponding to each of the charter school networks (district schools are the omitted category) and β_j is an array of coefficients giving the effect of attending network j (measured in standard deviations), γ is the "peer effect" coefficient, and $SchoolwideIncome_i$ is the schoolwide percentage of students eligible for free or reduced priced meals for class i, and ε_{ij} is an error term.

The "peer effect" on achievement is well known: when students are surrounded by high-income peers it tends to raise achievement, while lower peer income tends to lower achievement, other things being equal. It is, of course, not the case that money per se is responsible for this relationship. Rather, higher-income families tend to have better-educated parents whose characteristics and behavior are more conducive to their children's academic success. And a peer group composed of such advantaged students facilitates the educational process for everyone, regardless of a given child's own family background.

That does not imply, however, that all schools (or charter networks) are equally susceptible—some do a better job than others, for instance, of overcoming negative peer effects. When they manage to do so, it is a sign of their superior practices, not an indication that the peer effect itself was absent or abated. Expressed in econometric jargon, average family income is thus a "fixed" effect.[30]

For each of our eight SES categories, we have average test scores spanning grades from elementary school through high school and for many different academic subjects. Since we are interested in knowing how well the various charter networks perform overall, it makes sense to combine their results across subjects and grades. While raw test scores are not directly comparable across grades and subjects, it is possible to mathematically transform them to make them so—a process known as "standardization" that was described earlier in this paper.[31]

Note that our model omits the school as a level in the hierarchy, since we are interested in comparing charter networks to one another rather than comparing the schools within them. If our results revealed little variation in performance across networks, that would suggest that charter networks are not academically important constructs in California, and that whatever variation in test scores exists must be confined to the student, class, and school levels. This, as the following sections make plain, does not appear to be the case.

Also, because our California Standards Test observations are class-level score averages rather than individual student scores, their relative weights in the regressions must be allowed to vary based on the number of students' scores from which each average was calculated. The statistical analysis software used for this report, Stata, provides the "analytic" weighting option specifically for this situation. Note that analytic weights are not equivalent to simply replicating the given observation n times and then running the regression normally (where n is the number of students from which the average is calculated). Such replication of observations is achieved through "frequency" weights in Stata. A technical discussion of analytic weighting can be found on the Stata website.[32]

Finally, because the curriculum and teaching strategies in use within a given charter network are usually quite similar from one classroom to another, the independence-of-observations assumption of Ordinary Least Squares regression is violated (possibly resulting in artificially inflated precision of our coefficients). This concern is dealt with by using Huber/White robust standard errors, clustering on charter network.

Because we lack the necessary data to control for individual student income[33] and income-related schoolwide peer effects,[34] we cannot apply the above model to Advanced Placement test results. Given this data limitation, we simply report the ratio of passing AP scores per student for black and Hispanic students, which at least minimizes

Since we are interested in knowing how well the various charter networks perform overall, it makes sense to combine their results across subjects and grades.

Are the differences in achievement among charter school networks bigger than can be plausibly accounted for by within-subgroup selection bias?

the confounding effects of minority status on AP performance. To allow readers to take schoolwide income into consideration in evaluating these effects, we also report the percentage of students qualifying for free or reduced price lunches alongside the AP pass ratios.

The Problem of Selection Bias

A difficulty with measuring academic performance at a single point in time, rather than achievement gains over time, is that it confounds two different factors contributing to student achievement: the instruction provided by the schools and the characteristics of the students themselves. An effective school that happens to attract less motivated students might post below-average scores but nevertheless be doing a better job of serving those students than other schools would do. Such an uneven apportioning of unmeasured student characteristics among schools is known as selection bias.

The most we can do to mitigate selection bias using the STAR data is to compare the performance of the charter networks within SES subgroups (e.g., compare low-income Hispanics to each other, non-low-income Asians to each other, etc.). Some selection bias could remain, however, if different charter networks systematically attract families *within a given SES subgroup* that have differing levels of educational commitment, for example. Social scientists, particularly econometricians, have developed sophisticated methods for attempting to control for such within-subgroup selection bias, but the data to apply these methods are not available for the charter schools in our subject population.

This does not necessarily mean that our results will be inconclusive. The more sophisticated econometric methods are employed in education studies because the size of the instructional (or "treatment") effects being measured are often of similar magnitude to the confounding effect of selection bias, and so it is necessary to use very precise tools if researchers are to distinguish the one from the other.

But if the magnitudes of the treatment effects are substantially greater than any plausible level of bias, that degree of precision is no longer essential. In the context of the present study, therefore, we have to ask: are the differences in achievement among charter school networks (and between those networks and the district public school average) bigger than can be plausibly accounted for by within-subgroup selection bias? If so, we will have evidence of an instructional effect.

To answer that question, Appendix A reviews the selection bias effect sizes reported in the literature, and Appendix D compares them to the effect sizes identified in the present study.

Appendix C: Regression Results

The output of our eight SES subgroup regressions is presented in Table C1. The values in the rightmost eight columns represent the effect of attending the given charter network, measured in standard deviations above or below the statewide mean of district public schools for the given SES subgroup. In assessing the magnitude of these values, readers can refer to the earlier sidebar on standardized effect sizes. Empty cells are an indication that a given network had no test results for a particular SES subgroup. The California Department of Education does not report subject/grade results for groups of 10 or fewer students, to maintain student privacy.

As a point of comparison, the scores for two of the nation's most elite and academically selective public schools, San Francisco's Lowell High School and Gretchen Whitney High School outside of Los Angeles, were included in the regression. The import of their results is discussed in Appendix D. The average results for all independent charter schools (those not belonging to a multi-school network) are also reported in the row labeled "Other Charter."

Table C1
Charter Network Performance (California Standards Test) Relative to the Statewide Conventional Public School Mean

Charter Network	CST Academic Effect Sizes (SD)							
(+ Lowell & Whitney	Low Income				Not Low Income			
for comparison)	Asian	Black	Hispanic	White	Asian	Black	Hispanic	White
Agape Corporation	**-1.36**	**-1.12**	**-0.97**	**-1.35**				
Albert Einstein Academies			**0.29**				**0.30**	**-0.10**
Alliance College-Ready		**0.68**	**1.16**			**0.51**	**0.97**	
American Heritage Education			**0.41**	**1.01**			**-0.13**	**0.36**
American Indian Public Charters	**3.21**	**5.05**	**4.38**					
Aspire	**-0.80**	**0.91**	**1.14**	**0.46**	**-0.84**	**0.97**	**0.94**	**0.52**
Aveson Charter Schools						**-1.80**	**-1.45**	**-0.71**
Big Picture Learning			**-1.16**					**-1.02**
Bright Star			**1.55**				**-0.22**	
California Montessori Project				**-1.14**			**-0.86**	**-0.62**
California Virtual Academy		**-0.26**	**-0.56**	**-0.36**		**-1.07**	**-1.25**	**-1.12**
California Virtual Ed Partners			**-0.09**			**-0.12**	**-0.12**	**-0.53**
Camino Nuevo Charter Academy			**1.70**				**1.87**	
Celerity Education Group		**1.24**	**1.63**					
Century Community		**-0.18**	**-1.31**			**-0.13**	**-0.38**	
Charter Academy of the Redwoods			**-0.06**	**1.67**				
CiviCorps Schools		**-0.86**				**-0.66**		
Community Learning Center							**-0.80**	**-0.20**
Connections Academy			**-0.40**	**-0.24**			**-0.13**	**-0.48**
Crescendo Schools		**0.98**	**-0.03**			**1.77**		
Downtown College Preparatory			**0.06**				**0.10**	
East Oakland Leadership			**1.35**				**1.80**	
Education for Change		**-0.35**	**0.45**					
EJE Academies Charter School			**0.12**				**0.55**	
Environmental Charter Schools		**1.30**	**0.87**				**1.41**	
Envision Schools		**-0.08**	**-0.31**	**-0.75**		**-0.14**	**-0.20**	**-0.25**
Escuela Popular del Pueblo			**-1.21**				**-3.18**	
Gateway Community Charters	**-1.55**	**-0.95**	**-1.00**	**-1.28**		**-1.28**	**-1.34**	**-0.41**
Golden Valley Charter Schools							**-1.00**	**-0.88**
Green Dot Public Schools		**-0.10**	**0.21**			**0.64**	**-0.59**	
High Tech	**-0.71**	**-0.53**	**-0.12**	**-0.12**		**-0.27**	**-0.14**	
Inner City Education Foundation		**0.16**	**0.50**		**0.08**	**0.09**		
Innovative Education Management	**-2.35**		**-1.17**	**-1.32**	**-0.78**		**-1.44**	**-1.53**
Jerry Brown's Charter Schools	**0.38**	**0.61**	**-0.01**			0.05		

Continued next page

Table C1 *Continued*

Charter Network (+ *Lowell & Whitney for comparison)*	CST Academic Effect Sizes (SD)							
	Low Income				Not Low Income			
	Asian	Black	Hispanic	White	Asian	Black	Hispanic	White
King-Chavez Public Schools			**0.07**				**0.24**	
Knowledge is Power Program	**0.88**	**1.81**	**1.58**		**1.34**	**1.11**	**1.92**	
Larchmont Charter School			**0.51**					**0.62**
Leadership Public Schools		**0.28**	**0.23**				**1.06**	
Lighthouse			**0.56**					
Literacy First Charter Schools			**0.73**	**0.05**			**-0.25**	**-0.04**
Lowell High (selective)	**1.51**		**1.61**	**2.77**	**1.18**	**0.76**	**1.79**	**2.13**
Magnolia Schools		**1.26**	**0.78**	**0.83**		**-1.00**	**-0.52**	**-0.74**
Mare Island Technology Academy			**-0.98**			**-0.51**	**-1.03**	**-0.28**
New City Public Schools			**-1.22**					
New Designs Charter School		**0.31**	**0.45**			**-0.10**	**0.80**	
New Jerusalem Charter Schools			**-0.56**	**-1.87**			**-1.41**	**-1.22**
Nova Academy			**0.37**					
Oakland Charter Academies			**3.76**					
Other Charter	**0.56**	**0.08**	**0.01**	**-0.46**	**-0.02**	**0.14**	**-0.20**	**-0.41**
Para Los Ninos			**-0.38**					
Partnerships to Uplift Communities			**0.95**				**0.76**	
Rocketship Education			**2.46**					
Rocklin Academy Charter Schools					**-0.59**			**-0.09**
Santa Barbara Charter School								**-1.64**
Semillas Community Schools			**0.04**					
Sherman Thomas			**0.84**					
St. Hope Public Schools	**0.27**	**1.59**	**1.17**			**1.59**	**0.81**	
Synergy Academies			**1.66**				**0.77**	
The Accelerated School		**1.23**	**0.73**				**-0.07**	
The Classical Academies				**0.72**			**0.44**	**-0.48**
The Learner-Centered School								**0.34**
Today's Fresh Start		**0.36**	**1.27**					
Tracy Learning Center					**-1.87**	**-1.15**	**-1.05**	**-1.02**
University Charter Schools, CSU			**-1.31**				**-0.42**	**-0.37**
Value Schools			**0.40**					
Watts Learning Center		**1.65**	**-1.07**					
Western Sierra Charter Schools								**-0.94**
Whitney High (selective)	**1.87**		**2.41**		**1.20**			**2.43**
Wilder's Foundation		**2.76**				**2.20**		
Wiseburn 21st Century		**1.08**	**0.35**	**-0.43**		**0.17**	**0.15**	**-0.42**

Effects significant at the $p < 0.05$ level or better are shown in boldface, but the number of significant findings is likely overestimated due to the lack of independence of observations across subjects within grades, discussed in endnote 31. Even taking that into account, the precision of our estimates likely exceeds the levels normally seen in intersectoral education outcome studies. The reason is that such studies are typically based on small samples of students (often just a few hundred and seldom more than a few thousand) and the models used to analyze them must extrapolate from those small samples to the population at large, which lowers the precision of estimates. In the present study, by contrast, we have data for the entire population of California charter and district public schools, obviating the need to generalize to a larger population. The number of observations is also very large, ranging from tens of thousands to hundreds of thousands (with each observation, in turn, representing an average of the performance of multiple students).

Note that in order to obtain an overall academic effect associated with attending one of California's charter networks, it is necessary to combine the individual SES effects. There are several ways in which this could be done, each with its advantages and disadvantages.[35] The simplest of these is offered in column 2 of Table 1, above: an unweighted average.

Appendix D: Assessing Selection Bias

In light of the results reported in the body of this study, two questions remain:

- Is it possible that the lack of correlation between grant funding and performance is an artifact of selection bias?
- Can the performance of the top charter networks be discounted as the result of selection bias?

We can answer the first question empirically by assigning to each charter network effect in Table 1 a random selection bias value in the range identified in the scientific literature, and then recomputing the correlation. If we repeat this process many (say, 2,000) times and take the highest and lowest correlation values out of all those iterations, we can obtain a reasonable upper bound on the impact that selection bias could have. This is known as a "Monte Carlo simulation," or "random robustness testing."

As noted in Appendix A, the largest selection bias effect in the literature (after controls for student SES and peer effects) is about 0.3 SD. To be conservative, we can raise that to 0.5 SD. Randomly applying a selection bias effect between −0.5 SD and +0.5 SD to each charter network effect, and repeating this process 2,000 times, we find a maximum correlation between grant funding and performance of 0.18 and a minimum correlation of −0.03. Both of these remain within the "negligible" category. Hence, it is unlikely that selection bias is responsible for the lack of correlation between grant funding and charter network performance.

One way of answering the second question is to compare the largest selection bias effect sizes from the literature to the charter network effects identified in Table 1. Clearly, the top networks have effects an order of magnitude larger than the high end of the selection bias effects reported in the literature, suggesting that selection bias could explain at most only a small fraction of their performance advantage. Indeed, the top 29 charter school networks have effects larger than the high bound on selection bias effects.

Another test of this question is to compare the performance of the top charter networks—which must accept all applicants or employ random lotteries if oversubscribed—to the performance of two of the nation's most elite and academically selective public schools: San Francisco's Lowell High School and Gretchen Whitney High School outside of Los Angeles. Both of the latter schools re-

It is unlikely that selection bias is responsible for the lack of correlation between grant funding and charter network performance.

The top four charter networks outperform both of the elite, selective public schools, and a fifth charter network outperforms Lowell.

ceive many times more applicants than they can accept, and both consider students' prior test scores as a criterion for admission. Of the 2,500 students who applied for fall 2007 admissions at Lowell (the most recent year available), only 648 were accepted, a rejection rate of nearly 75 percent.[36] Whitney is a similarly academically selective institution, having been featured repeatedly in *Newsweek*'s annual short list of "public school elites," where it is described as "a suburban version of the New York–area superschools, with very competitive admission."[37]

This deliberate, systematic selection process undoubtedly produces a much more powerful effect on those schools' performance than does any incidental self-selection by parents that might occur in a charter school setting. Nevertheless, the top four charter networks outperform both of the elite, selective public schools, and a fifth charter network outperforms Lowell (see Table 1, in the body of the text). In many cases, the gaps are quite large. Moreover, an additional four or five charter networks perform only slightly below the level of the selective public schools. These findings further undermine the notion that selection bias could play a major role in explaining the success of the top charter networks.

Appendix E: Does Growth Necessarily Beget Mediocrity?

This study evaluates the average academic achievement of charter school networks. But averages have an important statistical property: they vary less as the number of observations from which they are calculated grows. If you were to conduct 10 different surveys to find out the average height of American men, basing each survey on a random sample of only 15 observations, the resulting 10 averages would vary widely. But if each survey randomly sampled 15,000 men, the resulting 10 averages would all be very similar. In mathematical terms, the variance of a mean calculated from a sample drawn from a normally distributed population is inversely proportional to the square root of the sample size.

Given this statistical fact, we might expect to see greater variance in the average performance of small charter networks than of large charter networks, *even if the networks themselves do not in fact differ in their academic effectiveness*. And if this statistical principle is driving the observed performance differences between charter networks, we would expect network performance to approach the statewide public school average as total enrollment increases.

We can test for that possibility by running a fixed-effects time-series regression of the absolute value of the charter networks' performance on their enrollment.[38] This will tell us if, as the networks have grown over the years, their performance has generally tended toward the statewide public school mean. If growth intrinsically begets mediocrity, the coefficient on the enrollment term in this regression will be statistically significant and negative (larger enrollments leading to smaller academic effect sizes, whether positive or negative), and the regression will explain a nonnegligible amount of the variation in charter network performance.

Running this regression, we find that it explains virtually none of the variance in charter network performance (R-squared "within" is an infinitesimal 0.003), and that the enrollment term is statistically insignificant at even the loosest accepted level of confidence ($p = .27$). The evidence thus contradicts the theory that the variance in charter network performance is a mere statistical artifact. Charter network performance has stagnated, improved, or declined independent of enrollment growth, most likely as a result of the pedagogical and managerial decisions charter school networks have made.

In a sense, this evidence makes the findings reported in the body of this paper all the more damning. We now know that growth is not an inherent barrier to the performance of charter school networks, yet the networks that have grown the most and been singled-out for scaling-up by philanthropists are no better than those that have not been singled-out for growth.

Notes

1. Julian R. Betts and Y. Emily Tang, "Value-Added and Experimental Studies of the Effect of Charter Schools on Student Achievement: A Literature Review," National Charter School Research Project, Center on Reinventing Public Education, University of Washington, December 2008.

2. Caroline M. Hoxby, Sonali Murarka, Jenny Kang, "How New York City's Charter Schools Affect Achievement," The New York City Charter Schools Evaluation Project, Cambridge, MA, September 2009.

3. One concern with charter lottery experiments is that only oversubscribed schools can be studied in this way, and it stands to reason that oversubscribed schools could be higher-performing than schools with many open places (if parents shun bad schools and favor good ones). That concern is abated in the Hoxby study since 94 percent of all charter students in New York City are enrolled in oversubscribed schools. It is of course still possible that average quality would go down if the number of charter schools expanded dramatically, but the fact that virtually all the existing charter students were included in the study is reassuring.

4. Center for Research on Education Outcomes (CREDO), " Multiple Choice: Charter School Performance in 16 States," CREDO, Stanford, CA, 2009.

5. Caroline M. Hoxby, "A Statistical Mistake in the Credo Study of Charter Schools," Stanford University and NBER, August 2009, http://credo.stanford.edu/reports/memo_on_the_credo_study%20II.pdf . (Note that, while dated "August 2009" the memo linked to above is a revised version released in the fall of 2009).

6. CREDO, "Fact vs. Fiction: An Analysis of Dr. Hoxby's Misrepresentation of CREDO's Research," Stanford University, October 7, 2009, credo.stanford.edu/reports/CREDO_Hoxby_Rebuttal.pdf. See also, Debra Viadero, "Scholars Spar Over Research Methods Used to Evaluate Charters," *Education Week*, October 8, 2009, updated May 8, 2010, http://www.edweek.org/ew/articles/2009/10/08/07credo.h29.html?tkn=MVRFgsPOB68KystIWPv%2Brcppn0GSaoJQqq%2Bz.

7. See the review of these studies in Susan Dynarski et al., *Charter Schools: A Report on Rethinking the Federal Role in Education* (Washington: Brookings Institution, December, 2010), http://www.brookings.edu/reports/2010/1216_charter_schools.aspx#_edn7. See also the review in Anna Nicotera, Maria Mendiburo, and Mark Berends, "Charter School Effects in an Urban School District: An Analysis of Student Achievement Gains in Indianapolis," paper prepared for the School Choice and School Improvement: Research in State, District and Community Contexts conference, Vanderbilt University, October 25–27, 2009, http://www.vanderbilt.edu/schoolchoice/conference/papers/Nicotera_COMPLETE.pdf.

8. Eric A. Hanushek et al., "Charter School Quality and Parental Decision Making with School Choice," *Journal of Public Economics* 91, no. 5–6 (June 2007): 823–48.

9. Note that four charter networks were omitted because they serve special student populations (returning dropouts, developmentally disabled children) associated with lower academic performance. It would be unfair to compare these networks directly to networks serving mainstream student populations, and there are not enough of them to meaningfully compare their performance relative to one another. The omitted networks are Opportunities for Learning, Options for Youth, Youth Build, and the CHIME Institute.

10. Foundation Center and Foundation Search America.

11. This is as far back as the data allow.

12. When grants were made to charter school networks that operate schools both within and outside of California, we prorated them on the basis of the fraction of their schools that is based in-state. So, for instance, a network that operated 10 schools in California and 10 in other states would be credited as having received $500,000 for its California operations if it received a $1 million dollar grant for its entire operation.

13. To protect student privacy, class-level CST score breakdowns by SES are only available for groups of 10 or more students. For the purposes of this study, however, in which test scores are aggregated across *multiple schools* within each network, this privacy criterion is unnecessarily restrictive, since many networks have a total of more than 10 students of a particular SES group in a given subject and grade even if they do not meet that threshold in any single one of their schools. An effort was made to obtain data for these smaller subgroups, but after only a single network proved willing to provide the scores out of dozens that were approached, that effort was abandoned.

Having obtained them, the data for that one SES group and network (low-income African American students at American Indian Charter Schools) are included in this study.

14. Leah Fabel, "Native Speakers Ace AP Language Exams," *Washington Examiner*, February 4, 2010, http://washingtonexaminer.com/news/nation/native-speakers-ace-ap-language-exams.

15. Note that total grant funding, rather than grant funding *per pupil*, is the correct measure. That is because enrollment is endogenous—it is a product, in part, of earlier grant funding. So, controlling for enrollment (which dividing by enrollment would do) would control away some of the very characteristics we are trying to measure: the charter network's ability to attract funding.

16. Any scale for characterizing correlations as "weak" or "strong" is necessarily arbitrary, but the use of such a scale is nevertheless essential in order to make judgments. To avoid bias, I offer, below, the scale most commonly referred to in the literature. By that scale, and indeed by any other of which I am aware, the correlation between charter network performance and grant funding is negligible.

Scale of Correlation Strength

< 0.2	Negligible
0.2–0.4	Weak
0.4–0.7	Moderate
0.7–0.9	Strong
> .9	Very strong

This scale is traceable to J. P. Guilford's *Fundamental Statistics in Psychology and Education* (New York: McGraw Hill, 1956), p. 145.

17. Derek Neal, "What Have We Learned about the Benefits of Private Schooling?" *Economic Policy Review*, 1998, pp. 80–81, http://www.ny.frb.org/research/epr/98v04n1/9803neal.pdf.

18. James Coleman, Thomas Hoffer, and Sally Kilgore, "Public and Private Schools: Report to the National Center for Education Statistics," National Opinion Research Center, Chicago, 1981.

19. J. Douglas Willms, "Catholic-School Effects on Academic Achievement: New Evidence from the High School and Beyond Follow-Up Study," *Sociology of Education* 58, no. 2 (1985): 98–114.

20. Karl L. Alexander and Aaron M. Pallas, "School Sector and Cognitive Performance: When Is a Little a Little?" *Sociology of Education* 58, no. 2 (1985): 115–28.

21. Christopher Jepsen, "The Effectiveness of Catholic Primary Schooling," *The Journal of Human Resources* 38, no. 4 (2003): 928–41.

22. Dan D. Goldhaber, "Public and Private High Schools: Is School Choice an Answer to the Productivity Problem?" *Economics of Education Review* 15, no. 2 (1996): 93–109.

23. Hanushek et al., pp. 823–48.

24. Adam Gamoran, "Student Achievement in Public Magnet, Public Comprehensive, and Private City High Schools," *Educational Evaluation and Policy Analysis* 18, no. 1 (1996): 1–18.

25. David N. Figlio and Joe A. Stone, "School Choice and Student Performance: Are Private Schools Really Better?" Institute for Research on Poverty Discussion Paper no. 1141–97, 1997.

26. Nicotera, Mendiburo, and Berends.

27. Kevin Booker et al., "The Effects of Charter High Schools on Educational Attainment." Paper prepared for the November 2010 meeting of the Association for Public Policy Analysis and Management, Boston, MA, pp. 18–19, https://www.appam.org/conferences/fall/boston2010/sessions/downloads/1017.4.pdf.

28. William Carbonaro and Elizabeth Covay, "School Sector and Student Achievement in the Era of Standards Based Reforms," *Sociology of Education* 83, no. 2 (2010): 160–82; and Joseph G. Altonji, Todd E. Elder, and Christopher R. Taber, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," Working Paper WP-00-17, Northwestern University, 2000, http://www.northwestern.edu/ipr/publications/papers/altonji.pdf.

29. Nicotera, Mendiburo, and Berends, p. 6.

30. When treated as a fixed effect, the magnitude of schoolwide mean family income is held constant across charter networks. This approach recognizes the superiority of schools that manage to perform unusually well despite low mean family income. If it were treated as a random effect, it would assume a different magnitude (i.e., coefficient) for each network, improperly ascribing important variation in network quality to the mean family income variable itself rather than to the coefficient of the charter network dummy variable.

31. Note that the process of combining standardized test scores across subjects in our sample violates one of the assumptions of OLS regression, the independence of observations, because a given student may have taken CST tests in more than one subject. The impact of this violation is to artificially inflate the precision of the

estimated coefficients. In practice, however, this impact is relatively minor due to the size of our dataset and the very high levels of precision it allows. For example, when the regressions are run using a single subject and the standardized scores are only averaged across grades (thus ensuring each student is tested only once), the vast majority of the charter network dummy coefficients that were statistically significant at accepted confidence levels using the entire dataset remain so for the single subject dataset. As a result, it is reasonable to combine results across subjects to obtain a comprehensive picture of each charter network's academic performance.

32. William Gould, "Clarification on Analytic Weights with Linear Regression," Stata.com, January 1999, http://www.stata.com/support/faqs/stat/crc36.html.

33. The College Board's income data rely on self-reporting by schools/students based on the completion of special forms as part of the AP process, and this self-reporting is highly uneven. It is not unusual for schools enrolling almost exclusively low-income students to nevertheless have none of their AP results flagged as "low-income." We therefore cannot rely on these data.

34. Though we do have the data on percentage of students qualifying for free and reduced price lunches for all California public schools, we lack detailed AP performance data for district public schools and hence would be unable to accurately estimate the parameter value for the schoolwide income control variable.

35. It would be possible, for instance, to pool the data for all eight SES groups into a single regression, but there is a concern with doing so: there may be significant differences in average educational capital in different SES households, due, for example, to cultural differences in parental expectations for homework. Arguably, the academic performance of children from households with higher educational capital is less affected by school quality (because these students learn more at home). If so, good schools with disproportionate numbers of students with high educational capital families will have their results biased downwards, while bad schools with many such families will have their results biased upwards. While this would not affect the overall conclusion of the study, it would bias the rankings. It is therefore desirable to find a single overall performance measure for each network that is less sensitive to SES enrollment breakdown. The unweighted mean of the SES results is one such measure, admittedly also imperfect.

36. "School Accountability Report Card School Year 2009-10," http://orb.sfusd.k12.ca.us/sarcs2/sarc-697.pdf.

37. Jay Mathews, "America's Best High Schools: The Elites," *Newsweek*, June 13, 2010, http://www.newsweek.com/2010/06/13/america-s-best-high-schools-in-a-different-class.html.

38. While a fixed-effects model is semantically the most appropriate, as it focuses on the relationship between enrollment and performance effect magnitude *within* each charter network, a random effects regression on these data yields the same results. Note that the z-scores whose absolute values were used in the time-series regression were computed from the pooled performance of all students within each school, controlling for percentage of students qualifying for free/reduced lunch. It was not possible to run the regressions separately by race/ethnicity and income-level (as was done in the body of the text, for the year 2010) because those demographic breakdowns were not reported for the entire span of years included in the time-series. This is not likely to have a noticeable effect on the findings of the time-series regression, however, because the fixed effects model looks only at the changes that occur *within* charter networks over time (not between networks) and the student demographics of charter networks tend not to vary dramatically from year to year.

STUDIES IN THE POLICY ANALYSIS SERIES

676. **Crony Capitalism and Social Engineering: The Case against Tax-Increment Financing** by Randal O'Toole (May 18, 2011)

675. **Leashing the Surveillance State: How to Reform Patriot Act Surveillance Authorities** by Julian Sanchez (May 16, 2011)

674. **Fannie Mae, Freddie Mac, and the Future of Federal Housing Finance Policy: A Study of Regulatory Privilege** by David Reiss (April 18, 2011)

673. **Bankrupt: Entitlements and the Federal Budget** by Michael D. Tanner (March 28, 2011)

672. **The Case for Gridlock** by Marcus E. Ethridge (January 27, 2011)

671. **Marriage against the State: Toward a New View of Civil Marriage** by Jason Kuznicki (January 12, 2011)

670. **Fixing Transit: The Case for Privatization** by Randal O'Toole (November 10, 2010)

669. **Congress Should Account for the Excess Burden of Taxation** by Christopher J. Conover (October 13, 2010)

668. **Fiscal Policy Report Card on America's Governors: 2010** by Chris Edwards (September 30, 2010)

667. **Budgetary Savings from Military Restraint** by Benjamin H. Friedman and Christopher Preble (September 23, 2010)

666. **Reforming Indigent Defense: How Free Market Principles Can Help to Fix a Broken System** by Stephen J. Schulhofer and David D. Friedman (September 1, 2010)

665. **The Inefficiency of Clearing Mandates** by Craig Pirrong (July 21, 2010)

664. **The DISCLOSE Act, Deliberation, and the First Amendment** by John Samples (June 28, 2010)

663. **Defining Success: The Case against Rail Transit** by Randal O'Toole (March 24, 2010)

Published by the Cato Institute, Policy Analysis is a regular series evaluating government policies and offering proposals for reform. Nothing in Policy Analysis should be construed as necessarily reflecting the views of the Cato Institute or as an attempt to aid or hinder the passage of any bill before Congress. Contact the Cato Institute for reprint permission. Additional copies of Policy Analysis are $6.00 each ($3.00 each for five or more). To order, or for a complete listing of available studies, write the Cato Institute, 1000 Massachusetts Ave., N.W., Washington, D.C. 20001 or call toll free 1-800-767-1241 (8:30-4:30 eastern time). Fax (202) 842-3490 • www.cato.org

CATO INSTITUTE