POLICY report

Goldwater Institute

No. 192 | March 15, 2004

Comparison of Traditional Public Schools and Charter Schools on Retention, School Switching, and Achievement Growth

by Lewis C. Solmon, Human Resources Policy Corporation, and Pete Goldschmidt, Center for the Study of Evaluation, UCLA

EXECUTIVE SUMMARY

There is a growing body of literature comparing the effectiveness of charter schools and traditional public schools. No consensus has yet been reached, but there are persistent concerns that performance differences might be due to "better" students attending charter schools. Researchers must therefore first determine whether charter school students and traditional public school students are substantially different before they can attribute any achievement differences to the education provided by these schools.

This study examines nearly 158,000 test scores of more than 60,000 Arizona students attending 873 charter and traditional public schools statewide over a three-year period. Its purpose is to determine the net effect of attending either type of school on Stanford Achievement Test, V9, (SAT-9) Reading achievement scores and total achievement growth over time.

Findings indicate that charter school students, on average, began with lower test scores than their traditional public school counterparts, and showed overall annual achievement growth roughly three points higher than their non-charter peers. Charter school students who completed the twelfth grade surpassed traditional public school students on SAT-9 Reading tests.

However, achievement growth varies by grade level. In the elementary grades, charter school students exhibited faster achievement growth than traditional public school students. Achievement growth in the middle grades was similar for both kinds of students, while high school achievement growth was higher for traditional public school students. One reason for this is that elementary charters are more likely to focus on academics, while middle and high school charters generally serve students who want vocational training, have been out of school, have learning or behavioral problems, or those who have been in the juvenile justice system.

There is growing public concern over early education and dropout prevention. This study finds that charter schools are providing solid early education that propels students through the advanced grades, even though they had lower achievement test scores on average than their traditional public school peers. Charters also reach at-risk students in the middle and high school grades who might otherwise have slipped though the cracks.[1]

GOLDWATER INSTITUTE

Comparison of Traditional Public Schools and Charter Schools on Retention, School Switching, and Achievement Growth

by Lewis C. Solmon, Human Resources Policy Corporation, and Pete Goldschmidt, Center for the Study of Evaluation, UCLA

Introduction

Much of the growing body of literature examining charter school effectiveness focuses on the methodological merits of each new analysis.[2] No consensus has been reached but there is concern that performance differences might be due to "better" students attending charter schools. Are students who actively seek out and transfer to alternative schools superior to students who remain in traditional public schools. Researchers refer to this possibility as *self-selection*, and some consider it a possible explanation of documented performance differences between charter and traditional public school students. Another possibility researchers offer is that charter schools "cream" off the top performing students from the traditional public schools. Researchers must therefore determine whether charter school students and traditional public school students are substantially different before they can attribute any achievement differences to the education that charter schools provide.

Researchers must also account for the fact that charter schools and traditional public schools may differ in several ways. For this reason, we examined three related issues. First, we examined the effect of different attendance patterns on student achievement growth, accounting for each student's initial achievement on SAT-9 Reading tests. Next, we examined whether the benefits of transferring to a different type of school—charter or traditional public—outweighed the disruption of switching schools. Finally, we investigated whether students in charter schools were more or less likely to repeat a grade than their traditional public school peers.

Essential for determining the impact of any charter school is to eliminate what researchers call *selection bias*. Thus, we must determine whether students in charter schools are like students not in charter schools, and whether any differences between them are due solely to attending a different type of school. If there is no selection bias, we would expect all students to exhibit the same achievement patterns in the absence of charter schools, although there would still be differences in achievement based on other student characteristics.

Ideally, researchers would randomly assign both teachers and students to traditional public schools and charter schools to remove any doubt concerning selection bias. Of course, this is infeasible. Ruling out selection bias is further complicated by the fact that many students do not simply attend either a charter or a traditional public school—many transfer back and forth.

This study uses longitudinal data

collected over a three-year period, which allow examination of each student's individual growth trajectory, and help to account for the potential confounding relationship between initial achievement levels and individual growth. Longitudinal data also account for differences in student demographics that are generally associated with performance differences. Further, because the most significant characteristics are included in the study models, no unobserved relationship remains between achievement and the errors, even when we compare varying attendance patterns of students who switched back and forth between charter and traditional public schools during the three-year study period.

An examination of test scores in any grade level must take into account the learning students have accumulated from previous grades, as well as students' home lives. We therefore assume a "value-added" model of measuring the impact of charter school attendance to make adjusted achievement comparisons based on various background factors. These data, such as demographic and personal characteristics, language proficiency, migrant status, and absenteeism, are available through school districts.

Thus, in classical experimental design, researchers overcome potential selection bias by using random selection. Since this approach is not feasible in this study, we use a value-added model with student information, as well as longitudinal data spanning three years, which allow for further bias controls.

Another important concern is accurately measuring achievement growth. Simple "pre/post test" approaches are inadequate because pretests and posttests may have errors, and errors within students are correlated. Also, this indirect method cannot determine whether there are influential effects over time. For this reason, we use a three-level hierarchical linear model design that directly measures achievement growth trajectories, beginning with students' initial achievement levels (what researchers call initial status), as well as the variation in initial achievement and achievement growth among students and schools over time. Appendices A-C provide a detailed description of the methodology we use in this study.

Our statewide sample uses data collected by the Arizona Department of Education. These data include several important student characteristics—as well as SAT-9 Reading achievement test scores—and represent the initial three years of extensive charter school existence in Arizona (1997-1998, 1998-1999, and 1999-2000). The student sample consists of 157,671 test scores, belonging to 62,207 students, attending 873 schools. Since many students transfer back and forth between traditional public schools and charter schools, these students make up one of eight attendance pattern groups.

We begin by examining several factors related to students' initial SAT-9

achievement and subsequent achievement growth, beginning with the 1997-1998 school year. Factors such as transferring schools, socioeconomic status, and being a non-primary English language speaker all affect student achievement and growth. Thus, we have made statistical adjustments for differences in student populations when comparing charter school and traditional public school achievement and achievement growth.

More students are designated special education among the charter school population than the traditional public school population. In addition, almost four times as many traditional public school students are designated as gifted than charter school students.

Examining these factors also allows examination of two additional ways charter schools and traditional public schools may differ: namely, the factors related to students transferring between schools and the factors related to students being held back a grade.

Factors Related to Student Achievement

Transferring schools is an important factor that can affect student achievement and achievement growth. Many wonder whether the potential benefits of switching schools outweigh the initial disruption. In our sample, there were eight possible attendance patterns for the three-year study period, which are summarized in Table 1.[3]

These patterns occur because students switch to and from charter schools (C) and traditional public schools (T). For example, a student who attended a public school for all three years would have an attendance pattern of TTT. A student who was in a public school during 1997-1998 but switched to charter schools thereafter would have an attendance pattern of TCC.

There is considerable variation among attendance patterns, especially when other student characteristics are taken into account. Nevertheless, the results presented in Table 2 demonstrate that no attendance pattern group consists wholly of traditionally underachieving students, such as migrants, those categorized as special education, and non-primary English speakers. Comparing students who attended traditional public schools all three years (TTT) to those who attended charter schools all three years (CCC), we see that both schools have similar student populations.

For example, 1.3 percent of the traditional public school students are migrants, compared to only 0.1 percent of charter school students. Yet more students are designated special education among the charter school population than the traditional public school population—5.8 percent compared to 3.8 percent. In addition, almost four times as many traditional public school students are designated as gifted than charter school students—7.6 percent compared to only 2.0 percent.

Although the group of students who remained in charter schools all three years had more special education students and fewer gifted students, the group of students who remained in traditional public schools all three years

Table 1. Attendance Patterns

Possible attendance patterns	School attended in year 1 (1997-1998)	School attended in year 2 (1998-1999)	School attended in year 3 (1999-2000)	Acronym for student group
1	charter	charter	charter	CCC
2	charter	charter	traditional public	CCT
3	charter	traditional public	traditional public	CTT
4	charter	traditional public	charter	CTC
5	traditional public	traditional public	traditional public	TTT
6	traditional public	traditional public	charter	TTC
7	traditional public	charter	traditional public	TCT
8	traditional public	charter	charter	TCC

had more minority students and non-primary English language speakers, in addition to more migrant students. Forty-two percent of the traditional public school students are minority students, compared to 35 percent of the charter school students.

Of particular interest is the number of school transfers that took place over the study period. Overall, 32 percent of students changed schools between 1997-1998, and about 40 percent of students changed schools between 1998-1999. There are also significant differences in the percentages of students who were held back a grade, when viewed by attendance pattern group. The results in Table 2 illustrate the complexity of the relationship between attendance patterns and the possibility of being held back. For example, students who attended charter schools in the first two years and then transferred to a traditional public school in the third year (CCT) were held back at a much lower rate between the second and third years, 1998-1999 and 1999-2000, the year they switched from charters to public schools.

This pattern reflects a more general pattern: the 1998-1999 percentages for being held back are similar regardless of the type of school students attend in the final year, 1999-2000. This similarity seems to indicate that students tend to switch schools to avoid repeating a grade, but that this strategy is more prevalent for students in charter schools.

Initial Student Achievement and Achievement Growth

Since achievement scores in SAT-9 Reading, Mathematics, and Language tend to be consistent across content areas, we focused on the results for SAT-9 Reading. There tends to be greater variation in reading scores by important student characteristics—such as race/ethnicity and language

Table 2. Student Demographics by Attendance Pattern (percent)

		Years in district	Absent	Migrant	Special education	Gifted	Minority	School change 1997-1998	School change 1998-1999	Non PLE	Held Back[a] 1997-1998	Held Back[a] 1998-1999	Percent of sample	N
CCC	Mean	1.02	2.04	0.1	5.8	2.0	35.0	27.3	44.2	8.3	5.6	5.0	4.1	2,531
	s.d.	0.72	0.70	2.8	23.4	14.1	47.7	44.5	49.7	26.0	23.0	21.8		
CCT	Mean	1.26	2.28	0.2	4.7	2.4	44.5	22.9	30.8	13.2	10.0	0.9	3.8	2,375
	s.d.	1.27	0.96	4.1	21.2	15.4	49.6	42.1	46.2	32.7	30.0	9.6		
CTT	Mean	2.49	2.18	0.8	4.7	4.3	50.6	99.6	68.9	10.9	7.8	2.3	2.5	1,545
	s.d.	1.98	0.89	8.8	21.1	20.4	49.9	6.2	46.3	30.0	26.9	14.9		
CTC	Mean	2.09	2.25	0.0	5.7	4.9	46.7	89.3	70.5	10.5	9.8	10.7	0.2	122
	s.d.	1.72	0.78	0.0	23.4	21.7	50.1	31.0	45.8	29.8	29.9	31.0		
TTT	Mean	1.31	2.11	1.2	3.8	7.6	41.9	23.2	36.8	12.1	1.8	1.3	80.1	49,806
	s.d.	1.20	0.64	10.7	19.1	26.5	49.2	42.2	48.2	31.8	13.2	11.3		
TTC	Mean	1.23	2.30	0.8	5.0	5.1	51.3	32.7	100.0	12.0	5.6	14.2	1.2	744
	s.d.	1.06	0.72	8.9	21.8	22.0	50.0	46.9	0.0	31.7	23.1	35.0		
TCT	Mean	1.11	2.25	0.2	2.9	1.5	47.9	98.1	26.2	9.4	12.2	1.8	4.8	2,988
	s.d.	0.92	0.78	4.1	16.8	12.2	49.8	13.7	44.0	28.2	32.7	13.2		
TCC	Mean	1.00	2.09	0.3	4.8	2.9	40.9	95.8	80.7	7.1	6.4	10.0	3.4	2,096
	s.d.	0.51	0.65	5.8	21.4	16.8	49.2	20.1	39.5	24.9	24.5	30.0		
Total	Mean	1.30	2.12	1.0	3.9	6.6	42.3	31.6	39.5	11.7	3.1	1.9	100	62,207
	s.d.	1.20	0.67	9.9	19.4	24.8	49.3	46.5	48.9	31.3	17.3	13.8		

Notes:

[a]This means that students who were held back in the 1997-1998 school year repeated the same grade in the 1998-1999 school year, and students who were held back in the 1998-1999 school year repeated the same grade in the 1999-2000 school year.

Non PLE represents non-primary English language speakers.

N is number.

s.d. is Standard Deviation.

Table 3. Mean SAT-9 Reading Achievement

Attendance pattern	1997-1998	1998-1999	1999-2000	Percent change 1997-1998 to 1999-2000
CCC	662	677	688	3.9
CCT	677	689	688	1.6
CTT	664	676	683	2.9
CTC	657	673	682	3.8
TTT	681	690	694	2.0
TTC	666	675	679	2.0
TCT	672	679	677	0.9
TCC	661	674	682	3.1

proficiency—than in mathematics scores.[4]

The overall raw SAT-9 Reading scale score results are presented in Table 3. There are relatively small differences between 1997-1998 initial student SAT-9 Reading achievement and achievement growth between 1997-1998 and 1999-2000 among attendance groups. These unadjusted differences suggest that no group is vastly different than any other.

Table 3 also demonstrates that while there was achievement growth for all attendance pattern groups, the level of growth varied. Examining mean scores alone, however, cannot account for these differences in growth. Using multilevel model results, we are able to examine more carefully the differences in initial student achievement and growth by accounting for differences in the composition of each attendance pattern group. These results have additional advantages as well.

With multilevel model results we can include factors associated with school type selection, and can account for the fact that student growth varies within and between schools. Model-based results also avoid penalizing groups that have higher proportions of traditionally low-scoring students because they adjust scores for differences in group compositions.[5]

The results in Table 4 indicate that the mean initial achievement for students in 1997-1998 was 661.[6] Since both the initial student achievement and the growth rates can vary randomly among students, the model separates the total variation in both initial achievement and growth into within-students, among-students, and among-school components.[7]

The average achievement growth rate is approximately 11.6 (scale score) points per year. This rate equates to an annual effect-size gain of approximately 0.26, roughly a quarter of a standard

deviation gain per year.

However, the results in Table 4 also indicate that both the initial student achievement and the annual achievement growth vary significantly among students and schools.[8] Comparing a student who is one standard deviation below average on initial achievement to a student in the same school who is one standard deviation above average on initial achievement yields a gap of about 57 points.[9] That is, we would expect the below-average student to have an initial status of about 633, while we would expect the above-average student to have an initial achievement of about 689.

There is substantial variability among students in reading achievement growth as well. A student demonstrating growth one standard deviation above average is expected to have an annual gain of about 14.8 points, while a student demonstrating growth one standard deviation below average is expected to have an annual gain of about 8.4 points. The unconditional model results in Table 4 also indicate that the variation among students accounts for about 39 percent of the total variation in initial achievement, and about 18 percent of the total variation in growth.

Table 4. Random Coefficient Model—SAT-9 Reading

Fixed effects	Coefficient	SE (standard error)	*p*[a] (probability)
Average 1997-1998 initial status	661.3	1.25	0.000
Average change (growth)	11.6	.259	0.000

Random effects	Variance component	*df* (degrees of freedom)	*p*[a]
Level 1			
Within-student (temporal variation—residual)	209.4		
Level 2			
Within-school variation—initial status	802.9	39,907	0.000
Within-school variation—growth	10.5	39,907	0.000
Level 3			
Between-school variation—initial status	1,252.8	848	0.000
Between-school variation—growth	47.1	848	0.000
Variation between Students			
In 1997-1998 initial status	39.0%		
In growth	18.2%		

These results, however, should only be considered as a baseline since important variables, such as grade level, are not included in this model. Grade level is an important variable to consider because scale scores are designed to increase with each successive grade. Since an average student will have a higher scale score in fourth grade than third grade, students' 1997-1998 initial achievement will depend upon which grade level they were in. Moreover, we include grade level as an independent variable describing student growth—not as a time-varying covariate—on the assumption that student achievement growth may differ depending on the grade a student is in at the beginning of the period under study.[10]

Initial Achievement and Achievement Growth by Grade Level

The preliminary analyses tested several potential factors affecting initial achievement and achievement growth. We now focus on the variation attributable to students. Tables 5 and 6 present the results of the full model we used to examine the effects of attendance patterns on student SAT-9 achievement and achievement growth, respectively.[11] They show only the grade-by-attendance-pattern interaction upon achievement and growth, as other interactions by attendance pattern were not significant.[12] The results indicate that grade has a significant impact on both initial status and reading achievement growth ($p<.05$), and that the effect of grade varies by attendance pattern.

The results in Table 5 show that student demographic characteristics and their relationship to achievement growth are consistent with previous research. The treatment groups have a statistically significant relationship to both initial status and achievement growth, and the effects of the treatment groups are dependent upon grade.

The effects of student demographic characteristics on initial achievement in 1997-1998 are consistent with expectations. For example, student absenteeism has a negative effect on 1997-1998 reading achievement. Further, students classified as special education, minority, migrant, or non-primary-language-as-English are all negatively associated with 1997-1998 reading achievement.

The achievement gap between students whose primary language is not English and students whose primary language is English is about 13 points, which is an effect size of about 0.28. The student demographic characteristics are important because they statistically adjust for differences in initial achievement. They also put growth estimates into context. That is, growth estimates taken together with initial achievement estimates identify whether performance gaps exist and whether those gaps are growing or narrowing over time. This, of course, is also true for the treatment group variables of interest.

Table 5. Final Estimation of Fixed Effects on Initial Status—Linear Model

Fixed effect	Coefficient	s.e.	p-value
Initial Status	642.5	1.08	0.00
Years in district	-1.8	0.12	0.00
Absent	-2.7	0.18	0.00
Migrant	-4.2	1.30	0.00
Special education	-32.0	0.59	0.00
Gifted	28.3	0.48	0.00
Effect of attendance patterns on initial status			
CCC	-11.8	2.51	0.00
CCT	-17.5	2.83	0.00
CTT	-22.5	2.92	0.00
CTC	-16.8	6.03	0.01
TTC	2.3	2.61	0.37
TCT	-3.8	1.48	0.01
TCC	-4.6	1.30	0.00
Interaction of grade and attendance pattern			
CCC	3.6	0.35	0.00
CCT	3.1	0.36	0.00
CTT	3.8	0.39	0.00
CTC	3.4	1.27	0.01
TTC	-1.3	0.53	0.01
TCT	-0.2	0.27	0.45
TCC	1.6	0.31	0.00
Grade	8.7	0.15	0.00
Minority	-10.7	0.29	0.00
English not primary language	-12.8	0.44	0.00
School change total	2.1	0.40	0.00
Grade school change interaction	0.0	0.08	0.89
Held back	-11.5	0.55	0.00

Table 6. Final Estimation of Fixed Effects on Growth—Linear Model

Fixed effect	Coefficient	s.e.	p-value
Annual Growth	19.1	0.31	0.00
Years in district	0.2	0.05	0.00
Absent	-0.6	0.09	0.00
Migrant	-1.3	0.62	0.03
Special education	0.6	0.26	0.02
Gifted	-0.6	0.20	0.00
Effect of attendance patterns on growth			
CCC	3.3	0.66	0.00
CCT	5.6	1.01	0.00
CTT	7.6	0.93	0.00
CTC	5.2	2.18	0.00
TTC	-2.8	1.02	0.01
TCT	-1.2	0.68	0.08
TCC	0.6	0.49	0.19
Interaction of grade and attendance pattern			
CCC	-0.7	0.13	0.00
CCT	-0.6	0.17	0.00
CTT	-1.3	0.17	0.00
CTC	-0.9	0.49	0.06
TTC	0.3	0.21	0.20
TCT	0.1	0.14	0.30
TCC	-0.5	0.12	0.00
Grade	-1.7	0.06	0.00
Minority	0.1	0.13	0.54
English not primary language	1.0	0.20	0.00
School change total	0.3	0.16	0.03
Grade school change interaction	-0.2	0.03	0.00
Held back	0.6	0.25	0.02

In general, students in the CCC group are estimated to exhibit achievement growth about 3.3 points faster per year than TTT students. Given these estimates, we would expect students in the CCC group to exhibit faster achievement growth than students in the TTT group until about eighth grade.

Achievement growth results related to initial achievement results indicate that differences among students—and correspondingly, among attendance pattern groups—are likely due to pre-existing differences among students, including selection. We would be more concerned with differences in initial achievement if they were related to growth, since growth is our primary focus. However, there seems to be little association between initial student achievement in 1997-1998 and its growth over the three-year period because the student-level correlation between initial achievement and achievement growth is -11.

Table 6 shows the effect grade level has on students who attended traditional public schools all three years (TTT). The expected annual achievement gain is approximately 19.1 points, but at each successive grade level this gain declines by about 1.7 points. This means that traditional public school students starting in third grade have an estimated achievement growth about 10 points greater than that of high school freshmen.

The results in Table 6 also indicate the expected achievement gains for each of the eight attendance pattern groups. Using the students who attended traditional public schools all three years (TTT) as the comparison group, all of the remaining attendance pattern groups have growth rates that are significantly different, except for two groups of students who started out in traditional public schools—namely the TCT and TCC attendance groups. Table 7 details the difference in estimated achievement growth between the TTT group and the remaining attendance groups.

We next highlight the difference in estimated achievement growth by comparing the results of students who attended charter schools all three years, (CCC), to the TTT group.

In general, students in the CCC

Table 7. Difference in Estimated Achievement Growth

Attendance group	Difference from TTT group in estimated achievement growth
CCC	3.3
CCT	5.6
CTT	7.6
CTC	5.2
TTC	-2.8

group are estimated to exhibit achievement growth about 3.3 points faster per year than TTT students. However, this effect varies by grade, with achievement growth diminishing about 0.7 points per grade. Given these estimates, we would expect students in the CCC group to exhibit faster achievement growth than students in the TTT group until about eighth grade.[13]

The results for the student background characteristics are generally smaller in effect size, and as noted, are not affected by attendance patterns. Yet student characteristics do have a statistically significant relationship to achievement growth. For example, minority student achievement growth is not significantly different from non-minority student growth. Nevertheless, since minority students were about 10 points behind non-minority students in achievement during 1997-1998, they are not closing the existing achievement gap. The 1997-1998 achievement gap for non-primary English speakers is diminishing at an estimated growth rate of about one point greater per year than for primary English-speaking students. The results for migrant students, however, are less positive because their estimated achievement growth rate is about 1.3 points slower than non-migrant students. This further exacerbates the 1997-1998 reading achievement gap.

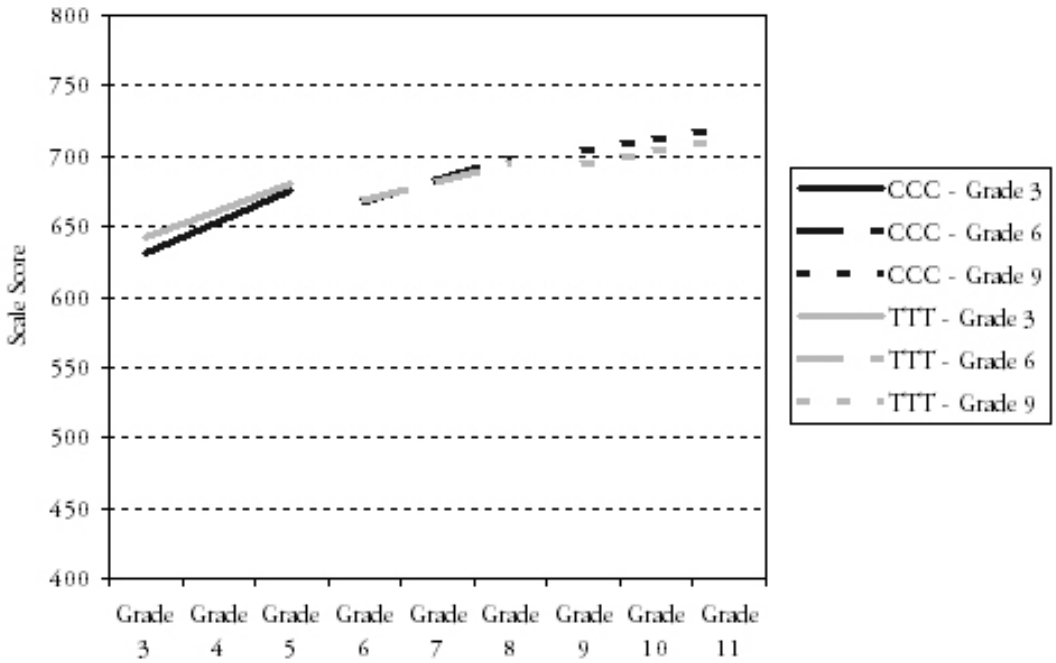
Students who transfer schools demonstrate slightly faster achievement growth of about 0.3 points per year. Students who have been held back also demonstrate achievement growth 0.6 points faster per year (Table 6) than students not held back. Nevertheless, students who are held back start with an 11.5-point achievement disadvantage (Table 5), hence it would still take more than 19 years for held-back students to catch up.[14]

Figure 1 presents results for students who are similar in all aspects except the attendance pattern group to which they belong. These results, based on the models presented in Tables 5 and 6, highlight how several factors combine to produce an intricate picture of the three-year achievement patterns for students attending only charter schools and students attending only traditional public schools.

To keep the figure interpretable, we limit the comparison to the groups of students who attended charter schools (CCC) or traditional public schools (TTT) all three years. Figure 1 demonstrates that SAT-9 Reading achievement growth varies by both attendance pattern and grade. Consistent with the raw means (and the initial 1997-1998 achievement levels), third grade students in the CCC group start below students in the TTT group, but exhibit faster growth. This growth results in CCC students exhibiting about equal achievement levels in sixth grade, and higher achievement levels in ninth grade. Achievement growth in middle school remains slightly higher for CCC students, while high school achievement growth is higher for TTT students.

Consistent with the raw means (and the initial 1997-1998 achievement levels), third grade students in the CCC group start below students in the TTT group, but exhibit faster growth. This growth results in CCC students exhibiting about equal achievement levels in sixth grade, and higher achievement levels in ninth grade.

Figure 1. Comparison of SAT-9 Reading Achievement Change Between 1997-1998 and 1999-2000 by CCC and TTT Attendance Patterns



Achievement Growth and the Timing of Student Transfers

The previous analysis used a three-level hierarchical linear model. To explore more fully the relationship among attendance patterns, achievement growth, school change, and potential assimilation effects, we used a piecewise growth model that explicitly accounts for a step-like growth pattern.[15] This model also implicitly links the timing of charter/traditional switching and school changes within an attendance pattern.[16] The results are displayed in Tables 8, 9, and 10.

The results for initial student achievement are not substantively different from the previous linear model results. Table 8 demonstrates again that performance gaps exist according to student characteristics and attendance pattern groups. Traditionally disadvantaged students perform significantly lower than their non-disadvantaged classmates. However, we again focus on achievement growth—now partitioned into growth between the first and second year and the first and third year. In this way, the attendance-pattern, school-change, and held-back variables can be linked specifically to either the first or the second year.

For example, the school-change variable reflects whether a student transferred schools between 1997-1998 and 1998-1999. Hence, a student's

reading achievement score will likely be affected by this school change. Further, a student in the CTT group, for instance, must have switched school type between 1997-1998 and 1998-1999, and this transfer will potentially affect the 1998-1999 SAT-9 achievement score. The interpretations for students who are held back are similar to those for students who change schools, since a student who is held back in 1997-1998 repeats the grade in 1998-1999.

The achievement growth results from the piecewise model in Tables 9 and 10 are substantively similar to those in the previous linear model. The results for the attendance patterns are also similar, but they warrant further explanation.

The results for the first year's growth indicate that there are differential effects of attendance patterns. The mean change from 1997 to 1998 was approximately 22 points. Once again, students in all but the TTC and the TCT attendance patterns demonstrated statistically significant differences in achievement growth from the TTT group. Again, these differences were affected by grade. For example, a third-grade student in the CCC group is expected to demonstrate 6.4 points greater growth than a TTT third-grade student. As shown previously, the benefit of faster achievement growth accruing to CCC students dissipates in middle school, by about seventh grade in the piecewise model, and by about eighth grade in the previous linear model.

The piecewise model also supplies additional information about students being transferred and held back. The achievement growth of students who transferred schools between the 1997-1998 and 1998-1999 school years is expected to be about 3.3 points lower than for students who did not change schools. Students who are held back are expected to demonstrate a slightly faster achievement growth of about 2 points. However, it is important to bear in mind that students who had been held back had lower average achievement (by approximately 12 points) in 1997-1998. For this reason, the faster growth represents some catching up, but it does not close the achievement gap.

Table 10 represents the second portion of the growth model, which presents estimates for the total achievement growth gain from 1997 to 1999 of roughly 39 points. This figure is consistent with the linear model, which estimated a mean achievement growth rate of about 19 points per year (hence a mean gain of 38 points by 1999). The piecewise model, however, demonstrates that there was a greater achievement growth gain between 1997 and 1998 (22 points) than between the latter two years (16 points).

The results for the second portion of the growth model are generally consistent with those for the first year results, with some notable differences. The results in Table 10 demonstrate that third-grade students in the CCC group gain more than third-grade students in the TTT group. Although the

achievement growth gain is generally lower between the latter two years, the CCC students are expected to gain about 5.2 points more over the remaining two-year period.

As with our previous results, grade level is an important element that must be considered. Not only is grade level important in moderating the effect of attendance patterns, but it also moderates the effect of students transferring to different schools. Overall, there is a positive gain of 2.6 points

Table 8. Final Estimation of Fixed Effects for Initial Status—Piecewise Model

Fixed Effect	Coefficient	s.e.	p-value
Initial Status	641.5	2.7	0.00
Years in district	-1.9	0.3	0.00
Absent	-2.6	0.2	0.00
Migrant	-4.3	1.1	0.00
Special education	-32.1	0.9	0.00
Gifted	28.2	1.0	0.00
Effect of attendance patterns on initial status			
CCC	-13.0	5.5	0.02
CCT	-17.7	5.3	0.00
CTT	-24.6	4.9	0.00
CTC	-17.6	8.4	0.04
TTC	0.7	3.1	0.83
TCT	-3.4	2.0	0.08
TCC	-5.8	2.4	0.02
Interaction of grade and attendance pattern			
CCC	4.0	1.2	0.00
CCT	3.2	1.0	0.00
CTT	4.1	1.0	0.00
CTC	3.4	1.7	0.04
TTC	-1.0	0.7	0.15
TCT	-0.2	0.4	0.70
TCC	1.8	0.5	0.00
Grade	8.7	0.6	0.00
Minority	-10.6	0.5	0.00
English not primary language	-12.9	0.6	0.00
School change total	3.0	1.1	0.01
Grade by school change interaction	-0.1	0.2	0.56
Held back total	-11.8	0.8	0.00

when students change schools, but this gain is reduced by each subsequent grade level. Hence, students in third grade benefit by about 2.6 additional points in achievement growth when switching schools. Switching schools after eighth grade, however, has a negative effect. This implies that reasons for switching schools are most likely different for younger and older students.

Table 9. Final Estimation of Fixed Effects,1997-1998 to 1998-1999—Piecewise Model

Fixed Effect	Coefficient	s.e.	p-value
Growth from 1997-98 to 1998-99	22.4	0.9	0.00
Years in district	0.4	0.2	0.02
Absent	-0.9	0.1	0.00
Migrant	-1.2	1.3	0.36
Special education	0.8	0.6	0.14
Gifted	-0.6	0.4	0.18
Effect of attendance patterns on growth			
CCC	6.4	1.5	0.00
CCT	5.5	1.8	0.00
CTT	13.3	2.5	0.00
CTC	8.3	4.6	0.07
TTC	1.3	2.1	0.54
TCT	-2.0	1.3	0.13
TCC	4.2	1.1	0.00
Interaction of grade and attendance pattern			
CCC	-1.6	0.3	0.00
CCT	-0.7	0.3	0.04
CTT	-2.1	0.4	0.00
CTC	-1.0	1.1	0.34
TTC	-0.6	0.4	0.17
TCT	0.3	0.2	0.29
TCC	-1.2	0.3	0.00
Grade	-1.8	0.2	0.00
Minority	0.0	0.2	0.89
School change 1997-1998	-3.3	0.7	0.00
English not primary language	1.3	0.3	0.00
Held back 1997-1998	2.0	0.6	0.00
Grade by school change interaction	0.0	0.1	0.63

Table 10. Final Estimation of Fixed Effects,1997-1998 to 1999-2000—Piecewise Model

Fixed Effect	Coefficient	s.e.	p-value
Growth from 1997-1998 to 1999-2000	38.3	1.0	0.00
Years in district	0.6	0.2	0.00
Absent	-1.1	0.2	0.00
Migrant	-3.0	1.3	0.02
Special education	1.1	0.7	0.12
Gifted	-1.2	0.5	0.02
Effect of attendance patterns on growth			
CCC	5.2	1.8	0.01
CCT	9.7	2.5	0.00
CTT	13.3	2.7	0.00
CTC	8.7	4.2	0.04
TTC	-6.9	1.9	0.00
TCT	-2.5	1.5	0.10
TCC	0.8	1.2	0.48
Interaction of grade and attendance pattern			
CCC	-0.8	0.4	0.06
CCT	-1.2	0.6	0.04
CTT	-2.2	0.5	0.00
CTC	-1.3	1.1	0.25
TTC	0.8	0.4	0.08
TCT	0.6	0.3	0.10
TCC	-0.8	0.3	0.02
Grade	-3.8	0.2	0.00
Minority	0.2	0.3	0.56
School change 1998-1999	2.6	0.8	0.00
English not primary language	1.9	0.4	0.00
Held back 1998-1999	2.9	0.9	0.00
Grade by school change interaction	-0.5	0.1	0.00

Why Charter School Achievement Is Faster in the Early Grades

The previous findings suggest a correlation between attendance at a charter school and higher achievement and achievement growth in the early grades, even though charter school students generally have lower initial achievement levels. Yet, this growth slows in middle school.

It is likely that charter schools have a greater impact on achievement growth at the elementary level because the nature, mission, and type of students in charter schools differ at the elementary, middle, and high school levels. Using data from the 2000-2001 school year, we divided schools into three groups: elementary or elementary/middle schools; K-12 schools; and middle, high, or middle/high schools.

We then sorted charter schools into three categories. We labeled the first category "academic," which included general, college-bound, accelerated-academically, back-to-basics, Montessori, and performing arts schools. We labeled the second category "at-risk," which included at-risk, special education, bi-literate, alternative, and alternative/technical schools. "Vocational" was the third category, and included technical, career, vocational, and agricultural schools.

Of the 406 charter schools for which we had information on type and target population, 44.6 percent of elementary or elementary/middle were in the academic category, compared to 34.2 percent of the K-12 schools, and 26.4 percent of the middle, high, or middle/high schools. However, 10.3 percent of the elementary group was in the at-risk category, as was 23.7 percent of the K-12 schools, and 47.7 percent of the middle, high, and middle/high schools. The remaining schools were in the vocational category: 1.7 percent elementary, 2.6 percent K-12, and 8.8 percent middle or high. Although we do not have comparable data for the traditional public schools, it seems unlikely that almost half of them would be focused on at-risk students.[17] Moreover, most traditional high schools have no focus at all, but include the full cross section of students, from academic to at-risk.

These patterns held when we looked at the number of students in each type of charter school at each grade level, providing some explanation as to why charter school gains for students at the elementary level are greater than those in higher grades. Elementary charter schools are substantially more likely to focus on academics, while charters with higher grade levels are more likely to be serving students with serious learning or behavioral problems, those who have been out of school, those who have been in the juvenile justice system, or those seeking vocational training.

The previous findings suggest a correlation between attendance at a charter school and higher achievement and achievement growth in the early grades, even though charter school students generally have lower initial achievement levels. Yet, this growth slows in middle school.

These data also help explain varied achievement growth among attendance groups of students who switched schools, compared to students who remained in traditional public schools over the three-year study period.

The CCT group—students who attended charter schools during the first two years then switched to traditional public schools—outperforms the TTT group for the longest period of time, and the test score growth rate of the CCT group does not decline to the TTT rate until twelfth grade. Perhaps the CCT students are good at matching their needs to their schools; on average, they stay in charters while they are effective (until about the eighth grade), and then switch to traditional public schools for middle and high school—the grade levels at which traditional public schools have higher achievement growth rates than charter schools.

For the CTT group—students who attended charter schools during the first year then switched to traditional public schools for the final two years—test-score growth falls to the TTT level by the end of eighth grade (compared to twelfth grade in the CCT group). Even though the CTT group started with a higher average growth rate, once these students reach middle school their growth declines more quickly than the CCT group. Specifically, achievement growth for the CCT group grows 5.6 points per grade, slowing down only -0.6 points each subsequent grade level. Thus it would take 6.33 years to fall to the TTT level. The CTT group has a slightly higher average growth than the CCT group, 7.6 points. However, its growth slows down more dramatically—by -1.3 points each subsequent grade level. Thus, it would take only 2.9 years for the CTT group to fall to the TTT achievement level.

Because charter schools focus more on academics in the earlier years, charter students do better. When charter schools begin serving more at-risk students in middle/high school years—with less focus on academics—we would expect to see the advantage for charter schools dissipate. That is, we expect to see the growth of charter students slow down. Since the CTT students leave charter schools sooner, their growth declines faster.

We can speculate about other reasons why these patterns of starting in charters and then moving back to traditional public schools after one or two years demonstrates the most rapid growth. Early attendance at a charter school might provide the basis for subsequent achievement in traditional public schools. For example, if students learn to read in charters, that prepares them for later studies. This could also explain why the TCC group does poorly, and never demonstrates growth significantly different from the TTT group. Their poor preparation in the early grades is not overcome once they switch to charter schools at the middle/high school levels because at higher grades charters are less likely to focus on academics.

If attendance at charter schools in the early grades prepares students better academically, this could explain the advantage of CCT and CTT over TCC, and even more so over TCT and TTC, which are consistently below TTT. The CCC group starts out with faster achievement growth than the TTT group, but falls back to the TTT level by middle school. Finally, among constant switchers, the CTC group does much better than the TCT group. This indicates that TCT students are not doing well in traditional public schools, and switch to charter schools, which are not the panacea they had hoped for. So these students switch back to traditional public schools. Unfortunately, all that switching does not help these TCT students. On the other hand, the CTC students start in charter schools, think they could do better in traditional public schools but do not, and so they return to charter schools.

Thus, charter school students start with lower scores. They have patterns of faster achievement growth in elementary school, growth patterns roughly equivalent to students in traditional middle schools, and slower growth in high school. Even so, charter school students who completed the twelfth grade surpassed traditional public school students.

The Effect of Switching Schools on Achievement Growth

Given that school change is related to achievement growth, we examine whether attendance patterns—after accounting for other student background information—have any impact on the probability that a student will switch schools.[18] We also examine the factors associated with changing schools. As noted, there are likely to be effects on achievement associated with switching schools. The question arises as to whether the benefit of a better match between student to school type outweighs the disruption associated with switching schools.[19]

Thus, charter school students start with lower scores. They have patterns of faster achievement growth in elementary school, growth patterns roughly equivalent to students in traditional middle schools, and slower growth in high school. Even so, charter school students who completed the twelfth grade surpassed traditional public school students.

The answer is particularly ambiguous as these data represent the initial period of charter school existence. And as noted previously, there was a significant amount of transferring by students among schools—not only among attendance pattern groups, but also within attendance pattern groups. Approximately half of the sample changed schools at least once between 1997 and 1999. About 24 percent of students changed schools in both years. Of the students who changed schools in 1997-1998, 75 percent changed schools again during 1998-1999. This pattern seems to indicate that parents were attempting to find the most appropriate academic match for their children.

Table 11 presents the results for switching schools during 1997-1998 and 1998-1999, respectively. It also shows results for student characteristics and the effects of charter school attendance. The results for school change in 1997-1998 indicate that for students scoring at the mean of SAT-9

Table 11. Fixed-Effect of Logistic Multilevel Model for School Change[a]

	1997-1998 through 19998-1999				1998-1999 through 1999-2000			
	TPS		Charter		TPS		Charter	
Student characteristic	Log-odds	Percent[a]	Log-odds	Percent[a]	Log-odds	Percent[a]	Log-odds	Percent[a]
Base	0.023	50.6	**-0.439**	39.7	**-0.187**	45.3	**-0.607**	31.1
	(0.061)		(0.150)		(0.068)		(0.154)	
SAT-9 reading[b]	**0.002**	4.0	**-0.005**	5.4	**0.001**	7.0	**-0.002**	16.5
	(0.000)		(0.001)		(0.000)		(0.001)	
Grade 6 vs. grade 8	**0.673**	-29.9	**-0.233**	-10.4	**0.430**	-15.6	**-0.399**	17.3
Grade 9 vs. grade 9	**0.673**	15.6	**-0.233**	21.1	**0.430**	15.3	**-0.399**	19.6
	(0.035)		(0.063)		(0.027)		(0.050)	
Absent [b]	**0.069**	2.3	0.013	2.2	**-0.061**	-2.0	0.002	-1.7
	(0.013)		(0.034)		(0.012)		(0.029)	
Migrant	**-0.120**	-3.0		-2.8	**0.139**	3.5		5.9
	(0.042)				(0.051)			
Special education	0.046		-0.213		0.039		0.092	
	(0.041)		(0.132)		(0.040)		(0.111)	
Gifted	**-0.277**	-6.9	**0.994**	17.7	0.043		0.286	
	(0.048)		(0.215)		(0.040)		(0.171)	
Minority	-0.012				0.012			
	(0.016)				(0.016)			
Non-primary lang. English	**-0.132**	-3.3		-3.1	-0.031			
	(0.028)				(0.027)			
Held back 1997-1998	**0.593**	14.3	-0.090	14.7	**-0.586**	-13.8	**0.428**	-3.3
	(0.115)		(0.189)		(0.068)		(0.126)	
Held back 1998-1999	-		-		**0.840**	20.4	**0.389**	29.6
	-		-		(0.085)		(0.191)	
School change 1997-1998	-		-		**0.710**	17.4	0.198	16.8
					(0.052)		(0.125)	

Notes: Coefficients in bold are statistically significant (at p<.05). The results present the student effects in the first column and the effect of charter school attendance (cross-level interactions) in the subsequent columns. The statistically significant marginal effect associated with each student characteristic is presented in percentage points. [a]The base percent is the probability of switching schools for an "average" student. The remaining percents in this column are the percentage point difference between a student who is described by the variable and a student who is not. (Except for variables denoted by [2]. [b]Results in the percent column indicate the percentage point difference between a student who is one standard deviation above the mean on this variable and a student who is one standard deviation below the mean on this variable.

achievement, absenteeism and grade increase the likelihood that they will change schools.[20] The mean rate of change for an average student is estimated to be about 50.6 percent in traditional public schools, and about 39.7 percent in charter schools. Migrant, gifted, and students whose primary language is not English are less likely to change schools, by three percent, seven percent, and three percent, respectively. However, gifted students attending charter schools have much higher school-switching rates in charter schools. They are about 18 percent more likely to switch than non-gifted students attending charter schools.[21]

Students who have been held back are also significantly more likely to change schools. In fact, they are 14 percent more likely to change schools as students who have not been held back. In addition, students have higher probabilities of switching schools as they move up in grade level.

The results for school change in 1998-1999 are similar to those for 1997-1998. However, the 1998-1999 model includes information from the prior year. The overall estimated probability that a student will switch schools is reduced from the previous year. Students in traditional public schools have an estimated probability of switching schools of about 45 percent, while for charter school students it is about 31 percent. This result would be consistent with the notion that students and parents are looking for the best school match. Over time, we would expect to observe less switching by students seeking a better school match to meet their needs, and more switching due to academic or discipline problems.

Turning to results that differ from 1997-98, migrant students were more likely to change schools in 1998-1999 than non-migrant students. The student characteristics that affect being held back are consistent with those from 1997-1998. That is, if students were held back in the most recent year, they were significantly more likely to change schools—20 percent more likely in traditional public schools, and 30 percent more likely in charter schools. However, if students had been held back a year prior—which increased the likelihood of switching schools in 1997-1998—and had not been held back again, they were significantly less likely to change schools. This implies a pattern of being held back, switching schools, and then staying. Even so, students who changed schools in 1997-1998 are about 17.4 percent more likely to switch schools again in traditional public schools, and 16.8 percent more likely to switch schools again in charter schools.[21]

The variables with the most impact at charter schools are grade level and being held back. Again, the effect of grade level is mitigated if a student attends a charter school. Charter school students are significantly more likely to switch schools if they are held back than traditional school students. Students attending charter schools who were held back in 1998-1999 were about 1.5 times

A primary concern is that better students account for superior charter school performance. To the contrary, our achievement results indicate that charter students generally started off with lower achievement than their traditional public school counterparts. Despite this initial deficit, charter school students showed overall annual achievement growth roughly three points higher than their non-charter peers.

as likely to change schools as students attending traditional public schools. Furthermore, charter school students who were held back were almost twice as likely to switch schools than their charter school classmates who were not held back. Since being held back is one of the strongest predictors of switching schools, we next examine factors associated with being held back.

Factors Associated with Being Held Back

Table 12 presents factors related to being held back, pertaining to both the 1997-1998 and 1998-1999 school years.[22] The results indicate that in most instances, measured student characteristics are statistically related to students being held back. Also, in almost all instances in 1997-1998 (except migrant and students whose primary language was English), traditional public school and charter school students have different probabilities of being held back. (There are fewer differences in 1998-1999.) The base rate for traditional public school students in 1997-1998 is 2.6 percent, while for charter school students it is 12.2 percent.

Not surprisingly, the better students performed on the SAT-9 Reading exam, the less likely it was that they would be held back. Also, the longer students had attended their schools, the less likely they were to be held back. Minority students in traditional public schools were 1.1 percent more likely to be held back, compared to their non-minority classmates. In charter schools, this difference was only 0.2 percent.

The results for 1998-1999 indicate that students were less likely to be held back between the second and third year of data, especially students in charter schools. This result further supports the theory that students and parents search for a good school match, and that some settling down eventually takes place. Nevertheless, students who attended charter schools were, on average, 2.6 times as likely to be held back as students attending traditional schools in 1998-1999.

Overall, the reduction in the estimated probabilities for being held back in 1998-1999 was much greater for charter school students than for those attending traditional public schools. Special education students are more likely to be held back in both traditional public schools and charter schools, but this phenomenon is significantly more likely in charter schools. The results associated with student characteristics for 1998-1999 are similar to those for 1997-1998. The one significant change is that migrant students are significantly more likely to be held back. In general, the results for 1998-1999 indicate smaller differences by student characteristics, and smaller differences between traditional public school and charter school students.

Table 12. Fixed Effects from Multilevel Logistic Model for Held Back[a] Students

		1997-1998			1998-1999		
Fixed Effect		Coefficient	s.e.	Percent[a]	Coefficient	s.e.	Percent[a]
Base	TPS	**-3.63**	0.07	2.6	**-3.85**	0.06	2.1
	Charter	**1.66**	0.15	12.2	**0.52**	0.11	3.4
SAT-9 reading[b]	TPS	**-0.02**	0.00	-2.0	**-0.01**	0.00	-1.1
	Charter	**0.02**	0.00	-1.2	**0.01**	0.00	-0.3
Grade	TPS	**0.17**	0.03	0.5	**0.12**	0.02	0.3
	Charter	**-0.14**	0.04	0.3	0.04	0.03	
Years in district[b]	TPS	**-0.24**	0.04	-1.1	0.01	0.03	
	Charter	**0.18**	0.05	-1.4	-0.03	0.05	
Absent[b]	TPS	**0.16**	0.02	0.6	**0.29**	0.02	1.0
	Charter	**-0.13**	0.05	0.4	**-0.26**	0.06	0.2
Migrant	TPS	**-0.50**	0.18	-1.0	**0.52**	0.15	1.4
	Charter	0.56	0.78	0.32	0.72		
Special education	TPS	-0.02	0.08	0.26	0.10	0.6	
	Charter	**0.33**	0.15	3.7	-0.21	0.27	
Gifted	TPS	**-0.17**	0.09	-0.4	**-0.55**	0.12	-0.9
	Charter	**0.58**	0.18	5.1	0.25	0.41	
Minority	TPS	**0.36**	0.05	1.1	**0.38**	0.05	0.9
	Charter	**-0.34**	0.11	0.2	-0.22	0.13	
Non-PLE	TPS	-0.10	0.06	-0.03	0.07		
	Charter	0.12	0.17		**-0.77**	0.25	-1.9

Notes: Coefficients in bold are statistically significant (at p<.05).

[a]The base percent is the probability of switching schools for an "average" student. The remaining percents in this column are the percentage point difference between a student who is described by the variable and a student who is not. (Except for variables denoted by [2].)

[b]Results in the percent column indicate the percentage point difference between a student who is one standard deviation above the mean on this variable and a student who is one standard deviation below the mean on this variable.

Conclusion

This study examined the net effect of attending Arizona charter or traditional public schools on SAT-9 Reading achievement and achievement growth. The results cannot answer unequivocally whether charter school attendance is superior to traditional public school attendance. However, they do shed light on some of the most common concerns about charter school performance.

A primary concern is that better students account for superior charter school performance. To the contrary, our achievement results indicate that charter students generally started off with lower achievement than their traditional public school counterparts. Despite this initial deficit, charter school students showed overall annual achievement growth roughly three points higher than their non-charter peers, whom charter students surpassed on SAT-9 Reading tests by the end of twelfth grade.

However, achievement growth varies by grade level. In the elementary grades, charter school students exhibited faster growth than traditional public school students. Achievement growth in the middle grades was similar for both kinds of students, while high school achievement growth was higher for traditional public school students. The likely reason for this is that the nature, mission, and type of students attending charter schools differed at the elementary, middle, and high school levels. Elementary charters typically focus on academics, while middle and high school charters generally serve at-risk students.

Early attendance at a charter school might provide the basis for subsequent achievement in traditional public schools, especially if students have learned to read at a charter school. This may be why students who transfer from charter schools to public schools in the later grades are prepared for later studies. This could also explain why students who spend the first year in traditional public schools then switch to charter schools for the final two years (TCC), never demonstrate growth significantly different from students who attend traditional public schools all three years (TTT). For students who spend their first two years in traditional public schools, then switch to charter schools (TTC), their poor preparation in the early grades is not overcome once they switch to charter schools at the middle/high school levels because, at higher grades, charters are less likely to focus on academics.

Results also show that the long-term benefits of switching schools outweigh the short-term disruption. Third-grade students in either school type benefit by nearly three additional points in achievement growth from transferring, but switching schools has a negative effect after eighth grade. This decline likely reflects the natural attrition in achievement growth as students get older, and the fact that older students are more likely to transfer to avoid discipline problems or repeating a grade.

APPENDIX A—Detailed Description of Methodology

In classical experimental design, researchers overcome selection bias by randomly placing students in treatment and control groups.[23] In most educational settings, this is simply not feasible. To account for differences in student characteristics that may moderate the effects of the interventions, covariates are used in the analyses.[24] We examine whether there are, in fact, systematic differences among students and schools that might account for differences in outcomes other than those attributable to charter school attendance. Using the longitudinal nature of the data and focusing on achievement growth, we are able to model mean differences between charter school students and traditional school students in both initial status and achievement growth. These models will also take into account differences in school composition.

Examining test scores in any grade level confounds student and school effects with accumulated learning from previous grades and home factors.[25] We therefore assume a "value-added" model of measuring the impact of charter school attendance. The value-added approach is both a common and well-documented approach to evaluating intervention/program effects.[26] Value-added analysis is akin to making "adjusted comparisons."[27] Hence, additional information such as student background factors must be included in the analyses in order to account for their effects on achievement. The background factors are available from school district data containing demographic and personal characteristics, including language proficiency, migrant status, and absenteeism.

Since we use a value-added model and have student information, we are able to counterbalance the fact that students are not randomly assigned to schools. More important, the longitudinal nature of the data enables us to estimate growth trajectories for each student so that students become their own controls. Such controls are not possible with a simple cross-sectional data set.

To further avoid potential bias, we examine model residuals.[28] We are less interested in differences in student achievement in initial status than we are in achievement growth, except to the extent that growth depends upon initial status. Using the panel data allows us to consider charter school attendance as a treatment. In fact, each student can then be considered to be in a specific treatment group, depending on his or her attendance pattern. These groups' initial status and achievement growth rates can then be compared to a quasi-control group (students attending only traditional public schools). Including treatment group indicator variables to model both the initial status and achievement growth yields estimates that identify whether there were any a priori differences between treatment and non-treatment groups in initial status. Including these variables also contrasts the treatment versus

the non-treatment group in growth trajectories and estimates an average treatment effect over the span of the data.[29] Since we are modeling individual growth over time, we can incorporate important factors such as mobility, which can negatively affect achievement results.[30]

The disadvantage of this type of data is that it suffers some attrition. Naturally, we examine both the background characteristics and achievement data to compare those students who dropped out of the analysis against those who remained. Generally, this is accomplished by comparing means, variances, and correlations between the two groups. An additional benefit of using random coefficient growth models is that unlike traditional repeated measure designs, the data can be unbalanced. That is, students with an incomplete set of test scores can remain in the analysis. Compared with traditional analyses that use gains (or differences, for example), complete pairs of data are needed to calculate any pairs of gains. By modeling individual growth trajectories, however, we can generate empirical Bayes estimates for each student's growth over the entire span of the data, whether or not they have data from all of the time points.[31] Further, students need not be considered missing completely at random since the less restrictive assumption of missing at random is reasonable.[32] Using the missing at random assumption allows student missingness to be conditioned on observed variables—including prior test scores—in the model. Given that the less restrictive missing at random assumption is tenable, maximum likelihood estimates will produce unbiased estimates.[33]

Finally, an important consideration is the test metric used in the analysis. This is particularly important when using a longitudinal evaluation strategy. It is important to use a metric that is an equal interval scale, which rules out percentile ranks.[34] Percentile ranks indicate above what proportion of students an individual student's score places him or her. That is, a student scoring at the sixty-fifth percentile, for example, has a score that is better than sixty-five percent of the students in the population. However, small gains in extreme test scores lead to large changes in percentile ranks, because few students are at either end of the distribution. Because there is not a one-to-one relationship between actual changes in test scores and changes in percentile ranks, percentile ranks are not appropriate for longitudinal analysis. It is therefore best to use an equated scale score because metric matters when examining student growth trajectories.[35]

As noted, a longitudinal analysis design is the major component of comparing student achievement in charter and traditional schools. To examine the outcomes noted above, the following methods are applied. As a first step, basic descriptive statistics and cross-tabulations are analyzed to determine whether there are any idiosyncrasies in the data, and also to derive some preliminary relationships that later focus on more complex models.

Descriptive statistics are an important first step, as they are necessary to check whether the assumptions required for a particular analysis are met by the variables. This is to say, means, standard deviations, and plots of data are examined to determine whether data are normally distributed and/or whether there are outliers, for example.

Further, we use cross-tabulations, as they are a straightforward method to examine potential relationships among categorical data. Cross-tabulations allow us to determine if there is potential for joint effects. For example, the cell frequencies of race/ethnicity by language proficiency will obviously have some cells with much higher proportions than other cells. This instance may lead us to consider whether the "effects" of race/ethnicity and language proficiency are additive or multiplicative.

We apply more advanced analytic techniques, since the literature indicates that simple pre/post designs fail to fully capture the processes through which change takes place. Similarly, the use of multiple t tests to examine growth from one period to another is both lacking in its explanatory ability, and misleading due to the increased likelihood of falsely rejecting null hypotheses.[36]

Much of the previous research on achievement gains or growth use a pre/post design to determine the effects of educational or programmatic interventions. Yet, two time points are generally inadequate for studying growth.[37] This inadequacy is due to potential measurement error in the pretest, the covariance structure of random errors, and the indirect method through which traditional analyses determined whether there are effects over time.

Another technique used in the quasi-experimental setting to examine changes in achievement growth and compare this growth between a treatment and a non-treatment group is the differences in differences approach.[38] This approach has been applied in many different settings, including wage effects due to minimum wage law changes, as well as previous comparisons of traditional and charter school student achievement growth. While this approach falls within the value-added framework, a primary assumption is that data from the first time point must be collected long enough before the treatment is implemented, so as not to reflect any adjustments due to the treatment.[39] While this may or may not be plausible in this instance, we use a less restrictive approach.

The methodology we employ is a hierarchical linear model design that has the advantages of directly modeling growth trajectories and being more flexible than traditional analyses. This approach estimates both fixed and random effects. That is, a hierarchical linear model design provides both fixed estimates for initial status and growth, as well as variance estimates that describe the variation in initial status and

growth among students and schools.[40] Since observations are nested within individuals, time intervals need not be constant or the same across individuals as in traditional repeated measures analyses, and, as noted, the number of observations per person may vary.[41] A hierarchical linear model design allows flexible specification of the covariance structure at every level of the analysis.[42]

The hierarchical linear model analysis we employ is based on a three-level model, where at level 1 each student's development is represented by a growth trajectory that depends on a set of parameters. The outcomes are nested within students. At level 2, these individual growth parameters become outcomes that depend upon student-level characteristics, and at level 3, student characteristic effects become outcomes dependent upon school-level characteristics. In other words, for each student we estimate a growth trajectory, which is potentially some function of initial status. Initial status is a student's SAT-9 achievement in 1997-1998 and an estimated growth parameter that measures the average annual growth over the span of the data, which is three years. The growth trajectory could be described with an individual parameter (linear growth), quadratic (curved), or take some other functional form. Generally, with data consisting of only a few occasions, a linear growth model is acceptable.[43] Appendix B provides a brief description for the framework of the three-level hierarchical linear model we use, while Appendix C contains the model's descriptive statistics that we employ in the analyses.

Both the parameter estimates for initial status and growth are allowed to vary among students (within schools). Hence, these estimates become outcomes at level 2. At level 2, we use student characteristic variables to model students' initial status and growth. In other words, we use the student characteristic variables that account for between-student variation in initial status and growth. Of course, in this analysis, the focus is on which treatment group (attendance pattern) a student is in. The coefficients at level 2 are allowed to vary among schools. For example, we may find that a student's growth depends on whether or not that student's primary language is English. The estimated parameter for this effect is allowed to vary among schools. In this way, we are examining whether school context affects the differences in growth accruing with students whose primary language is or is not English. That is, we are examining whether school context can differentially impact relationships between student characteristics and growth.

It is important to use a three-level model in this context because we are particularly interested in examining the effect of charter school attendance on student achievement status and growth. By using a three-level model, we are able to divide the variation in achievement into between-student, between-school, and error components. This is particularly important to do because data containing multiple levels of aggregation can lead to errors in interpretation of results when these

multiple levels are ignored.[44] For example, socioeconomic status—measured at the student level—represents a measure of a student's home resources, while the aggregated mean student socioeconomic status at a school measures the average resources available in a community.[45]

Ignoring the nested nature of the data and simply analyzing outcomes aggregated to school level upwardly biases results of student-level predictors because within schools student-level variation is lost upon aggregation.[46] Not only are the student effects biased, but it also becomes unclear whether the estimated effect is due to a group effect or whether the aggregated variable proxies for an unrepresented student effect.[47]

A more pervasive source of bias is the problem of self-selection. Given that students self-select into either charter or traditional public schools, the estimated charter school effects may be over-estimated, if in fact there is a dependency between achievement and selection.[48] If we assume that all of the potentially important school-selection variables are available to model, then we are concerned with whether the expected value of the residuals is in fact zero, after controlling for attendance patterns, selection, and student background variables.[49] We examine the residuals to determine whether they are randomly distributed with a mean of zero, and whether there is any relationship between the conditional residual and the attendance patterns.[50]

APPENDIX B—Brief Description of Three-level Hierarchical Linear Model (HLM)

The notation is as follows:

Level 1: Y is the outcome (i.e. SAT-9 Reading, Mathematics, or Language). The coefficient is a, predictor is π, and the random effect is e.

Level 2: Coefficient is ß, predictors are x, λ, and the random effect is r.

Level 3: Coefficient is γ, predictor is Z, and the random effect is u.

Hence at level 1:
$Ytij = \pi_{0ij} + \pi_{ij}\alpha_{tij} + e_{tij}$, (1) where Y_{tij} is the outcome at time t for student i in school j, α is the time parameter measured in years. Other modeling options are available, but given the short time span of the data, a simple linear model adequately reflects growth over the period under investigation. Preliminary analyses indicate that growth is not linear across grades, and this is incorporated into the level 2 part of the model. Since growth trajectories are assumed to vary across students, at level 2 for the initial status at time = 0:[51]

$\pi_{0ij} = ß_{00j} + ß_{01j}X_{1ij} + \ldots + ß_{0Pj}X_{Pij} + r_{0ij}$, (2a) where there are p = 1 to P student-level predictors.

For the growth trajectories, $\pi_{1ij} = ß_{10j} + ß_{11j}X_{1ij} + \ldots + ß_{1Pj}X_{Pij} + r_{1ij}$, (3a) There p = 1 to P student-level predictors, as well.[52] The key element of interest among students is the effect of belonging to a particular treatment group. Hence, we add indicator variables for treatment group membership. These groups represent the different attendance patterns of students in the sample. Equations 2 and 3 are modified to include c = 7 groups.

$$\pi 0ij = ß_{00j} + ß_{01j}X_{1ij} + \ldots + ß_{0Pj}X_{Pij} + \Sigma\lambda_{1cj}Group_{1cj} + r_{0ij}, \quad (2b)$$
$$\pi 1ij = ß_{10j} + ß_{11j}X_{1ij} + \ldots + ß_{1Pj}X_{Pij} + \Sigma\lambda_{1cj}Group_{1cj} + r_{1ij}, \quad (3b)$$

Using the λ's in both the intercept and the slope models yields estimates that identify whether there were any a priori differences between treatment and non-treatment groups in initial status, contrasts the treatment versus the non-treatment group in growth trajectories, and estimates an average treatment effect over the span of the data.[53]

Our focus is on the effect of treatment group membership on achievement growth; however, given that students are nested within schools, we include a third level in

the analysis that accounts for school membership during the initial year of the sample. Hence, for initial status for each school j:

$ß_{00j} = \gamma_{000} + u_{00j}$, (4) so that school mean initial status can vary randomly around the grand mean initial status, γ_{000}. Each student level effect is fixed:

$ß_{0pj} = \gamma_{0p0}$. (5)

The growth trajectory for school j is:

$ß_{10j} = \gamma_{100} + u_{10j}$, (6) so that school mean school growth trajectory can vary randomly around the grand mean growth trajectory, γ_{100}. Each student level effect on the growth trajectory is fixed:

$ß1pj = \gamma_{1p0}$. (7)

In other words, an average status and an average growth trajectory are estimated. These estimates are allowed to vary among school and students.

In general, the following steps are employed in building a parsimonious model that explains student growth and the variation in growth among students. Equations 1 through 7 are combined to build a three-level growth model, which describes SAT-9 achievement growth trajectories for each student, how these growth trajectories vary among students, and subsequently among treatment groups to which the students are assigned. The first step is to use an unconditional model (a model with only a growth parameter, but no other predictors) to examine various growth trajectories and provide baseline statistics to evaluate various level 2 and level 3 models. This also provides an estimate of the mean intercept and an estimate of the mean growth trajectory.

Additionally, the unconditional model determines whether these estimates are significant and whether they vary significantly between students and schools. Further, this provides an estimate of the true correlation between the initial status and the growth rate. Normal pre/post designs generally provide spurious negative correlations because the error variance of the pretest is negatively correlated with growth.[54]

APPENDIX C—Descriptive Statistics for Three-level Hierarchical Linear Model (HLM)

LEVEL 1. DESCRIPTIVE STATISTICS

VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
GRADE3	157671	5.30	2.26	0.00	9.00
YEAR0	157671	0.85	0.78	0.00	2.00
READSS	157671	685.21	43.45	482.00	872.00
MATHSS	157671	681.10	46.31	460.00	873.00
CHARTER	157671	0.13	0.34	0.00	1.00
CHARTCUM	157671	0.24	0.61	0.00	3.00
NOYEAR3	157671	0.29	0.45	0.00	1.00
GRADE3SQ	157671	33.22	22.00	0.00	81.00
YEARTWO	157671	0.37	0.48	0.00	1.00
YEARTHRE	157671	0.24	0.43	0.00	1.00
SC97_98	157671	0.33	0.47	0.00	1.00
SC98_99	157671	0.45	0.50	0.00	1.00
HB97_98	157671	0.03	0.16	0.00	1.00
HB98_99	157671	0.02	0.14	0.00	1.00
SAME12	157671	0.67	0.47	0.00	1.00
SAME23	157671	0.75	0.43	0.00	1.00

LEVEL 2. DESCRIPTIVE STATISTICS

VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
YRSNDIST	62207	1.30	1.20	0.00	8.00
ABSENT	62207	2.12	0.67	0.00	9.00
MIGRANT	62207	0.01	0.10	0.00	1.00
SPECED	62207	0.04	0.19	0.00	1.00
GIFTED	62207	0.07	0.25	0.00	1.00
CHARTCUM	62207	0.32	0.75	0.00	3.00
NOYEAR3	62207	0.39	0.49	0.00	1.00
CCC	62207	0.04	0.20	0.00	1.00
CCT	62207	0.04	0.19	0.00	1.00
CTT	62207	0.02	0.16	0.00	1.00
CTC	62207	0.00	0.04	0.00	1.00
TTT	62207	0.80	0.40	0.00	1.00
TTC	62207	0.01	0.11	0.00	1.00
TCT	62207	0.05	0.21	0.00	1.00
TCC	62207	0.03	0.18	0.00	1.00
GRADECCC	62207	0.12	0.72	0.00	9.00
GRADECCT	62207	0.21	1.14	0.00	9.00
GRADECTT	62207	0.11	0.81	0.00	9.00
GRADECTC	62207	0.01	0.19	0.00	7.00
GRADETTT	62207	4.02	2.80	0.00	9.00
GRADETTC	62207	0.05	0.54	0.00	8.00
GRADETCT	62207	0.25	1.22	0.00	9.00
GRADETCC	62207	0.11	0.13	0.00	8.00
GRADE3	62207	4.89	2.25	0.00	9.00
MINORITY	62207	0.42	0.49	0.00	1.00
SC97_98	62207	0.32	0.46	0.00	1.00
SC98_99	62207	0.39	0.49	0.00	1.00
NONPLE	62207	0.12	0.31	0.00	1.00
HB97_98	62207	0.03	0.17	0.00	1.00
HB98_99	62207	0.02	0.14	0.00	1.00
MSGRDE99	62207	0.35	0.48	0.00	1.00
SCTOTAL	62207	0.71	0.82	0.00	2.00
GRDSCINT	62207	3.71	4.06	0.00	12.00
SCTOTCCC	62207	0.03	0.21	0.00	2.00
SCTOTCCT	62207	0.02	0.16	0.00	2.00
SCTOTCTT	62207	0.04	0.27	0.00	2.00
SCTOTCTC	62207	0.00	0.08	0.00	2.00

LEVEL 2. DESCRIPTIVE STATISTICS (Continued)

VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
SCTOTTTT	62207	0.48	0.75	0.00	2.00
SCTOTTTC	62207	0.02	0.00	0.00	2.00
SCTOTTCT	62207	0.06	0.28	0.00	2.00
SCTOTTCC	62207	0.06	0.33	0.00	2.00
SAME12	62207	0.69	0.46	0.00	1.00
SAME23	62207	0.78	0.42	0.00	1.00
MNTYCCC	62207	0.01	0.12	0.00	1.00
MNTYCCT	62207	0.02	0.13	0.00	1.00
MNTYCTT	62207	0.01	0.11	0.00	1.00
MNTYCTC	62207	0.00	0.03	0.00	1.00
MNTYTTT	62207	0.34	0.47	0.00	1.00
MNTYTTC	62207	0.01	0.08	0.00	1.00
MNTYTCT	62207	0.02	0.15	0.00	1.00
MNTYTCC	62207	0.01	0.12	0.00	1.00
NPLECCC	62207	0.00	0.05	0.00	1.00
NPLECCT	62207	0.01	0.07	0.00	1.00
NPLECTT	62207	0.00	0.05	0.00	1.00
NPLECTC	62207	0.00	0.01	0.00	1.00
NPLETTT	62207	0.10	0.29	0.00	1.00
NPLETTC	62207	0.00	0.04	0.00	1.00
NPLETCT	62207	0.00	0.07	0.00	1.00
NPLETCC	62207	0.00	0.05	0.00	1.00
HBTOT	62207	0.05	0.23	0.00	2.00

LEVEL 3. DESCRIPTIVE STATISTICS

VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
N_BREAK	873	71.26	121.64	1.00	915.00
ES	873	0.15	0.35	0.00	1.00
ES6	873	0.28	0.45	0.00	1.00

Appendix D—Description of Data Used

The statewide sample uses data collected by the Arizona Department of Education. These data include several important student characteristics, as well as Stanford Achievement Test, V9, and (SAT-9) reading scores for each student. The SAT-9 was first administered in Arizona in the 1997-1998 school year in grades three through twelve, and subsequently in grades two through eleven. Arizona students did not have unique student identification numbers that spanned years, so the Arizona Department of Education matched students across years using 21 unique combinations of last names, first names, birth dates, an, gender. Overall, 87 percent and 88 percent of all cases were matched between the 1997-1998 and 1998-1999, and 1998-1999 and 1999-2000 school years, respectively. The actual sample was complied first by selecting students based on the first two years of available data (1997-1998 and 1998-1999), and then adding student information from year three (1999-2000) to create a longitudinal dataset with potentially three test occasions per student.

Specifically, we included all students who attended charter schools in either 1997-1998 or 1998-1999 in the sample. Because students can switch school types between years, this creates the basis for three of the attendance patterns based on charter school (C) attendance or traditional public school (T) attendance.

These patterns are: charter school in 1997-1998 and charter school in 1998-1999, CC; charter school in 1997-1998 and traditional public school in 1998-1999, CT; or traditional public school in 1997-1998 and charter school in 1998-1999, TC. The original intent was to create a quasi-experimental condition comparing students who began in traditional public schools (T) and switched to charter schools (C), against students who began and stayed in traditional public schools (TT).

To facilitate comparison, the TT group was created based on matching four criteria of students in the TC group: SAT-9 Reading quartile, SAT-9 Mathematics quartile, grade level, and 1997-1998 T attendance. Thus, students in the TT pattern were similar to students in the TC pattern in 1997-1998 on the afore-mentioned criteria.

Using quartiles, however, still allows for significant variation in individual scale scores so that we can examine differences in initial status among the attendance patterns. All matchable students (based on names, birth date, and gender) who were in the first two years of the sample and had valid SAT-9 Reading scores were included for the third year (1999-2000). Approximately 35 percent of the sample did not have valid data for the third year. Students who moved or graduated accounted for a majority of the sample portion lacking valid third-year data. Problems matching names, birth dates, or gender, and dropping out of school

accounted for the remainder of this sample portion. We included these students in the three-tiered, longitudinal, hierarchical linear model analysis.

These data represent the initial three years of extensive charter school existence in the state. The student sample consists of three years of data that provide 157,671 test scores (level 1), that belong to 62,207 students (level 2), who are considered as belonging to one of eight attendance pattern groups. These students attend 873 schools (level 3). We focus on the student-level results since they estimate the marginal effects of the attendance patterns.

NOTES

1. Lewis C. Solmon, "Findings From the 2002 Survey of Parents with Children in Arizona Charter Schools: How Parents Grade their Charter Schools," December 30, 2002.

2. Darcia Harris Bowman, "Vast Majority of Charter Schools Studies Show Positive Findings," Education Week, November 8, 2000.

3. Students without any SAT-9 scores are excluded, but missing student characteristic data were replaced with the unconditional probability of being in that particular group.

4. J. Wang and P. Goldschmidt, "Opportunity to learn, language proficiency, and immigrant status effects on mathematics achievement," *Journal of Educational Research* 93 (2) (1999): 101-112.

5. In fact, some researchers argue that unadjusted comparisons are meaningless. See H. Goldstein and D. Spiegelhalter, "League Tables and their Limitations: Statistical Issues in Comparisons of Institutional Performance," *Journal of the Royal Statistical Society*, Series A 159 (1996): 385-443.

6. The first step in this analysis examines the unconditional random coefficient model results and uses these as a basis for the ensuing analyses and interpretations. The mean, 661, is among students across all grades and is less meaningful than the initial status once grade is explicitly included in the model.

7. This is the advantage of using a random coefficient model. The within-student variation is the temporal variation, or residual. We are most concerned about the variation among students. Had we been more interested in between-school variation, we would need to add two additional levels to the analysis to account for the cross-classified structure of the data—that is, to account for the fact that some students are nested within more than one school.

8. We focus on the between-student variation, as we are interested in the effect the treatment groups—namely attendance patterns—have on student initial status and growth.

9. The standard deviation is the square root of 802.9 or 28.3, thus + or - 1 standard deviation = 57.

10. Both grade and years cannot be used as time-varying covariates because they would be practically colinear.

11. We proceeded through a series of models that helped explain the contribution of each set of variables, but we present only the results based on the final model in Tables 5 and 6. The intermediate results are available upon request.

12. The criterion for excluding a set of interaction variables was if the change in deviance between a model with the

interactions and without the interactions was not statistically significant. See S. W. Raudenbush and A.S. Bryk, *Hierarchical Linear Models: Applications and Data Analysis Methods*, 2nd ed. (Thousand Oaks, CA: Sage Publications, 2002).

13. Grade has been recoded so that when the actual grade = 3, the grade variable in the model equals 0. In this way, the initial status estimate equals initial status for third-graders (the first grade taking the test in 1997-1998). Given that fact, we can see that the average achievement growth gap between CCC and TTT students is 3.3 points/year (when the grade variable equals 0, or third grade). This gap decreases by 0.7 points per grade, so that in 3.3/0.7, or about 4.7 additional grades, the achievement growth gap will be 0. Adding 4.7 grades to the grade variable equates to grade 7.7, or approximately eighth grade.

14. At an achievement growth rate of 0.6 points per year, it would take "held-back" students 19.2 years to make up for the fact that they start, on average, 11.5 achievement points behind.

15. Raudenbush and Bryk, *Hierarchical Linear Models*.

16. We change equation (1) into a piecewise growth model such as: $Y_{tij} = \pi 0_{ij} + \pi_{1ij}\alpha_{1tij} + \pi_{2ij}\alpha_{2tij} + e_{tij}$ where $\alpha 1$ is coded 1 if the year is 1998-1999 and 0 otherwise, and $\alpha 2$ is coded 1 if the year is 1999-2000 and 0 otherwise. This formula change yields the incremental steps in growth from the initial status over the duration of the data rather than an average growth rate. Yet, both the π's can be modeled as before to include student characteristics and attendance group indicator variables.

17. In the 2000-2001 school year, 39 percent of public school districts nationwide had alternative schools and programs for at-risk children. They served an estimated 613,000 students, roughly 1.3 percent of all students enrolled in public elementary and secondary schools. See John Wirt, Susan Choy, Stephen Provasnik, Patrick Rooney, Anindita Sen, and Richard Tobin, *Indicator 27: Public Alternative Schools for At-Risk Students*, in *The Condition of Education 2003*, National Center for Education Statistics, June 2003, p. 57.

18. We examined whether there were any school-switching-by-attendance-patterns effects on achievement, and there were none.

19. E. A. Hanushek, J.F. Kain, J.F., and S.G. Rivkin, "The Costs of Switching Schools." Paper presented at the annual meeting of the Society of Labor Economists, Boston, May 1999.

20. Continuous Level 1 (student) and Level 2 (school) variables are grand mean centered. Results in Table 11 also control for (at level 2) mean school SAT-9 reading, years in the district, absent, migrant, special education, gifted, minority, and non-primary language English.

21. It is important to reiterate that the there is no statistical difference between traditional public schools and charter schools on the effect of switching schools the prior year.

22. Again, this means that students who were held back in the 1997-1998 school year repeated the same grade during the 1998-1999 school year, and students who were held back in the 1998-1999 school year repeated the same grade during the 1999-2000 school year. As with Table 11, we use a multilevel logistic model to examine the various relationships. Table 12 presents both the log odds associated with a particular variable and the marginal percentage point change associated with each variable (for log odds that are statistically significant, only).

23. D. T. Campbell and J. Stanley, *Experimental and Quasi-Experimental Designs for Research* (Chicago: Rand McNally, 1963).

24. E. Pedhazur, *Multiple Regression in Behavioral Research* (Fort Worth: Harcourt Brace Jovanovich, 1982).

25. R. H. Meyer, "Value-Added Indicators of School Performance," in E. A. Hanushek and D. Jorgenson, eds., *Improving America's Schools: The Role of Incentive* (Washington, DC: National Academy Press, 1996).

26. E. A. Hanushek, "Conceptual and Empirical Issues in the Estimation of Educational Production Functions," *Journal of Human Resources* 14 (3) (1979): 351-388.

27. Goldstein and Spiegelhalter, "League Tables."

28. Pedhazur, *Multiple Regression*, 1982.

29. W. D. Osgood and G. Smith, "Applying Hierarchical Linear Modeling to Extended Longitudinal Evaluations," *Evaluation Review* 19 (1) (1995): 3-39.

30. Hanushek et al., "The Costs of Switching Schools."

31. Raudenbush and Bryk, *Hierarchical Linear Models*.

32. Joop Hox, *Multilevel Analysis: Techniques and Applications* (Mahwah, NJ: Lawrence Erlbaum Associates, Inc., 2002).

33. Hox, *Multilevel Analysis*, 2002.

34. Blaine Worthen, Karl White, Xitao Fan, and Richard Sudweeks, *Measurement and Assessment in Schools*, 2nd ed. (New York: Longman Publishers, 1999).

35. M. Seltzer, K. Frank, and A. Bryk, "The Metric Matters: The Sensitivity of Conclusions about Growth in Student Achievement to the Choice of Metric," *Educational Evaluation and Policy Analysis* 16 (1994): 41-49.

36. R. E. Kirk, *Experimental Design* (Pacific Grove, CA: Brooks/Cole Publishing Company, 1982).

37. Raudenbush and Bryk, *Hierarchical Linear Models*; and D. R. Rogosa, D. Brandt, and M Zimowski, "A Growth Curve Approach to the Measurement of Change," *Psychological Bulletin* 92 (1982): 726-74.

38. David Card and Alan Krueger, *Myth and Measurement: The New Economics of the Minimum Wage (*Princeton, NJ: Princeton University Press, 1995).

39. Daniel S. Hammermesh, "What a Wonderful World this Would Be," *Industrial and Labor Relations Review* 48 no. 4 (1995): 835-838.

40. Initial status is student SAT-9 reading achievement in the 1997-1998 school year, while growth is average gain per year from 1997-1998 to 1999-2000.

41. Raudenbush and Bryk, *Hierarchical Linear Models.*

42. T. A. B Snijders and R. J. Bosker, *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*, (London: Sage Publishers, 1999).

43. Raudenbush and Bryk, *Hierarchical Linear Models.*

44. M. Aitkin and N. Longford, "Statistical Modeling Issues in School Effectiveness Studies," *Journal of the Royal Statistical Society* 149 (1) (1986): 1-43; and L. Burstein, "The Analysis of Multi-level Data in Educational Research and Evaluation," *Review of Research in Education* 4 (1980): 158-233.

45. Burstein, "The Analysis of Multi-level Data."

46. D. Freedman, R. Pisani, and R. Purves, *Statistics* (New York: WW Norton and Company, 1978).

47. Burstein, "The Analysis of Multi-level Data."

48. J. Heckman and V. J. Hotz, "Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training," *Journal of the American Statistical Association* 84 (408) (December 1989): 862-874.

49. Heckman and Hotz, "Choosing Among Alternative Nonexperimental Methods."

50. Pedhazur, *Multiple Regression*, 1982.

51. Initial status is achievement in the 1997-1998 school year.

52. "Growth trajectories" refers to the change in achievement from 1997-1998 to 1999-2000.

53. Osgood and Smith "Applying Hierarchical Linear Modeling."

54. N. Bloomquist, "On the Relation between Change and Initial Value," *Journal of the American Statistical Association* 7 (1977), 746-749.

The Goldwater Institute was established in 1988 as an independent, nonpartisan research and educational organization dedicated to the study of public policy in Arizona. Through research papers, commentaries, policy briefings and events, Goldwater scholars advance public policies based on the principles championed by the late Senator Barry Goldwater during his years of public service—limited government, economic freedom and individual responsibility. Consistent with a belief in limited government, the Goldwater Institute neither seeks nor accepts government funds and relies on voluntary contributions to fund its work.

GOLDWATER
INSTITUTE
in defense of liberty

500 East Coronado Rd., Phoenix, AZ 85004 | *Phone* (602) 462-5000 | *Fax* (602) 256-7045 | *email* info@goldwaterinstitute.org