POLICY REPORT

Another Look at the Effect of Charter Schools on Student Test Scores in North Carolina

CRAIG M. NEWMARK

APRIL 2005

John Locke Foundation

Another Look at the Effect of Charter Schools on Student Test Scores in North Carolina

CRAIG M. NEWMARK
Associate Professor of Economics
North Carolina State University
craig_newmark@ncsu.edu

APRIL 2005

Executive Summary

A 2004 study on the academic impact and effectiveness of charter schools in North Carolina authored for the Terry Sanford Institute by Robert Bifulco and Helen Ladd reached some harsh conclusions regarding the performance of the charter schools. Using three different models that compare state end-of-grade (EOG) test scores for regular public school students and charter school students, Bifulco and Ladd conclude that North Carolina charter schools are not only failing to improve their students' academic performance, but are actually hurting it.

The analysis here suggests that the Bifulco and Ladd's research faces serious problems: some of which this report attempts to adjust for, some of which are revealed by statistics that the researchers do not report, and others that serve to question the relevance of the comparisons being made.

While the problems are numerous, there are three that should be highlighted. First, Bifulco and Ladd do not consider specific student characteristics that could significantly affect the results of their study. For example, the fact that charter schools have a much smaller percentage of gifted students than regular public schools is ignored. In each year studied, the percentage of gifted regular public school students is at least 4 percentage points higher than in charter schools. In the last year of the study, 2002-2003, this percentage is more than three times that of charters: 13.62 percent for regular schools compared to 4.2 percent for charters. Failure to incorporate this fact into their study biases it against charter schools.

Bifulco and Ladd also do not distinguish among types of charter schools, many of which were established to serve "at risk" students. Adjusting the Bifulco and Ladd study to account for these differences in student population cuts the difference between charters and regular public schools by more than half. In fact, such an adjustment nearly eliminates the differences in reading scores. When comparing at-risk charter students to those attending special education or alternative regular public schools, the "charter school effect" is not significantly different from zero.

Furthermore, Bifulco and Ladd fail to report an important statistical result which measures the extent to which the model used explains the results implied in the data. This is known as the "R-squared" statistic. The analysis in this report finds that the R-squareds are quite low, meaning that a large part of the results obtained are left unexplained by the models used. Given that so much is unexplained, we should treat the findings of these models very cautiously. Indeed, with respect to the specific model described by Bifulco and Ladd as their preferred approach (the "fixed effects model") there are certain aspects of the R-squared results that are so low that it is difficult to see how anyone could have confidence in the inferences generated.

Ultimately, there is a fundamental question regarding the appropriateness of comparing EOG test results for charter and regular public schools. It is not clear whether personnel in charter schools and regular public schools are equally motivated to increase student EOG scores. Many charter schools use alternative instructional materials and emphasize different topics and subjects at different times. While charter students are forced by state law to take EOG tests, many do not consider them important and focus on other measures such nationally standardized tests. Since this difference is very difficult to adjust for, to the extent that charters focus on different measure of performance, both Bifulco and Ladd's results and the results obtained here will be biased in favor of regular public schools.

Acknowledgments

This research was funded by a grant from the John Locke Foundation. I thank the Foundation and especially John Hood, Kory Swanson, and Roy Cordato for their support, assistance, and patience.

Data was provided by the North Carolina Education Research Data Center. I thank the Center's Director, Beth Glennie, for her help.

Finally, I thank Steve Margolis, Doug Pearce, Wally Thurman. and Baker Mitchell for helpful discussions.

I. Introduction: The Problem of Evaluating Charter Schools

Do students in charter schools achieve academically better, worse, or the same as students in regular public schools? Attempts to answer this question face two significant problems.[1] Students in charter schools differ systematically from students in regular public schools. Charter school students would be different if for no other reason than that their parents or guardians chose to remove them from regular public schools and enroll them in charters. This might well indicate that charter students have "parents who are extraordinarily concerned about their children's success at school and are highly motivated to intervene on their children's behalf."[2] If so, and if charter students earn higher test scores than students attending regular public schools, those higher test scores might not indicate that charter schools teach better but only that charter students have advantaged backgrounds. However, many observers have pointed out that charter students tend to be students who are experiencing some sort of difficulty. If so, and if charter students test less well than regular public school students do, those lower test scores might not indicate that charter schools instruct poorly but only that charter students have problems that tend to lower their test scores.[3]

What we would like to do is compare the test scores of charter school students to regular public school students after having subtracted differences in scores caused by differences in students' personal characteristics, family backgrounds, and prior schooling. As Caroline Hoxby points out, the best way to do this would be to randomly assign students to schools: "[T]he most credible research is that in which choice students are compared to students who applied to the same choice program but who were randomly not assigned to a voucher or charter school."[4] But because this is not usually possible—only one such study has been done to date for charter schools (Hoxby and Rockoff 2004)—researchers try to control for these other differences through statistical techniques. This poses a significant problem for investigators because the accuracy of every one of these statistical techniques depends upon specific assumptions and these assumptions are difficult to verify. It is thus difficult to determine which of these other techniques is accurate and which, if any, is "best."

The other major problem in assessing the academic performance of charter schools is that charter schools are quite dissimilar. They vary considerably in the students they try to serve, in their objectives, and in their methods. In a study of North Carolina charter schools, George W. Noblit and Dickson Corbett wrote (2001, p. I-6):

"[C]harter schools vary tremendously in terms of their primary reasons for existence. . . . [T]he schools were quite diverse as reflected in their distinctive missions. These included one or more of the following: challenging gifted students; assisting students having difficulty in traditional public schools; maintaining small class and/or school size; facilitating individualized instruction; enhancing local control; providing arts-enriched or multiple intelligence-enhanced academic opportunities; increasing academic and/or behavioral discipline; returning to "the basics"; incorporating research-based instructional models or curricula; and/or attending to cultural enrichment."

What we would like to do is compare the test scores of charter school students to regular public school students after having subtracted differences in scores caused by differences in students' personal characteristics, family backgrounds, and prior schooling.

Other investigators report considerable heterogeneity in charter schools.[5] This means that comparing the academic performance of charter schools in total to the performance of regular public schools is potentially misleading; at the very least, such comparisons obscure important information. Identifying which charter schools seem to perform better than other charter schools, and why, would be very useful.[6]

II. Review of Previous Research

A. STATES OTHER THAN NORTH CAROLINA

There is a rapidly growing literature on the academic performance of charter school students. One must apply results from other states cautiously to North Carolina since there are considerable differences among states in charter school laws, in the age of charter schools, in schooling options available other than neighborhood public schools (magnet schools, private schools, home schools, use of vouchers), in the proportion of charters that are targeted to special student populations, in the proportion of charter schools that are converted regular public schools, and in the proportion of charter schools that are run by for-profit companies.[7] And it is beyond the scope of this paper to review this research in detail. Nonetheless, five aspects of this research are relevant to the present research and are worth noting.

Several studies conclude that the age of a charter school significantly affects its impact on students' test scores.

1. *Conclusions on charter performance*. Some studies find that charter schools raise students' academic performance relative to regular public schools. But some find that charter schools actually lower students' academic performance. Gary Miron concludes (2003, p. 8): "[T]he student achievement results from charter schools were mixed at best [footnote omitted]."[8]

2. *At-risk students*. In a study of Texas charter schools, Timothy J. Gronberg and Dennis W. Jansen report (2001, pp. 37–38 and 41–42) that at-risk students in charters seem to substantially outperform at-risk students in regular public schools. But Tim R. Sass contends (2004, p. 4) that this result may be misleading because Gronberg and Jansen estimate their model by ordinary least squares (OLS) in a setting in which OLS is known to yield biased results. Sass also finds (p. 20) that for Florida schools, at-risk students in charters seem to perform worse than at-risk students in regular public schools—though this result is not statistically significant.

3. *Continuing students*. Gronberg and Jansen group students in a potentially useful way. They examine students who remain in either regular public schools or charter schools for three consecutive years separately from students who move between regular public schools and charter schools. They find (p. 29): "One key feature in these tables is the strong performance—in terms of changes in test scores—for students continuing in charters. For every year in our sample, for both math and reading, continuing charter students . . . show greater increases in their test scores than do continuing traditional public school students [footnote omitted]."[9] Lewis Solmon and Pete Goldschmidt find (2004, pp. 12–13) that students who attend charter schools for three consecutive years exhibit higher test-score growth than students who attend regular public schools for three consecutive years.

4. *Effect by grade*. Sass reports (2004, pp. 19–20) that in Florida, the effect of charter schools on performance varies by grade. For reading, charter school students lag behind regular public school students the most in the elementary grades, next most in the middle-school grades, and least for the high-school grades. For math, this ordering is reversed: the gap between charters and regular public schools is greatest for high-school students and least for elementary-school students.

5. *Age of charter schools*. Several studies conclude that the age of a charter school significantly affects its impact on students' test scores. Eric A. Hanushek, John F. Kain, and Steven G. Rivkin report (2002, p. 18) that students attending charter schools in their first or second year of operation increase their reading scores substantially less than do regular public school students; for math, students attending charter schools in their first year of operation increase their scores substantially less. But students attending older charters (second-year or greater for math, third-year or greater for reading) progress as well as students in regular public schools. Sass reports (2004, pp. 15–16) a similar finding for Florida charter schools (although the time to equal performance in charters is longer: three to five years for reading and four to seven years for math). Michael Agostini states (2003, p. 1) that in California, "Results show that charter schools that have operated for five or more years outperformed non-charter public schools and younger charter schools." Other studies showing that newer charters (first-year, and in some studies also second-year) seem to perform significantly less well than older charters include Loveless (2002, pp. 33–34) (data from ten states); Gronberg and Jansen (2001, p. 41) (Texas); Dickman et al. (2003, p. 17) (Wisconsin); and Miron et. al. (2002, p. 151) (Pennsylvania).

B. STUDIES OF NORTH CAROLINA CHARTERS

To my knowledge, only five research studies have compared the test scores of charter students to regular public school students in North Carolina: Noblit and Corbett (2001); Greene, Forster, and Winters (2003); Hoxby (2004a, 2004b); and Bifulco and Ladd (2004). I will discuss the first four only briefly. The study by Bifulco and Ladd merits more discussion.

An early study of North Carolina charter schools, done for the State Board of Education, is Noblit and Corbett (2001). Noblit and Corbett report that charter school students improve their reading and math end-of-grade (EOG) test scores less rapidly than regular public school students (pp. IV-12 and IV-20). But Noblit and Corbett do not—and they emphasize that they do not—adjust the raw-score differences for any of the systematic differences between charter school students and regular school students (p. I-11). That is, they do not address, at all, the first problem discussed in Section I above. And they do not address the second problem, either: they note that there is considerable variation in charter school performance that they do not account for. And finally, their sample of charter schools is small, and their study includes data only from the first three years that charter schools were in existence in North Carolina.

Jay P. Greene, Greg Forster, and Marcus A. Winters (2003) and Hoxby (2004a, 2004b) do address the problem that charter students are different from other public school students. They address the problem in the same way: they compare the performance of each charter school to the performance of a close, untargeted, regular public school.[10]

But this approach has two serious drawbacks. First, it appears—although more data would be helpful—that some charter schools draw students from much larger regions than just their local neighborhoods. This is indicated by answers to two "Frequently Asked Questions About Charter Schools" on the Department of Public Instruction's website (http://www.ncpublicschools.org/charter_schools.facts.html): (1) "Must Charter School students reside within the county where the Charter School is located?" "No . . ." (2) "If there is more than one school system in a county, must the student reside within the school district where the charter school is located?" "In this case, the child may attend any charter school within the county."[11] To the extent a charter school draws from a wider region, the closest regular public school could provide a poor control. Second, with this method there is no good way to handle charters that target special populations of students. Greene, Forster, and Winters (2003) drop such charters from their analysis, but this limits the scope of their conclusions. Hoxby, on the other hand, in her first paper (2004a) retains targeted charters. But it is almost certain that students attending a targeted charter are not comparable to the students attending the neighborhood public school. (In her second paper that includes North Carolina schools, 2004b, Hoxby attempts to control for charters that target at-risk students. But her method of identifying at-risk charters seems too restrictive; see Newmark 2005.)

C. BIFULCO AND LADD

1. Introductory problems

Before discussing Bifulco and Ladd's methodology, I should note three problems that they cannot solve—and that I will not be able to solve—in studying the academic performance of North Carolina charter schools. Since these problems cannot be solved, after stating them I will ignore them. But the existence of these problems should warn those who would apply the results of academic studies to be cautious.

For the third- through eighth-grade students in their sample, Bifulco and Ladd (2004) are missing 9.5 percent of the possible test scores (p. 13): "Test scores from a particular year might be missing for a student because that student left the North Carolina public school system, was exempted from taking the test, or has a missing or an invalid test score for some other reason." Note that 9.5 percent is only a minimum for the percentage of student observations Bifulco and Ladd are missing; students who never entered the public school system, because they were in private school or were home-schooled for those grades, do not appear in their sample at all.

That some of these observations are missing is benign. For example, a student might have moved out of state because one of his parents took a better job. But if some of the observations are missing systematically, not randomly, their absence could distort the results.[12] It is impossible to determine whether this distortion is large or not. But we can consider two scenarios that will indicate some of the possibilities. Suppose that all of the missing observations are from students who would be doing poorly in regular public schools and are scoring well below average on the EOG tests—say, one standard deviation below the mean. Further, suppose that all of these students have recently left North Carolina regular public schools. Since the missing students equal only about 10 percent of the number of students in the regular public schools, their

To the extent a charter school draws from a wider region, the closest regular public school could provide a poor control.

absence raises the average score of regular public school students by 0.1 of a standard deviation. This is not a small amount, but neither is it very large. Depending on the model and the test (reading or math), Bifulco and Ladd conclude that regular public school students score between 0.095 and 0.265 of a standard deviation above charter students.

But consider another scenario. Half of the missing observations are missing for benign reasons. But the other half are from above-average students—say, 0.5 standard deviation above average. And suppose all of those students, whose parents are for one reason or another dissatisfied with the regular public schools, would be attending charter schools, if charter schools had more room and if the number of charter schools in North Carolina were not capped by law. Since charter students are about 1.5 percent of the total number of students in North Carolina schools, the missing students would amount to more than three times the number of students currently in charters. Such students would raise the average score of charter students about 0.4 of a standard deviation, a substantial amount.[13]

A final issue that has an unknowable impact on the results is whether personnel in charter chools and regular public schools are equally motivated to increase student EOG scores.

A second issue is one raised by Dale Ballou (2002): scores, and even the growth in scores (score "gains"), are not necessarily comparable across students. Ballou notes that raising a test score from 15 to 20 might not be of the same difficulty as raising a score on that test from 20 to 25. He notes further (p. 14) that this problem "applies to several popular methods of standardizing raw test scores that fail to account sufficiently for differences in test items—methods like recentering and rescaling to convert scores to a bell-shaped curve. . . ." (Such "recentering and rescaling" is also known as "standardizing.") In other words, Ballou is arguing that we cannot be sure that a student who raises his score from -0.5 standard deviation to the mean has made the same amount of progress as a student who raises his score from the mean to +0.5 standard deviation. But such standardized scores are precisely what Bifulco and Ladd—and I, following their lead—analyze.

A final issue that has an unknowable impact on the results is whether personnel in charter schools and regular public schools are equally motivated to increase student EOG scores. Richard Buddin and Ron Zimmer observe (2003, p. 37): "Charter schools may pursue goals other than academic achievement or they may emphasize material that is not well measured on standard achievement tests. In short, charter schools are not designed to be rescaled replicas of conventional public schools. Indeed, they may differ from other schools in many dimensions that may have direct and indirect implications for student achievement."

Here is one expert opinion that there may be differences between regular North Carolina public schools and charter schools in pursuing high EOG scores. Joe Maimone, headmaster of the Thomas Jefferson Classical Academy charter school, recently stated (Sherman 2004): "The reason charter schools exist is because parents were tired of all the focus being on these end-of-grade tests. We focus on making sure our kids learn what they need to be successful and how they perform on national tests like the SATs, not how they perform on this one test."

And here is an illustration that a charter school's motivation to increase EOG scores can be quite low. Deirdre Fernandes (2002) discusses East Winston Primary, a charter for at-risk preschool to fifth-grade students. For the 2000–01 school year, less than 8 percent of East Winston's students performed at grade level on the EOGs. But for the 2001–02 school year, that percentage rose to 76! What could account for such a dramatic improvement in such a short time? The school's executive director stated that the school's board of directors overhauled the teaching staff and hired new administrators who changed the curriculum to match the state's testing program.[14]

But East Winston's 2000–01 scores are in the data, as are the scores of other charter schools who might be, at least for some length of time, pursuing goals other than high EOG scores.[15]

2. Outline of Bifulco and Ladd's methodology

Bifulco and Ladd (2004) start with the most general model of student achievement. Achievement of student i in time t results from a production function, which is a function of the current and past values of all personal, family, and schooling inputs:

(1) $A_{it} = A_t[F_i(t), S_i(t), m_{i0}, e_{it}]$[16]

A_{it} represents the academic achievement of student i at age t. A_t is the production function for academic achievement at time t and is a function of four kinds of inputs. F_i is the set of the student's values for personal and family factors. (So, for example, F_1 might be the number of hours a parent spent reading to the child, and we would expect that to be positively related to A_{it}; F_2 might be the number of hours per day a student spent watching TV, and we might expect that to be negatively related to A_{it}.) $F_i(t)$ is a vector of both the current values and the past values of these factors, back to the student's birth. S_i is the set of the student's values for schooling inputs and $S_i(t)$ is a vector of both the current values and all the past values of those inputs. μ_{i0} represents the endowed characteristics of the student that do not vary with age; for example, the I.Q. the student was born with. Finally, e_{it} represents random error.

Assuming that the production function does not vary with student age (A_t = A for all time periods), academic achievement for student i at time t can be written

(2) $A_{it} = a_1F_{it} + a_2F_{it-1} + \ldots a_tF_{i1} + b_1S_{it} + b_2S_{it-1} + \ldots + b_1S_{i1} + g_tm_{i0} + e_{it}$

where the a's represent vectors of weights, weights showing the impact of each of the past personal/family factors on current academic achievement, the b's represent vectors of the weights on the schooling inputs, and g_t represents the weight of the individual endowment in year t. Equation (2) is known as the cumulative model. Estimating it is obviously difficult because of the data required. (For instance, the individual endowment m_{i0} will never be observed, so g_t can never be estimated, at least not directly.) Bifulco and Ladd, like other investigators, make simplifying assumptions to transform the model into one that can be estimated. Depending on the set of assumptions made, they derive three models that can be estimated: the levels model, the gains model, and the fixed effects model. (For more detail, see Bifulco and Ladd 2004, pp. 12 and 14–18.)

First, suppose all past personal/family factors and schooling inputs have zero effect on current achievement. And further suppose that the individual's endowment has zero effect on current achievement. This assumption thus sets all the a's from 2 through t equal to zero, all the b's from 2 through t equal to zero, and g equal to zero. Transform S_{it} into a zero-one variable indicating whether student i attended a charter school in year t (1 = yes). Finally, Bifulco and Ladd add a set of grade-by-year binary variables to "capture systematic differences across exams" (p. 15), a variable which indicates whether a student made a "structural" change of school—a promotion from elementary school to middle school or junior high—between year t and year t-1, and a variable which indicates whether a student made a nonstructural change of school between t and t-1. The latter two variables are included "[t]o distinguish the effects of charter schools themselves from the effects of movement among schools" (p. 12). This yields the levels model, equation (3), in which the student's achievement in year t is a function of his personal/family characteristics in year t, whether or not he attends a charter school in year t, two types of school change, grade-by-year binary variables, and random error at time t:

(3) $A_{it} = aF_{it} + bCH_{it} + dSTRCH_{it} + zNSTRCH_{it} + fh_t + e_{it}$ where CH_{it} represents the binary variable for charter school attendance, $STRCH_{it}$ and $NSTRCH_{it}$ represent structural school change and nonstructural school change respectively, and h_t is a vector of grade-by-year binary variables.

A second approach replaces the single restriction that past personal/family and schooling inputs have zero effect on current performance with three other restrictions. Start again with equation (2). Now assume that personal/family inputs are constant over time, $F_{it} = F_{it-1} = \ldots = F_{i1}$. Further assume that the initial endowment's impact on achievement is the same in all years, $g_t = g$. Finally, assume that all past school inputs have an immediate, one-time effect on current achievement and that this effect does not decline with time. Sass explains the last assumption (2004, p. 10) as requiring, for example, that the quality of a child's kindergarten schooling has the same effect on his achievement at age eighteen as at age five. Bifulco and Ladd show (p. 17) that these assumptions yield the gains model:

(4) $DA_{it} = A_{it} - A_{it-1} = aF_i + bCH_{it} + dSTRCH_{it} + zNSTRCH_{it} + fh`_t + e`_{it}$

where the dependent variable is now the gain in the student's achievement from year t-1 to year t.[17]

Bifulco and Ladd's third approach is to modify equation (3) by replacing current personal/family inputs with a vector of indicator variables that uniquely identify each individual student. This yields the fixed effects model:

(5) $DA_{it} = aST_i + bCH_{it} + dSTRCH_{it} + zNSTRCH_{it} + fh`_t + e`_{it}$ where ST_i is a vector of student fixed-effect variables, ST_i equal to one for the ith student, 0 otherwise. Bifulco and Ladd write (p. 17) that using gains as the dependent variable "eliminates the need to control for previous educational experiences" while including the student fixed-effect variables controls "for any unobserved differences between charter school students and traditional public school students that remain constant over time." They further emphasize that "the estimated effects of charter schools from this model are based on the experiences of only those students who have test-score gains observed at least once in a charter school and at least once in a traditional public school [footnote omitted]."

A second approach replaces the single restriction that past personal/family and schooling inputs have zero effect on current performance with three other restrictions.

Bifulco and Ladd obtain the end-of-grade test scores from the North Carolina Education Research Data Center. In order to make the test scores comparable across years and across grades, each individual student's test score in year t is "standardized": the mean of all scores for the student's grade level, in year t, is subtracted from the student's test score, and the result is divided by the standard deviation of all scores for that grade, in that year t. The standardized test scores for each grade in each year thus have a mean of zero and a standard deviation of one. Each student in year t has possibly two standardized test scores, one for reading and one for math.

Data on the independent variables that Bifulco and Ladd use were also obtained from the North Carolina Education Research Data Center, though this data ultimately originates from the National Center for Educational Statistics's "Common Core of Data, Public School Universe Survey."

The data Bifulco and Ladd use span the school years 1995–96 through 2001–02. Starting with third-graders in each of the years 1995–96, 1996–97, 1997–98, 1998–99, and 1999–2000, they include each student's scores through eighth grade, if available, or through the last year of the data, 2001–02, whichever comes first (p. 11 and p. 50).

3. Analysis of Bifulco and Ladd's models and procedures

I will first make comments that apply to all of Bifulco and Ladd's models, and then I will make comments specific to each of the three individual models.

General comments. The single binary variable for charter schools does not distinguish among the various types of charter schools. Noblit and Corbett observed (2001, p. I-7) that among North Carolina charter schools are schools targeted toward four special populations of students: at-risk students, economically disadvantaged students, students with special needs or disabilities, and academically gifted and/or college-bound students. (Of these four types of targeted charters, at-risk charters were by far the most numerous. At least during the period 1998 through 2000, the number of at-risk charters was greater than the number of the other three categories combined. Noblit and Corbett 2001, p. III-6.) Students at these charter schools could well test differently from students at mainstream charters. Bifulco and Ladd appear to recognize the potential usefulness of distinguishing among charters, but they stated in an earlier draft (Bifulco and Ladd, undated, p. 21) that the data did not allow them to distinguish charter schools by type of target population.

In fact, there is evidence of strong demand for charter schools, both nationally and in North Carolina. The Chicago Tribune states (2004) that in Chicago, waiting lists for entry into charter schools are so long, most charters have stopped recruting actively.

As noted above, Bifulco and Ladd include grade-by-year binary variables in their models to "capture systematic differences across exams." I do not understand why these variables are included. Bifulco and Ladd have standardized the test scores within each grade in each year. It seems to me that systematic differences across exams should be removed by these standardizations, but Bifulco and Ladd do not discuss the issue further.

I also question including the structural switching variable. Note that a much higher fraction of charter schools than regular public schools combine the elementary grades and middle grades. For example, for the 2003–04 school year, 31 of 97 charter schools combined grades K through 8, while only 98 of 2,166 regular (non-special) schools did (http://www.ncpublicschools.org/nceddirectory/). Charter schools offer a greater chance that students will not have to make a structural switch, thereby avoiding any drop in test score associated with such a switch. Avoiding this drop should be "credited" to the charter schools, but including a structural switching variable prevents that.[18]

I note that Bifulco and Ladd did not use all the data that are available. In their attempt to follow each cohort of students completely from third through eighth grade, they discarded some data for cohorts without complete data: they did not use data on fourth- through eighth-graders from 1995–96, sixth- through eighth-graders from 1997–98, seventh- and eighth-graders in 1998–99, eighth-graders in 1999–2000, third-graders in 2000–01, and third- and fourth-graders in 2001–02 (2004, p. 50). If a student had to repeat a grade, and thus had two or more observations for the same grade, Bifulco and Ladd dropped all but one of those observations (p. 36, footnote 9). And they apparently let the number of observations used to estimate each of their three models be dictated by the model that used the fewest observations. To estimate the fixed effects model, they need a student to have at least three years of data; neither the levels model nor the gains model requires this. But they seem to have estimated all three models using only the students available to estimate the fixed effects model (p. 43).[19]

As a final general comment, consider an observation about their results that Bifulco and Ladd themselves make. Each of their three models yields the result that not only are charter schools not improving their students' academic performance, they are actually hurting it. If this is true, Bifulco and Ladd observe (p. 34), it is hard to explain why parents continue to send their children to these schools. In fact, there is evidence of strong demand for charter schools, both nationally and in North Carolina. The Chicago Tribune states (2004) that in Chicago, waiting lists for entry into charter schools are so long, most charters have stopped recruiting actively. Patrick J. Wolf wrote recently (2002, p. 86): "The bulk of the evidence supporting charter schools is found in surveys of parental satisfaction and in the undeniable fact that an increasing number of parents are placing their names on the waiting lists of oversubscribed charter schools." Lee Anderson et al. report (2002, p. 18) that 62 percent of charter schools nationwide had waiting lists. Andrew J. Rotherham (2003) estimates that 70 percent do. I do not have a comparable up-to-date figure for North Carolina charters, but there are stories suggesting growing demand: a Pamlico County charter school added a new wing or a new building almost every year (Manual and McLaughlin 2002); a High Point charter school plans to expand from 250 students to 1,000 (Dominello 2004); a Brunswick County charter school that had 65 students in 2000 expected enrollment in 2004 of 525 students (Jones 2004).[20]

Bifulco and Ladd suggest two possible explanations for this paradox: either parents are choosing poorly or parents believe charters have some important nonacademic advantages ("the decision to enroll in a charter school is not motivated solely by concerns with academic achievement"). But the available evidence is inconsistent with parents choosing poorly. Joseph L. Bast and Herbert J. Walberg (2004) report that parents rate schools the same as experts and choose schools based on academic quality.[21] Hoxby argues (2003, pp. 45–46) that there is "substantial evidence" that parents do not assess schools ignorantly or superficially and that they evaluate schools on "academics, discipline, a supportive atmosphere, and safety." Hoxby writes (p. 45):

"[A] parent is revealing his belief that a choice school is better when he continues to send his child there, rather than the regular public school his child could freely and easily attend. Suppose we were to find that students' achievement was no better in choice schools. What would we then conclude, knowing that the parent still prefers the choice school? We might conclude that the parent valued some aspect of the school other than achievement (such as discipline or safety); we might conclude that the student's achievement was higher on some dimension not measured by standardized tests. Given that parents observe much more than an econometrician does about his child's schools, it would be foolish to conclude that the parent was simply wrongheaded."

If parents are not making systematic mistakes, two explanations remain. One—as suggested by Bifulco and Ladd, but beyond the scope of this study—is that charter schools offer important nonacademic advantages compared to regular public schools. This is an important area for further research. The other possibility is that Bifulco and Ladd have overstated the academic liability of charters. The rest of this paper addresses that possibility.

Comment on the levels model. Bifulco and Ladd, like other investigators, do not endorse this model. Its restriction that the student's endowment and all the student's past personal, family, and schooling inputs have zero effect on academic achievement seems too restrictive. This model also has the problem that we do not have data on the complete set of personal and family characteristics that affect achievement. If some of these unmeasured variables affect parents' decisions to send students to charters, the coefficient of charter schools will be biased and potentially misleading.

One small improvement might be possible, however, and that would be to increase the number and range of personal, family, and schooling inputs controlled for in estimating the model.

Comments on the gains model. This model, too, can be questioned. Two of its assumptions—that the amount of each personal and family input is constant over time and that each input's impact on achievement does not decay over time—are implausible. And for both assumptions there is some inconsistent evidence. Andrew J. Houtenville and Karen Smith Conway contend (2003) that, at the margin, parents should adjust their inputs into their children's education in response to the amount of the inputs the children receive from school. They report evidence that higher school inputs are associated with less parental effort; hence, family inputs are not constant. Sass tests the assumption (2004, p. 15) that the impacts of the inputs remain constant over time. With Florida data, he strongly rejects the assumption.

Given that parents observe much more than an econometrician does about his child's schools, it would be foolish to conclude that the parent was simply wrongheaded.

This second assumption can be relaxed if a general version of the gains model is estimated. In the "value-added" model, the dependent variable is current academic achievement, A_{it}, and lagged academic achievement, A_{it-1}, is included as an independent variable.[22] But for the value-added model to be valid, we must accept the assumption that the impacts of all inputs, both observed and unobserved, decline over time at exactly the same rate. This assumption can be rephrased: lagged achievement must be a sufficient statistic for the past values of all inputs, observed and unobserved. But Petra E. Todd and Kenneth I. Wolpin (2004, p. 20) report evidence that rejects that hypothesis. Even when lagged achievement is included as an independent variable, lagged values of the inputs affect current achievement.

And note that Bifulco and Ladd warn (p. 17) that "if unobserved differences between charter school and traditional public school students affect the rate of growth in student performance as well as its level . . . the gains model will generate biased estimates of the effect of attending a charter school." Because of this, Anneliese Dickman et al. (2003, p. 12) seriously question the applicability of the gains model to at least at-risk students.[23]

Comments on the fixed effects model. The fixed effects model is the one that Bifulco and Ladd prefer (p. 17). The model identifies the impact of charter schools on achievement by comparing the score gains of students who switch from regular public schools to charters and from charters to regular public schools. I think there are three potential problems with this model.

First, this model, too, requires a strong assumption. As noted by Hanushek, Kain, and Rivkin (2002, p. 7), who applied the model to Texas data, "The key identifying assumption is that entry into a charter school is not correlated with other changes in family or student circumstance that affect achievement." But we have very little information on why students switch into or switch out of charter schools. Bifulco and Ladd's model leaves entry into and exit from charter schools completely unexplained.

Parents will not necessarily wait for a problem to manifest itself in lower test scores or lower score growth before they switch their children to charter schools.

To their credit, Bifulco and Ladd acknowledge this potential problem and respond. They try to check whether students switching into charters have the same trend in test scores before they switch as do all other students (p. 22). Using only observations from students who will attend charter school—but before they actually attend—Bifulco and Ladd estimate this regression: $DA_{it} = A_{it} - A_{it-1} = at + h_t + e_{it}$ where DA_{it} is the student's score gain and t is a counter taking the value zero in 1997 and increasing by one each year thereafter. Since a would be zero if all the observations were used, Bifulco and Ladd argue that testing whether a is negative tests whether students who switch into charter schools have a worse trend in scores than other students do. They find a to be insignificantly different from zero for both math and reading score gains; they conclude that the identifying assumption is valid.

While I cannot see a straightforward way to do this test better, I have some doubts about it. Suppose a student has a score gain two years before he enters a charter school of +0.5 standard deviation and a score gain one year before he enters of -0.5 standard deviation. Bifulco and Ladd's equation would measure zero trend in this student's score gains, but the score gain immediately preceding entry, likely more relevant, is negative. Note, too, that instead of estimating the trends in score gains of individual students and then averaging them, Bifulco and Ladd's equation takes an average of the score gains and computes the trend in this average. We want an average of trends; they estimate a trend in averages. Does it matter? I believe that if the exact same students were observed every year, it wouldn't. But the exact same students are not observed every year. Students can be observed only when they enter third grade; as they enter the data, they change the set of students from the year before. And some students are only observed as they move from non-public schools into public schools or as they move from out-of-state schools to North Carolina public schools. As they enter the data, they change the set of students. Finally, we should expect parents to behave as other decision-makers and be forward-looking. Parents will not necessarily wait for a problem to manifest itself in lower test scores or lower score growth before they switch their children to charter schools. Bifulco and Ladd's test does not accommodate any such switches. (I concede that I don't know if these problems are large enough to affect the outcome of Bifulco and Ladd's test.)

A second potential problem with their fixed effects model is that it is common practice when estimating a model of fixed effects to test whether the fixed effects actually belong in the model. This is done by computing an F-statistic for the null hypothesis that the fixed effects, jointly, are all equal to zero (Greene 2000, p. 526). Bifulco and Ladd do not report the result of such a test.

A third potential problem for this model is indicated by the warning about statistical modeling given by Bruce D. Meyer (1995, p. 151): "If one cannot experimentally control the variation one is using, one should understand its source." Neither the fixed effects model, nor the literature on charter schools, provides any detailed information on why students switch into, and switch out of, charter schools. A limited bit of information on exit, but the best I have seen, comes from Noblit and Corbett (2001, p. I-7): "The primary reason [charter school] directors gave as to why most students left were discipline, transportation problems, the school's program did not meet student needs, and the school's environment was too structured."

This raises the possibility that charter schools and some switching students simply do not match well. To the extent that the fixed effect model focuses on these students, it focuses on the students least well served by charters, but reveals little about the students charters may serve well. It seems reasonable to ask: What about the scores, and the score growth, of students who remain in charters?

III. Replication and Extension of Bifulco and Ladd

In this section, I report how sensitive Bifulco and Ladd's results are to the introduction of a few changes, changes prompted by the discussion above. In particular, I consider these six changes:

1. More data. I include an earlier year, 1994–95, for which data are available and a later year, 2002–03, for which data has become available since Bifulco and Ladd did their analysis. I also include observations from incomplete cohorts (see p. 16 above) and all observations available from the years students fail a grade. Finally, I estimate each model with the maximum number of observations available for that model.

2. More individual characteristics. I control for students who are academically gifted, learning disabled (both generally, and for either math or reading specifically), behaviorally/emotionally handicapped, and mentally handicapped.

3. Distinctions among charter schools. I enter a binary variable that equals one for charter schools that focus on at-risk students and also a binary variable that equals one for charter schools that focus on the arts or cultural issues.

4. More school characteristics. I look at variables that control for different types of regular public schools: magnet, vocational, and alternative schools. I examine a variable that indicates whether a school is new or not in a given year and a set of variables that indicate how urban the school's location is.

5. Excluding grade-by-year binary variables and structural change of school. I question whether these variables should be included, so I explore how sensitive Bifulco and Ladd's results are to excluding them.

6. Gains model and fixed effects model specification tests. Todd and Wolpin note (2003, p. F29) that the key assumptions underlying the value-added model (recall that the gains model is a restricted version of the value-added model) can be tested by including lagged input measures in the model. If the assumptions are correct, lagged input measures should have no explanatory power. For the fixed effects model, I will test whether the fixed effects are jointly equal to zero or not.

Tables 1 and 2 present some descriptive information for Bifulco and Ladd's data and the corresponding information for my data. There is a good match qualitatively between the two data sets. Two points are worth noting. My data, even after temporarily dropping the two extra years I will use later, has more observations per student for the early cohorts of students than does Bifulco and Ladd's data. I don't know why. And Table 2 shows that in my larger data set, the average test scores of charter students are even lower than in Bifulco and Ladd's data set.

Levels model. Tables 3A (math scores) and 3B (reading scores) reproduce Bifulco and Ladd's estimates of the levels model, my attempt to replicate their results, and some of my modifications. Bifulco and Ladd include as independent variables gender, three variables for race (black, Hispanic, and white), several variables for the level of parent education, nonstructural change of school, and structural change of school. The second column of each table shows my attempt to replicate Bifulco and Ladd's results. Even with more than three times as many observations, the match is quite close. The most noteworthy feature of the two replication estimates is that the effect of charter schools is even more negative than Bifulco and Ladd report (for math, the coefficient on the charter school indicator falls from -.255 to -.307; for reading, it falls from -.158 to -185).

After testing variables for a number of personal and school characteristics, I found that while many had the expected signs and were statistically significant, few affected the size or significance of the charter school effect. In the interest of space and the reader's patience, I therefore do not report estimates for most of these variables.

After testing variables for a number of personal and school characteristics, I found that while many hada the expected signs and were statistically significant, few affected the size or significance of the charter school effect.

There are two exceptions. The third columns of Tables 3A and 3B show the result of adding whether or not a student was designated "academically/intellectually gifted." The results are impressive. Adding the academically gifted variable cuts the negative effect of charter schools on math scores almost in half (-.307 to -.179) and raises the R-squared of the estimated model by nearly half (.2654 to .3810). Adding the academically gifted variable cuts the negative effect of charter schools on reading scores by more than half (-.185 to -.075) and raises the R-squared of the estimated model from .2607 to .3462. The estimated model indicates that gifted students perform quite well: even after controlling for parental education, gifted students average more than one standard deviation above the mean in math and more than .9 of a standard deviation above the mean in reading. Their success affects the estimated impact of charter schools because a lower percentage of charter students than regular public school students are gifted. Table 4 gives details. In each year, the percentage of regular school students that is gifted is at least four percentage points higher than the percentage of charter students that is gifted. By the last year of the data, 2002–03, the percentage of regular school students that is gifted, 13.62, is more than three times the percentage in charters, 4.20.

A second variable that has some impact on the charter school coefficient is a variable that designates charters that focus on at-risk students.[24] This variable enters the model negatively and statistically significantly and lowers the size of the charter school coefficient by 28 percent for math (from -.179 to -.129) and by two-thirds for reading (from -0.075 to -0.025).[25]

While students attending at-risk charter schools score lower than other charter students and considerably lower than regular public school students, neither group might be the best comparison group for such students. Table 5 shows the results of comparing at-risk charter students to just those students attending special education or alternative regular public schools. For most of the estimates, the charter school effect is now insignificantly different from zero. At-risk charter students seem to perform approximately as well as alternative school students.

The conclusion from this look at the levels model is straightforward. Adding just two relatively crude measures of the difference between charter students and regular public school students greatly reduces the negative impact of charters reported by Bifulco and Ladd. The negative effect for math scores is cut almost in half (-.255 to -.129) and the negative effect in reading is almost eliminated (-.158 to an insignificant -.025). This suggests at least the possibility that if more precise measures of the differences between charter students and regular public school students were available, the impact of charter schools would be still more positive. We would like to have a measure of whether an individual student is at risk, for example, not just whether he attends a school with a substantial at-risk population. And we would also like to know if a higher proportion of regular public school students might be "almost gifted" than students in charters.

Gains model. Tables 6A and 6B present some results for the gains model. My estimates include more than twice as many observations as Bifulco and Ladd used. Once again the replication columns show a good match with their results. Unlike the levels model, however, the replication estimates show the charter effect as less negative than Bifulco and Ladd's. For math scores, the charter coefficient rises from -.076 to -.058; for reading scores, from -.062 to -.035.

Also unlike the levels model, deleting the questionable structural change and grade-by-year binary variables has an impact worth noting. After deleting those variables, the coefficient on charter schools rises to -.039 for math scores and to -.026 for reading scores. Comparing -.076 to -.039 and -.058 to -.026, we see that my estimates eliminate about half of the negative effect of charter schools that Bifulco and Ladd reported using the gains model.

Doubt about the robustness of the negative effect of charter schools in the gains model is raised by three observations. Unlike the levels model, giftedness enters the gains model negatively (and significantly). This is surprising and raises a question about either the test scores or the model specification or both. Next, note that the R-squareds are quite low: .02 for my three estimated models (Bifulco and Ladd do not report R-squareds for any of their models). This means that these models leave an enormous amount of the variation in score gains unexplained. Given that so much is unexplained, we should treat the findings of these models very cautiously.[26] Last, Table 7 shows the result of including lagged values for the independent variables that change over time. Under the assumptions used to derive the gains model, lagged values of the independent variables should not be significant. But the model fails this specification test: for math scores, the first and second lags of the charter variable are significant; for reading, the first lag (and the third lag as well if the structural change and grade-by-year variables are removed).

(It is interesting to note that the sum of the coefficients of the charter school variables—contemporaneous plus the lagged values—is positive in Table 7 for both math and reading. We cannot necessarily attribute this to the change in specification, however. Because enough data for a given student must be present to include three lagged values, Table 7 is estimated on a much smaller number of observations. We will see a similar change when the cumulative model is estimated in Section IV below.)

Fixed effects model. Table 8 shows the results for the fixed effects model. I include about 75 percent more observations than Bifulco and Ladd. Once again, I can rather closely reproduce their result for charter school effect. Bifulco and Ladd obtain a coefficient of -.160 for math scores, while I obtain -0.138; they get -0.095 for reading, I get -0.079.

There are two indications that caution should be exercised in accepting the estimates from the fixed effects model. The "within" R-squareds for the models are truly abysmal. (A "within" R-squared is the fraction of the variance explained by the model, not including the fraction explained by the fixed effects.) It is hard to see how anyone could be confident of the inferences generated by such a poorly fitting model.

The model also fails the specification test mentioned earlier. The F-statistic for the test that the fixed effects are all jointly equal to zero is .37 for math and .34 for reading. Neither permits us to reject the null hypothesis that all the fixed effects are zero. So, statistically at least, including the fixed effects in the model is not justified. And without the fixed effects, the charter school coefficients are much less negative: -.044 for math and -.033 for reading.

IV. Two Other Models

Since the levels model, the gains model, and the fixed effects model all have shortcomings, I report the results of estimating two other models.

Cumulative model. Todd and Wolpin (2004) argue that the cumulative model is a more satisfactory specification than the levels model or the value-added model. The cumulative model is the levels model augmented by lagged values of the variables that change over time. It removes the assumption of the levels model that past inputs have zero effect on current achievement. It allows past inputs to have long-lived effects on achievement, and it does not arbitrarily restrict the time pattern of effects. However, it shares the levels model's weakness that some inputs will not be measurable and hence not included, possibly biasing the estimated coefficients of the variables that are included.

Table 9A reports the results of estimating the cumulative model, once with two lags included and once with three lags. Across the four estimated models there are fourteen coefficients for the charter variable or its lagged values; twelve of them are positive. (And one of the two negative signs switches to positive if the structural change and grade-by-year variables are omitted.) But most are not individually significant. Testing that the coefficients of the charter variable and its lags are jointly different from zero, we find that for the math-score models, we cannot reject the null hypothesis that they are all zero. But for both reading-score models, we can reject the null hypothesis at the .02 level. Charters have a positive effect, in total, on reading test scores.

Even the insignificantly-different-from-zero effect in the math-score equations contrasts with the negative effects observed in the levels, gains, and fixed effects models. But it is not clear that this difference results from the model itself. Estimating several lags requires several observations on each student; therefore, fewer students and fewer observations can be included. To explore this idea further, Table 9B reports estimates of the sample levels model but with the sample of observations restricted to be the same as the sample used for Table 9A. The charter effect on math scores is once again indistinguishable from zero, and the charter effect on reading scores is positive, and for the two smallest samples, significantly so.

These results suggest that sample composition is important: the charter effect is somewhat more positive for students we observe for longer periods.

"Stayers" model. The fixed effects model is identified by students who switch out of or into charter schools. Switchers could well be students who, as a group, match relatively poorly with charter schools: this could be true of students switching out, as evidenced by the very fact that they are leaving, and it could be true of students switching in, because they might be in some difficulty that affects test performance and that also induces them to switch. This prompts the thought: What about the students for whom charter schools might be a good fit? A model focusing on these students overestimates the positive effects of charters on students as a whole, but it serves as a useful counterweight to the fixed effects model, which probably underestimates them.

These results sugest that sample composition is important: the charter effect is somewhat more positive for students we observer for longer periods.

Table 10 reports estimates from a stayers model. The sample is restricted to students who have stayed in either the regular public school system or the charter schools—not necessarily the same school, just the same system—for exactly N years, where N varies between two and six, inclusive. The dependent variable is the change in a student's standardized test score from the first year the student is observed to the Nth year. Because students entering charter schools have lower test scores on average when they enter, we must beware of regression toward the mean. I enter the student's first-year score as an independent variable to control for this (see, for example, Barnett et al. 2004). (If the first-year scores are omitted, the effect of charter schools becomes substantially more positive.)

The charter effects are mostly positive, though their significance is mixed. If the structural change and grade-by-year variable are dropped, the effects become more positive by .01 to .02 and the negative sign for math scores, three years, becomes positive and the negative sign for math scores, two years, becomes essentially zero. Once again, though, it is not clear that the more positive results are due to the model. The effects of the model are confounded with the effects of analyzing a smaller number of students, students whom we see for longer periods.

V. What Distinguishes Poorly Performing Charter Schools?

Probably more important for public policy than determining whether charter schools on average affect academic performance for the worse, the better, or the same as regular public schools is determining what distinguishes better-performing charters from worse-performing ones. My results suggest one major factor, the same one identified in research on charter schools in other states (see p. 5 above): the age of the charter school.

Table 11A reestimates the levels, gains, and fixed effects models after disaggregating the charter binary variable into six variables: a binary variable equaling one if a school is a first-year charter, a variable equaling one for a second-year charter, and so on through sixth-year charters. First-year charters have by far the most substantial negative effects. The negative effects of second-year charters are also sizeable. But for third-year and older charters, in the levels and gains models, the effects are small and sometimes not significant. While older schools have significant negative effects in the fixed effects model, the size of the negative effect is much reduced from the simple fixed effects model. (These results are contrary to Bifulco and Ladd's [2004, p. 44]; they report large, significant negative effects for third-, fourth-, and especially fifth-year charters. They report these effects, however, only for the fixed effects model.) Table 11B shows that if first- and second-year charters are deleted from the data, the charter effect is insignificant in the levels and gains models.

My results suggest one major factor, the same one identified in research on charter schools in other states (see pg. 5 above): the age of the charter school.

I thought that the poorer results from first-year charter schools might be a vintage effect: the first cohort of charters in 1997–98 might well have struggled because charters were new and because competition for a place among the 100 permitted charters had not yet become intense. Table 12 indicates, however, that this hypothesis is incorrect. After disaggregating the data by year, the size of the first-year effect remains relatively constant. In fact, for 2002–03, the first-year effect is even more negative than for 1997–98.

I tested another possible explanation for the first-year effect: perhaps the first-year effect was due to a small number of charter schools that were mismanaged or unpopular and that subsequently closed. To investigate this possibility, I removed from the data all eighteen charter schools that had closed by the 2002–03 school year, and I then reestimated the levels, gains, and fixed effects models. The results from the reestimated models were qualitatively similar for the levels and fixed effects models. For the gains model, the negative effect of charters was smaller and was no longer significant for math. But overall, the first-year effect seems to be caused by more than just failed charters.

Obviously, determining the causes of this effect should be a top priority for further research.[27]

Based on Sass's and Bifulco and Ladd's work, I investigated one other determinant of performance, grade level. Table 13 shows the results. Charter schools seem—tentatively—to have a more negative impact on third- through fifth-graders, particularly third-graders, than on sixth- through eighth-graders. The reason for this is not at all clear and is another candidate for further research.

VI. A Preliminary Look at Exit from Charter Schools

The last bit of empirical work I present follows a good idea in the Hanushek, Kain, and Rivkin (2002) paper. They investigated whether Texas charter schools that seemed to be adversely affecting the academic performance of their students were more likely than better-performing charters to lose students. If so, given the direct tie between charter-school funding and student enrollment, poorly performing charters would have to improve or they would go out of business. Any problem of poorly performing charters would fix itself.

Table 14 reports the results of some preliminary work—I stress preliminary—toward a similar investigation for North Carolina charters. A binary variable equaling one for a student's exit from the charter system is the dependent variable. Both student characteristics and school characteristics are included as independent variables.

Some factors have the expected impact. Among the factors that decrease the probability that a student will exit the charter system are as follows: higher math and reading scores for the student, higher levels of the student's parents' education, the charter being an at-risk charter, and the charter being attended by a higher fraction of gifted students.

But other factors have surprising impacts, surprising at least to me. Charter school age is positively related to the probability that a student exits. So is the fact that a student is gifted, and so is the school's average level of the parents' education.

Charter school age is positively related to the probability that a student exits. So is the fact that a student is gifated, and so is the school's average level of the parents' education.

As for the variables of interest, if entered as the sole school score, a higher average math score for the entire school *raises* the probability of student exit (first column). So does, if entered as the sole school score, a higher average reading score (second column). But if both school-average math and reading scores are entered, student exit is positively related to the math score but negatively related to the reading score (third column).

When we take these results together, I find them more than a little puzzling. One factor that complicates the explanation of school exit is the possibility that parents are attracted to nonacademic advantages of charter schools. But such advantages might be difficult to quantify and include in a regression.

We have, yet again, another topic for further research.

VII. Conclusion

Assessing the academic impact of charter schools is difficult. I don't contend that I have done it better than Bifulco and Ladd did. I do contend that each of their three models has important weaknesses and that their conclusion should be qualified: excluding at-risk charters, the effect of charter schools on academic performance is closer to zero. But students attending at-risk charters seem to perform about as well as students in public alternative schools, quite possibly a more appropriate baseline. And there are some identifiable groups of students for which charter schools seem to work relatively well: students we observe in the data longer, students who continue in charters for several years, and students who attend charters that are at least three years old.

Further work is needed. We would benefit from more information in the following areas.

1. The beginning aptitude and characteristics of students. It would be useful to have a more finely grained classification than just "gifted" or "not gifted"—possibly the score on some other aptitude test—and it would be useful to know whether a student attending an at-risk charter school is, individually, at risk.

2. The switching process. Who switches into charter schools and why? Who leaves and why?

3. Young charter schools. The apparent persistence of a negative effect in first-year charters is curious. If more information were available, perhaps charter school administrators and faculty could take actions to ameliorate it.

4. Young students. Similarly, it would be very worthwhile to understand the apparent lower performance of third-graders. Last, as important or perhaps more important than these, we need additional measures of student success. Brian P. Gill et al. (2001, p. 113) summarize this need well: "Future studies should include measures that reflect the richer set of academic outcomes that schools are expected to produce. At the very least, researchers should examine academic attainment (including continuation in school, graduation, and college attendance) in voucher and charter schools."

Notes on the Data

The vast majority of the data used in this study was obtained in machine-readable form from the North Carolina Education Research Data Center. (Common Core data for the 2002–03 school year were an exception. These data were not available from the Center when I began my analysis, so I obtained them directly from a Department of Education web page, http://nces.ed.gov/ccd/pubschuniv.asp.) I converted the Center's SAS data sets to STATA data sets, combined school and student data sets, combined data sets for individual years, edited, and further processed the data.

The major steps in editing the data were as follows:

1. For each of 15,013 observations, there was another observation with the same student ID, year, and grade level. For each of twenty-four observations, there were two such matching observations, and for each of thirteen observations, there were three such matching observations. (In some cases, but not all, the duplication appears to have

CHARTER SCHOOLS CLASSIFIED AS TARGETING AT-RISK STUDENTS

LEA CODE	NAME	CITY	SOURCE(S) FOR AT-RISK CLASSIFICATION*
01A	Lakeside School	Elon College	(1) and (2)
06A	Grandfather Academy	Banner Elk	(1) and (2)
06B	Crossnore Academy	Crossnore	(1) and (2)
26B	Alpha Academy	Fayetteville	(2)
32B	Healthy Start Academy	Durham	(2)
32C	Carter Community Charter	Durham	(2)
32J	Ann Atwater Community Charter	Bahama	(3)
34A	Lift Academy	Winston-Salem	(4) and (5)
34E	East Winston Primary	Winston-Salem	(6)
41A	Imani Institute Charter	Greensboro	(2)
49C	Developmental Day School	Statesville	(7)
49D	Success Institute Charter	Statesville	(2)
53A	Provisions Academy	Sanford	(1)
60B	Sugar Creek Charter	Charlotte	(8)
60C	Kennedy Charter	Charlotte	(1)
68L	School in the Community	Chapel Hill	(9)
74A	Right Step Academy	Greenville	(10)
78A	CIS Academy	Lumberton	(11)
83B	Laurinburg Homework Center	Laurinburg	(1)
90A	Union Academy	Monroe	(2)
92G	East Wake Academy	Zebulon	(12)
92I	SPARC Academy	Raleigh	(13)
92Q	Hope Elementary	Raleigh	(3)
96B	Change for Youth Academy	Greensboro	(14)
97C	Wilkes County Alternative Charter	Roaring River	(15)

occurred because the student changed schools during the school year.) For such observations, I averaged the multiple math scores if multiple scores were available, did the same for reading scores, and then deleted the duplicate observations.

2. For each of 248 observations, there was another observation with the same student ID and year, but different grades. Unsure whether these were in-year promotions or recording errors, I simply deleted these observations.

3. I noted several thousand students for whom either their gender, race, or the level of their parents' education changed between third and eighth grades. Since most of the regressions run for this paper use several hundred thousand observations or more, and since gender, race, and parents' education are not the foci of the study, I expected that the errors in these observations would not affect my results, so I retained these observations.

4. For some of the alternative public schools in the data, the designation as "alternative" was missing for some years. If a school was designated an alterative school for at least one year, and it had missing values for other years, I assigned the years with missing values the alternative designation.

Two examples of the processing I performed are as follows:

1. One variable that was not obtained from the North Carolina Education Research Data Center was the binary variable designating whether a charter school focused on at-risk students or not. I created this variable by consulting multiple sources. The main sources were the minutes of the North Carolina State Board of Education (http://www.ncpublicschools.org/sbe_meetings/), school websites, Allen and Cooper (2004), and Manual and McLaughlin (2002). I classified twenty-four schools as "at-risk schools" for all the years they appeared in the data and one school as an "at-risk school' for part of the sample period. My designations are subject to error, of course. I thought to confirm the designations by consulting each school's charter documents, but Matthew Lanner at the North Carolina Office of Charter Schools, 919-807-3496, informed me that the charter documents probably would not be helpful.

2. Bifulco and Ladd's calculation of the structural change in school variable (2004, p. 12) seems unnecessarily complex. I calculated it in a simpler fashion. The Common Core of data includes information on the highest grade in each school. If a student is in the highest grade a school offers in a given year and if he does not fail that year, then next year he will make a structural change of school. (In a very small number of instances, this calculation had to be modified because the school increased its highest grade.)

*** SOURCE(S) FOR AT-RISK CLASSIFICATION**

(1) N.C. State Board of Education, "EEO7 – Recommendations of Charter Schools to Become Alternative Charter Schools," www.ncpublicschools.org/sbe_meetings/0403/0403_EEO07.pdf.

(2) Jeanne Allen and Autumn Cooper, eds., *National Charter School Directory 2004*.

(3) North Carolina Department of Public Instruction, "Charter School Cap Met with Recent Approval of Six New Schools," March 1, 2001, http://www.ncpublicschools.org/news/00-01/030101.html.

(4) Winston-Salem/Forsyth County Board of Education, Minutes, January 7, 2002, http://mts.admin.wsfcs.k12.nc.us/admin/Minutes/01-02/bm02-0107.html.

(5) Doug Haynes, "Let Charter Schools Work," *Business Leader Online*, September 1998, http://www.businessleader.com/bl/sep98/eyeongov.html.

(6) Deirdre Fernandes, "East Winston Primary Celebrates Scores," *Winston-Salem Journal*, September 16, 2002, http://www.journalnow.com.

(7) Melanie Looney, ed., *Charter School Closures: The Opportunity for Accountability,* p. 15, http://edreform.com/_upload/closures.pdf.

(8) N.C. State Board of Education, "EEO1 – Renewal Recommendations for Charter Schools Established in 1998 and Placed on a One-Year Renewal Delay," http://www.ncpublicschools.org/sbe_meetings/0212/0212_EEO01.pdf. "Students at Sugar Creek have many of the social demographics associated with school failure" (p. 15).

(9) Addresses.com website, http://www.addresses.com/schools_search_by_city/Chapel+Hill/NC/15.html. Listed as an "alternative school."

(10) Addresses.com website, http://www.addresses.com/schools_search_by_city/Greenville/NC/15.html. Listed as an "alternative school."

(11) CIS Academy of Robeson website, http://www.cisnc.org/cisrc/code/programs.htm.

(12) East Wake Academy website, http://www.eastwakeacademy.org/.

(13) John Manual and Mike McLaughlin, "The Charter School Experience in North Carolina," *North Carolina Insight*, July 2002. Describes school as dealing with "primarily at-risk" students after the 1998–99 school year.

(14) Addresses.com website, http://www.allemailaddresses.com/schools_by_city/Goldsboro,NC/11575.html. Listed as an "alternative school."

(15) Addresses.com website, http://www.addresses.com/schools_by_city/Roaring+River,NC/12309.html. Listed as an "alternative school."

Notes

[1]Hill, 2005.

[2] Loveless, 2003, p. 28. See also Solmon, Paark, and Garcia 2001, p. 2: that parents enroll students in charter schools may indicate ". . . indicates greater parent involvement in and concern for their children's education."

[3] For example, Gene V. Glass, professor of education at Arizona State University, states (Freedman, 2004):, "The majority of urban charter school students have come to charter schools as a result of failure in or dissatisfaction with a traditional public school."

[4]Hoxby, 2003, p. 46.

[5] Buddin and Zimmer (2004, p. 1): "One of the challenges of assessing charter schools is that there is no single charter school approach to educating students. By design, charter schools vary in their educational programs, curricula, instruction, and school settings." Sass (undated, p. 26): "Charter schools are quite diverse. . . ." Finn (2004): "[C]harter schools are astoundingly varied." Dickman, et al. (2003, p. 8), (referring to Wisconsin charter schools): "Taken as a whole, the range of offerings is very extensive compared to the traditional formats of most public schools."

[6] Loveless argues (2002, p. 35), "Charters are incredibly diverse. . . . The greater likelihood is that charters will be found to produce a wide range of outcomes. Some charters will be terrific places for educating children and others will be failures. Identifying the characteristics of excellent charter schools and encouraging their adoption should be the main objective of the next wave of charter school research and policy." Hassel and Terrell argue (2004, p. 16), "Charter schools are not a single kind of school, the way schools adopting a particular instructional model are. They vary greatly in their student bodies, instructional approaches, and organizational forms. As a result, it makes little sense to assess a state's charter school policy based on how well the average charter school is doing relative to the average district school [footnote omitted]."

[7] On the wide variation among states in the proportion of charters that are "targeted,", see Greene, et al., 2003, p. 15.

[8] See also Bulkley and Fisler, 2002, pp. 7—8. See Bettinger (2004), for a recent study that concludes that Michigan charter schools have either a zero or a negative effect on their students' academic performance.

[9] It is not clear to me whether Sass's earlier criticism of Gronberg and Jansen applies to this finding, too. Gronberg and Jansen do not supply a lot of detail about their methodology.

[10] Greene, Forster, and Williams Winters write (2003, pp. 6—7), "For each test score we had from a given charter school, we found the closest regular public school for which we had test scores from the same grade. We ignored regular public schools known to be targeted to particular populations, such as magnet schools and schools for juvenile delinquents. . . ." Hoxby states (p. 1), ""[C]harter schools are compared to the schools that their students would most likely otherwise attend: the nearest regular public school. . . ." She also writes (footnote 12), "The pool of potential comparison schools does not include non-regular public schools, such as alternative schools, schools for the disabled, schools that admit students based on examinations, and magnet schools to which a student must apply."

[11] Baker A. Mitchell, Jr., founder of Charter Day School in Brunswick County, observes that state law prohibits charter schools from excluding eligible children residing anywhere in North Carolina. He further notes that Charter Day School currently draws most of its students from two separate counties, and 10% are drawn from three additional counties. (Baker e-mail from Mitchell, forwarded to the author, January 5, 2005). My wife informs me that the charter school where she teaches, Raleigh Charter High School, also draws its students from five counties.

[12] McCaffrey, et al., in the context of evaluating teacher effectiveness, write (2004, p. 97), "However, given the large proportion of missing data in many achievement databases and known differences between students with complete and incomplete test data, it is possible that estimates may be highly sensitive to this (or other) assumptions about missing data."

[13] As we we'll see below, Bifulco and Ladd estimate a "gains" model that looks not at the level of test scores but at the growth of test scores. Such a model is less prone to this potential bias but not completely free of it: advantaged students not only score higher, their score growth should be higher. To the extent the statistical model does not completely control for all personal and environmental factors affecting student performance—which is virtually certain—the missing students could distort the results.

[14] One way that East Winston might have achieved this

dramatic improvement—a way that would not be reflect as well on its teachers and administrators—would be if it discouraged weaker students from reenrolling and replaced them with stronger students. I examined this possibility. There were 19 East Winston students in the data for the 2000–-01 school year. Their average normalized test score was -2.65 (that is, two and two-thirds standard deviations below the mean). The following year, there were 19 new students at East Winston; their average normalized test score was a much better -0.71. That seems consistent with the student-turnover hypothesis. However, 9 of the 19 students from 2000–-01 also reenrolled in East Winston and, importantly, their scores improved significantly, too: from an average normalized score of -2.91 to -1.10. Thus, the improvement in the reenrolling students' average score, 1.81 standardized units, thus almost equals the difference between the new students' average score and the old students' average score, 1.94 standardized units. I conclude that student turnover doesn't explain East Winston's remarkable one-year improvement.

[15] It should be noted that in his study of Florida charters, Sass compares charter schools to regular public schools using scores from a test that is *not* used for accountability, thereby minimizing any possible bias from "teaching to the test" (2004, p. 14). Unfortunately, it is not seem possible to do this for third- through eighth- graders in North Carolina schools (the grades that Bifulco and Ladd, and I, examine).

[16] The notation is that of Sass (undated), which is similar to that of Todd and Wolpin (2003).

[17] The primes (′) associated with η and ε indicate that these terms are not of the same form as in equation (3), but this does not seem to have consequences for Ladd and Bifulco's estimation of (4).

[18] In this context, it is interesting to note that many school districts across the nation are experimenting with replacing middle schools with K–-8 schools in an attempt to improve education in the middle-school grades (Herszenhorn, 2004). North Carolina charter schools have already moved a significant degree in that direction.

[19] There is is a reason, of course, for doing this. It ensures that any differences in results among the three methods can be ascribed to the methods themselves, rather than to differences in the samples. But the cost of discarding data seems too high a price to pay for this advantage.

[20] Baker A. Mitchell, Jr. reports that Charter Day School currently has a "sizeable waiting list". (eBaker e-mail from Mitchell, forwarded to the author, January 5, 2005). Noblit and Corbett report (2001, p. III-6) that in 2000, 52% of North Carolina charter schools had waiting lists, although 85% of the waiting lists had 50 or fewer students.

[21] It's true that most of the evidence that Bast and Walberg cite is not about parents choosing charter schools. But evidence about parents' choice of private schools and use of vouchers should be relevant as the processes are similar.

[22] That the value-added model is a generalization of the gains model is easily shown. If the coefficient of A_{it-1} is restricted to equal 1, we obtain the gains model.

[23] For more on the limitations of the value-added model, see Todd and Wolpin (2003).

[24] For details on the construction of this variable, see the section "Notes on the Data" below (it follows the Conclusion).

[25] Removing the questionable structural change and grade-by-year binary variables decreased the size of the charter school effect still further—and for reading the deletions also lowered the coefficient's standard error—but these changes were all small.

[26] A number of researchers have noted that year-to-year changes in student test scores are quite noisy and must be used with great care when drawing policy conclusions. See, for example, Kane and Steiger (2002).

[27] Baker A. Mitchell, Jr., founder of Charter Day School in Brunswick County, suggests the following possible reasons for the relatively weaker EOG performance of first-year charters: in later years, charter schools' curricula are changed to better align with the EOG tests; charter schools learn to add supplemental test preparation programs to their daily schedules; and in the charter schools' early years administrators are busy with regulatory issues that reduce the time they can spend focusing on academic issues. (eE-mail from Mitchell, from Baker forwarded to the author, January 5, 2005).

Tables

TABLE 1: NUMBER OF STUDENTS AND OBSERVATIONS IN EACH COHORT (USING BIFULCO AND LADD'S TABLE 3)

	COHORT96	COHORT97	COHORT98	COHORT99	COHORT00
Number of students					
Bifulco & Ladd	93,349	98,404	102,869	105,292	106,106
Newmark	93,348	98,372	102,868	105,290	106,104
Avg. obs./student					
Bifulco & Ladd	4.9	5.3	4.5	3.7	2.8
Newmark	5.7	5.8	4.8	3.9	3.0
Avg. obs./student with valid reading scores					
Bifulco & Ladd	4.8	5.1	4.3	3.5	2.7
Newmark	5.6	5.5	4.6	3.7	2.8
Avg. obs./student with valid math scores					
Bifulco & Ladd	4.8	5.1	4.4	3.6	2.7
Newmark	5.6	5.6	4.7	3.8	2.8
Number of students observed at least once in a charter school					
Bifulco & Ladd	1145	1603	2009	2181	2035
Newmark	1172	1602	2008	2188	2047
Number of students observed at least once in a charter school and at least once in a traditional public school					
Bifulco & Ladd	1145	1603	1788	1794	1335
Newmark	1172	1602	1797	1809	1367

In the analysis in this paper, data from the 1994–95 and 2002–03 school years are also used, but are not included in the computations for this table, to allow a more direct comparison to Bifulco and Ladd's sample.

TABLE 2: DESCRIPTIVE STATISTICS FOR THE STUDY SAMPLE (USING BIFULCO AND LADD'S TABLE 4)

		ONLY IN TRADITIONAL PUBLIC SCHOOLS	AT LEAST ONCE IN A CHARTER SCHOOL	IN CHARTER AND IN TRADITIONAL PUBLIC SCHOOL
Ethnicity	% Black			
	Bifulco & Ladd	31.3	43.4	44.4
	Newmark	30.0	41.2	44.0
	% Hispanic			
	Bifulco & Ladd	3.1	1.0	0.9
	Newmark	3.9	1.2	1.2
	% White			
	Bifulco & Ladd	61.3	51.2	50.4
	Newmark	61.7	53.1	50.4
Parent Education	% Less than HS			
	Bifulco & Ladd	10.1	5.3	5.5
	Newmark	11.3	5.9	6.6
	% HS graduate + some college			
	Bifulco & Ladd	48.6	42.7	42.1
	Newmark	49.6	42.7	44.9
	% 2--year college deg.			
	Bifulco & Ladd	13.8	13.4	13.8
	Newmark	12.8	12.9	12.8
	% 4--year college deg.			
	Bifulco & Ladd	22.4	31.8	30.7
	Newmark	21.2	30.8	28.8
	% Grad. school degree			
	Bifulco & Ladd	5.1	6.9	7.9
	Newmark	5.1	7.8	6.9
Average reading score	1998			
	Bifulco & Ladd	0.001	-0.040	-0.032
	Newmark	0.001	-0.109	-0.149
	1999			
	Bifulco & Ladd	0.001	-0.068	-0.040
	Newmark	0.002	-0.117	-0.152
	2000			
	Bifulco & Ladd	0.002	-0.136	-0.103
	Newmark	0.003	-0.159	-0.181
	2001			
	Bifulco & Ladd	0.002	-0.100	-0.109
	Newmark	0.003	-0.130	-0.156
	2002			
	Bifulco & Ladd	0.003	-0.145	-0.147
	Newmark	0.003	-0.146	-0.161
Average math score	1998			
	Bifulco & Ladd	0.002	-0.158	-0.134
	Newmark	0.003	-0.212	-0.244
	1999			
	Bifulco & Ladd	0.003	-0.182	-0.144
	Newmark	0.004	-0.225	-0.251
	2000			
	Bifulco & Ladd	0.003	-0.263	-0.201
	Newmark	0.005	-0.272	-0.284
	2001			
	Bifulco & Ladd	0.003	-0.180	-0.179
	Newmark	0.005	-0.221	-0.231
	2002			
	Bifulco & Ladd	0.004	-0.208	-0.229
	Newmark	0.005	-0.240	-0.229

In the analysis in this paper, data from the 1994–95 and 2002–03 school years are also used, but are not included in the computations for this table, to allow a more direct comparison to Bifulco and Ladd's sample.

TABLE 3A: LEVELS MODEL FOR MATH SCORES

	BIFULCO AND LADD (P. 43)	REPLICATION	PLUS GIFTED	PLUS GIFTED AND AT RISK
Charter School	-0.255** (0.073)	-0.307** (0.055)	-0.179** (0.052)	-0.129* (0.056)
Gender (Female=1)	0.036** (0.002)	0.043** (0.002)	0.026** (0.002)	0.026** (0.002)
Black	-0.464** (0.023)	-0.478** (0.019)	-0.417** (0.016)	-0.417** (0.016)
Hispanic	-0.046 (0.024)	-0.092 (0.020)	-0.057** (0.017)	-0.058** (0.017)
White	0.155** (0.023)	0.160** (0.019)	0.129** (0.016)	0.128** (0.016)
High school	0.386** (0.005)			
Some college	0.603** (0.006)			
High school or some college		0.424** (0.004)	0.382** (0.004)	0.381** (0.004)
2-year college	0.705** (0.006)	0.723** (0.005)	0.617** (0.004)	0.616** (0.004)
4-year college	1.076** (0.008)	1.068** (0.007)	0.812** (0.006)	0.812 (0.006)
Graduate school	1.404** (0.014)	1.374** (0.012)	0.959** (0.011)	0.959** (0.011)
Change schools	-0.160** (0.005)	-0.141** (0.005)	-0.103** (0.004)	-0.102** (0.004)
Structural change	-0.044** (0.008)	0.002 (0.008)	-0.015* (0.008)	-0.015* (0.008)
Gifted			1.075** (0.005)	1.075** (0.005)
At--risk charter				-0.326** (0.089)
R-squared		.2654	.3810	.3811
Total observations	1,533,367	4,717,034	4,711,977	4,711,977
Total students	446,855	1,802,612	1,800,671	1,800,671
Total schools		2123	2123	2123

** Indicates statistical significance at .01 level, two-tail test. * Indicates statistical significance at .05 level, two-tail test.

The dependent variable is the End-of-Grade test score, normalized. All models include grade-by-year fixed effects. The figures in parentheses are robust standard errors, using a generalized Huber/White/Sandwich estimator (for my estimates computed using Stata's "robust cluster()" option), and are robust to clustering within schools.

TABLE 3B: LEVELS MODEL FOR READING SCORES

	BIFULCO AND LADD (P. 43)	REPLICATION	PLUS GIFTED	PLUS GIFTED AND AT RISK
Charter School	-0.158** (0.044)	-0.185** (0.035)	-0.075* (0.052)	-0.025 (0.032)
Gender (Female=1)	0.174** (0.002)	0.196** (0.002)	0.182** (0.002)	0.182** (0.002)
Black	-0.351** (0.023)	-0.344** (0.018)	-0.291** (0.016)	-0.291** (0.016)
Hispanic	-0.002 (0.025)	-0.066** (0.019)	-0.037* (0.017)	-0.037* (0.017)
White	0.235** (0.023)	0.239** (0.018)	0.213** (0.016)	0.212** (0.016)
High school	0.444** (0.005)			
Some college	0.679** (0.006)			
High school or some college		0.484** (0.004)	0.447** (0.004)	0.447** (0.004)
2-year college	0.784** (0.006)	0.799** (0.004)	0.707** (0.004)	0.707** (0.004)
4-year college	1.130** (0.008)	1.120** (0.006)	0.900** (0.006)	0.899** (0.006)
Graduate school	1.419** (0.011)	1.384** (0.010)	1.029** (0.009)	1.028** (0.009)
Change schools	-0.133** (0.005)	-0.137** (0.004)	-0.104** (0.004)	-0.103** (0.004)
Structural change	-0.048** (0.007)	-0.020** (0.007)	-0.035** (0.006)	-0.035** (0.006)
Gifted			0.923** (0.005)	0.923** (0.005)
At-risk charter				-0.324** (0.080)
R-squared		.2607	.3462	.3463
Total observations	1,527,157	4,704,054	4,699,046	4,699,046
Total students	445,562	1,800,441	1,798,491	1,798,491
Total schools		2122	2122	2122

** Indicates statistical significance at .01 level, two-tail test. * Indicates statistical significance at .05 level, two-tail test.

The dependent variable is the End-of-Grade test score, normalized. All models include grade-by-year fixed effects. The figures in parentheses are robust standard errors, using a generalized Huber/White/Sandwich estimator (for my estimates computed using Stata's "robust cluster()" option), and are robust to clustering within schools.

TABLE 4: PERCENTAGE OF STUDENTS IDENTIFIED AS ACADEMICALLY/INTELLECTUALLY GIFTED

	CHARTER SCHOOLS	REGULAR PUBLIC SCHOOLS
1997-98	7.90	12.22
1998-99	5.57	12.96
1999-00	5.39	13.52
2000-01	5.35	13.44
2001-02	4.39	13.66
2002-03	4.20	13.62

TABLE 5: SAMPLE RESTRICTED TO STUDENTS IN AT-RISK CHARTERS AND SPECIAL ED. OR ALTERNATIVE PUBLIC SCHOOLS

	Levels Model		Gains Model		Fixed Effects Model	
	Math	Reading	Math	Reading	Math	Reading
Charter School	0.031 (0.093)	0.090 (0.092)	-0.015 (0.035)	-0.010 (0.026)	-0.007 (0.172)	0.289 (0.191)
Gender (Female=1)	0.066** (0.018)	0.244** (0.018)	0.010 (0.009)	0.039** (0.012)		
Black	-0.307** (0.044)	-0.233** (0.085)	-0.096 (0.075)	-0.148* (0.076)		
Hispanic	-0.108 (0.061)	-0.081 (0.101)	-0.163* (0.082)	-0.136 (0.089)		
White	0.137** (0.075)	0.249* (0.099)	-0.132 (0.074)	-0.153* (0.073)		
High school or some college	0.244** (0.030)	0.341** (0.029)	0.029 (0.024)	0.053* (0.031)		
2-year college	0.410** (0.043)	0.581** (0.040)	-0.000 (0.026)	0.052 (0.031)		
4-year college	0.669** (0.064)	0.871** (0.054)	0.045 (0.032)	0.105** (0.030)		
Graduate school	0.877** (0.077)	1.125** (0.062)	0.004 (0.035)	0.115** (0.043)		
Change schools	0.006 (0.042)	0.023 (0.041)	-0.156** (0.024)	-0.061** (0.023)	-0.020 (0.097)	0.052 (0.121)
Structural change	0.022 (0.043)	0.023 (0.042)	-0.349** (0.039)	-0.219** (0.053)	-0.192 (0.155)	-0.090 (0.140)
Gifted	1.175** (0.094)	1.029** (0.077)	0.070* (0.031)	-0.053* (0.023)		
R-squared	.3772	.3201	.0662	.0030	.6942	.6910
Total observations	21,086	21,065	14,304	14,277	13,452	13,418
Total students	16,405	16,377	11,234	11,198	10,083	10,046
Total schools	105	105	100	100		

** Indicates statistical significance at .01 level, two-tail test. * Indicates statistical significance at .05 level, two-tail test.

For the levels model, the dependent variable is the End-of-Grade test score, normalized. For the gains and fixed effects models, it is the change in the normalized End-of-Grade test score. All models include grade-by-year fixed effects. The figures in parentheses are robust standard errors, using a generalized Huber/White/Sandwich estimator, and are robust to clustering within schools.

TABLE 6A: GAINS MODEL FOR MATH SCORES

	BIFULCO AND LADD (P. 43)	REPLICATION	PLUS GIFTED	PLUS GIFTED AND AT RISK
Charter School	-0.076** (0.021)	-0.058** (0.017)	-0.059** (0.017)	-0.056** (0.019)
Gender (Female=1)	0.009** (0.001)	0.008** (0.001)	0.009** (0.001)	0.009** (0.001)
Black	-0.019** (0.005)	-0.022** (0.005)	-0.023** (0.005)	-0.023** (0.005)
Hispanic	0.020** (0.006)	0.030** (0.005)	0.029** (0.005)	0.029** (0.005)
White	-0.020** (0.005)	-0.030** (0.004)	-0.029** (0.004)	-0.029** (0.004)
High school	-0.007** (0.002)			
Some college	0.005 (0.003)			
High school or some college		-0.002 (0.002)	-0.001 (0.002)	-0.001 (0.002)
2-year college	0.004 (0.003)	0.004 (0.002)	0.005 (0.002)	0.005* (0.002)
4-year college	0.029** (0.003)	0.015** (0.003)	0.018** (0.003)	0.018** (0.003)
Graduate school	0.058** (0.004)	0.029** (0.003)	0.034** (0.003)	0.034** (0.003)
Change schools	-0.030** (0.004)	-0.000 (0.003)	-0.001 (0.003)	-0.001 (0.003)
Structural change	-0.068** (0.008)	-0.113** (0.008)	-0.113** (0.008)	-0.113** (0.008)
Gifted			-0.011** (0.002)	-0.011** (0.002)
At-risk charter				-0.021 (0.038)
R-squared		.0232	.0232	.0232
Total observations	1,520,132	2,912,699	2,910,903	2,910,903
Total students	443,548	1,002,100	1,001,917	1,001,917
Total schools		2054	2054	2054

** Indicates statistical significance at .01 level, two-tail test. * Indicates statistical significance at .05 level, two-tail test.

The dependent variable is the change in the normalized End-of-Grade test score.All models include grade-by-year fixed effects. The figures in parentheses are robust standard errors, using a generalized Huber/White/Sandwich estimator, and are robust to clustering within schools.

TABLE 6B: GAINS MODEL FOR READING SCORES

	BIFULCO AND LADD (P. 43)	REPLICATION	PLUS GIFTED	PLUS GIFTED AND AT RISK
Charter School	-0.062** (0.009)	-0.035** (0.008)	-0.041* (0.008)	-0.042** (0.008)
Gender (Female=1)	0.001 (0.001)	0.001* (0.001)	0.002** (0.001)	0.002** (0.001)
Black	-0.029** (0.004)	-0.026** (0.003)	-0.029** (0.004)	-0.029** (0.004)
Hispanic	0.041** (0.005)	0.040** (0.004)	0.038** (0.003)	0.038** (0.004)
White	-0.011** (0.004)	-0.018** (0.003)	-0.017** (0.003)	-0.017** (0.003)
High school	0.005* (0.002)			
Some college	0.016** (0.003)			
High school or some college		0.004** (0.002)	0.006** (0.002)	0.006** (0.002)
2-year college	0.016** (0.002)	0.008** (0.002)	0.013** (0.002)	0.013** (0.002)
4-year college	0.022** (0.002)	0.010** (0.002)	0.022** (0.002)	0.022** (0.002)
Graduate school	0.027** (0.003)	0.012** (0.002)	0.031** (0.002)	0.031** (0.002)
Change schools	-0.018** (0.003)	0.003 (0.002)	0.002 (0.002)	0.002 (0.002)
Structural change	-0.065** (0.006)	-0.099** (0.006)	-0.099** (0.006)	-0.099** (0.006)
Gifted			-0.044** (0.002)	-0.044** (0.002)
At-risk charter				0.007 (0.029)
R-squared		.0164	.0170	.0170
Total observations	1,512,587	2,900,270	2,898,512	2,898,512
Total students	441,863	998,885	998,704	998,704
Total schools		2054	2054	2054

** Indicates statistical significance at .01 level, two-tail test. * Indicates statistical significance at .05 level, two-tail test.

The dependent variable is the change in the normalized End-of-Grade test score. All models include grade-by-year fixed effects. The figures in parentheses are robust standard errors, using a generalized Huber/White/Sandwich estimator, and are robust to clustering within schools.

TABLE 7: GAINS MODEL INCLUDING LAGGED VARIABLES

	MATH	READING
Charter School	-0.151** (0.031)	-0.109** (0.015)
Charter School, 1 Lag	0.134** (0.022)	0.088** (0.014)
Charter School, 2 Lags	0.049* (0.022)	0.018 (0.013)
Charter School, 3 Lags	-0.002 (0.019)	0.022 (0.012)
Gender (Female=1)	0.009** (0.001)	0.015** (0.001)
Black	-0.033** (0.007)	-0.029** (0.005)
Hispanic	0.014 (0.008)	0.023** (0.006)
White	-0.028** (0.007)	-0.025** (0.003)
High school or some college	0.001 (0.003)	0.002 (0.003)
2-year college	0.004 (0.003)	-0.003 (0.003)
4-year college	0.010* (0.004)	0.003 (0.003)
Graduate school	0.015** (0.004)	0.005 (0.004)
Change schools	-0.003 (0.005)	0.011** (0.004
Change schools, 1 lag	-0.003 (0.003)	-0.002 (0.003)
Change schools, 2 lags	-0.002 (0.003)	-0.005* (0.002)
Change schools, 3 lags	0.004 (0.003)	-0.009** (0.002)
Structural change	-0.094** (0.009)	-0.044** (0.004)
Structural change, 1 lag	0.010 (0.007)	-0.032** (0.004)
Structural change, 2 lags	0.013 (0.008)	-0.038 (0.004)
Structural change, 3 lags	0.040** (0.010)	-0.008 (0.006)
Gifted	-0.012** (0.002)	-0.036** (0.002)
At-risk charter	0.047 (0.067)	0.090 (0.052)
At-risk, 1 lag	-0.036 (0.037)	-0.092** (0.028)
At-risk, 2 lags	-0.011 (0.033)	0.036 (0.037)
At-risk, 3 lags	0.013 (0.031)	0.017 (0.033)
R-squared	.0064	.0065
Total obs.	1,196,799	1,192,257
Total schools	1883	1882

** Indicates statistical significance at .01 level, two-tail test. * Indicates statistical significance at .05 level, two-tail test.

The dependent variable is the change in the normalized End-of-Grade test score. All models include grade-by-year fixed effects. The figures in parentheses are robust standard errors, using a generalized Huber/White/Sandwich estimator, and are robust to clustering within schools.

TABLE 8: FIXED EFFECTS MODEL

	BIFULCO AND LADD (P. 43)	NEWMARK	BIFULCO AND LADD (P. 43)	NEWMARK
	Math Scores	**Math Scores**	**Reading Scores**	**Reading Scores**
Charter School	-0.160** (0.021)	-0.138** (0.020)	-0.095** (0.014)	-0.079** (0.011)
Change schools	-0.027** (0.005)	0.017 (0.004)	-0.013** (0.004)	0.019** (0.003)
Structural change	-0.061** (0.010)	-0.022** (0.007)	-0.056** (0.007)	-0.015** (0.004)
At-risk charter		0.040 (0.051)		0.042 (0.045)
R-squared, "within"		.0006		.0003
R-squared, total		.1239		.1163
F-statistic for fixed effects = 0		.366		.341
Total observations	1,502,339	2,674,927	1,494,885	2,662,357
Total students	425,654	743,362	424,066	740,216

** Indicates statistical significance at .01 level, two-tail test. * Indicates statistical significance at .05 level, two-tail test.

The dependent variable is the change in the normalized End-of-Grade test score. Bifulco and Ladd models include grade-by-year fixed effects and the figures in parentheses are robust standard errors, using a generalized Huber/White/Sandwich estimator, and are robust to clustering within schools. Newmark models do not include grade-by-year fixed effects—my computer lacked enough memory to include them—but do also use robust standard errors.

Samples for both are restricted to students who have at least three valid test scores, so that at least two score gains can be computed.

TABLE 9A: CUMULATIVE MODEL

	TWO LAGS		THREE LAGS	
	Math	Reading	Math	Reading
Charter School	-0.014 (0.056)	0.038 (0.039)	0.006 (0.044)	0.051 (0.039)
Charter School, 1 Lag	0.018 (0.025)	0.032* (0.015)	0.047 (0.027)	0.057** (0.018)
Charter School, 2 Lags	0.022 (0.020)	0.012 (0.015)	0.058 (0.037)	0.034* (0.017)
Charter School, 3 Lags			0.008 (0.025)	-0.006 (0.022)
Gender (Female=1)	0.023** (0.002)	0.162** (0.002)	0.023** (0.003)	0.165** (0.003)
Black	-0.400** (0.020)	-0.287** (0.020)	-0.406** (0.023)	-0.297** (0.025)
Hispanic	-0.065** (0.021)	-0.040 (0.022)	-0.043 (0.025)	0.003 (0.028)
White	0.104** (0.019)	0.204** (0.020)	0.101** (0.022)	0.194** (0.025)
High school or some college	0.346** (0.005)	0.422** (0.005)	0.330** (0.006)	0.406** (0.006)
2-year college	0.564** (0.006)	0.668** (0.006)	0.547** (0.007)	0.651** (0.008)
4-year college	0.739** (0.008)	0.835** (0.008)	0.707** (0.010)	0.808** (0.010)
Graduate school	0.883** (0.013)	0.958** (0.012)	0.852** (0.016)	0.925** (0.014)
Change schools	-0.112** (0.005)	-0.104** (0.005)	-0.111** (0.008)	-0.104** (0.007)
Change schools, 1 lag	-0.083** (0.004)	-0.074** (0.003)	-0.125** (0.041)	-0.079** (0.004)
Change schools, 2 lags	-0.088** (0.004)	-0.086** (0.004)	-0.073** (0.004)	-0.066** (0.004)
Change schools, 3 lags			-0.085** (0.005)	-0.089** (0.004)
Structural change	-0.028* (0.011)	-0.033** (0.007)	0.009 (0.013)	-0.008 (0.010)
Structural change, 1 lag	-0.028** (0.009)	-0.022** (0.007)	-0.006 (0.011)	-0.012 (0.008)
Structural change, 2 lags	-0.039** (0.010)	-0.027** (0.007)	-0.050** (0.013)	-0.036** (0.009)
Structural change, 3 lags			-0.152** (0.018)	-0.134** (0.016)
Gifted	1.111** (0.006)	0.928** (0.006)	1.121** (0.009)	0.924** (0.008)
At-risk charter	-0.326** (0.083)	-0.292** (0.088)	-0.343** (0.071)	-0.317** (0.007)
At-risk, 1 lag	-0.116** (0.037)	-0.112** (0.034)	-0.125** (0.041)	-0.109** (0.040)
At-risk, 2 lags	-0.096** (0.033)	-0.071* (0.029)	-0.097* (0.048)	-0.069* (0.034)
At-risk, 3 lags			-0.087 (0.049)	-0.069 (0.055)
F-statistic for charter and all lagged charter coefficients = 0	F(3, 1996) = .60 Prob > F = .62	F(4, 1883) = 1.20 Prob > F = .31	F(3, 1996) = 3.33 Prob > F = .019	F(4, 1883) = 5.31 Prob > F = .0003
R-squared	.4202	.3708	.4320	.3777
Total obs.	1,956,620	1,951,098	1,210,120	1,207,552

** Indicates statistical significance at .01 level, two-tail test. * Indicates statistical significance at .05 level, two-tail test.

The dependent variable is the End-of-Grade test score, normalized. All models include grade-by-year fixed effects. The figures in parentheses are robust standard errors, using a generalized Huber/White/Sandwich estimator, and are robust to clustering within schools.

TABLE 9B: USING SAMPLES THAT MATCH THE SAMPLES FOR THE CUMULATIVE MODEL

	SAMPLE FOR 2 LAGS		SAMPLE FOR 3 LAGS		SAMPLE FOR 4 LAGS	
	Math	**Reading**	**Math**	**Reading**	**Math**	**Reading**
Charter School	-0.005 (0.064)	0.053 (0.035)	0.045 (0.062)	0.088* (0.035)	0.073 (0.074)	0.104** (0.037)
Gender (Female=1)	0.024** (0.002)	0.162** (0.002)	0.024** (0.003)	0.166** (0.003)	0.013** (0.003)	0.159** (0.003)
Black	-0.405** (0.019)	-0.291** (0.020)	-0.414** (0.022)	-0.304** (0.024)	-0.407** (0.024)	-0.310** (0.029)
Hispanic	-0.064** (0.021)	-0.040 (0.022)	-0.042 (0.025)	0.004 (0.027)	-0.016 (0.027)	0.046 (0.031)
White	0.107** (0.019)	0.207** (0.019)	0.106** (0.021)	0.199** (0.024)	0.102** (0.023)	0.191** (0.028)
High school or some college	0.350** (0.005)	0.424** (0.005)	0.338** (0.006)	0.412** (0.006)	0.329** (0.007)	0.411** (0.008)
2-year college	0.569** (0.006)	0.671** (0.006)	0.557** (0.007)	0.658** (0.008)	0.545** (0.008)	0.656** (0.009)
4-year college	0.744** (0.008)	0.839** (0.008)	0.719** (0.010)	0.817** (0.010)	0.698** (0.011)	0.803** (0.011)
Graduate school	0.890** (0.013)	0.962** (0.012)	0.864** (0.016)	0.935** (0.014)	0.850** (0.017)	0.918** (0.015)
Change schools	-0.122** (0.006)	-0.112** (0.005)	-0.128** (0.008)	-0.118** (0.007)	-0.138** (0.009)	-0.125** (0.008)
Structural change	-0.021* (0.010)	-0.029** (0.008)	0.006 (0.011)	-0.009 (0.009)	-0.021 (0.020)	-0.054** (0.017)
Gifted	1.115** (0.007)	0.931** (0.006)	1.127** (0.009)	0.929** (0.008)	1.136** (0.009)	0.912** (0.008)
At-risk charter	-0.410** (0.091)	-0.368** (0.089)	-0.447** (0.085)	-0.406** (0.088)	-0.442** (0.088)	-0.387** (0.082)
R-squared	.4190	.3716	.4294	.3755	.4410	.3768
Total obs.	1,956,620	1,948,220	1,210,120	1,205,661	687,412	684,911

** Indicates statistical significance at .01 level, two-tail test. * Indicates statistical significance at .05 level, two-tail test.

The dependent variable is the End-of-Grade test score, normalized. All models include grade-by-year fixed effects. The figures in parentheses are robust standard errors, using a generalized Huber/White/Sandwich estimator, and are robust to clustering within schools.

TABLE 10: SAMPLE RESTRICTED TO STUDENTS WHO HAVE STAYED EITHER IN REGULAR PUBLIC SCHOOLS OR IN CHARTER SCHOOLS FOR SIX, FIVE, FOUR, THREE, OR TWO CONSECUTIVE YEARS

	Two Years		Three Years		Four Years		Five Years		Six Years	
	Math	Reading	Math	Reading	Math	Reading	Math	Reading	Math	Reading
Charter School	-0.022 (0.029)	0.051** (0.018)	-0.010 (0.024)	0.001 (0.021)	0.065 (0.052)	0.094** (0.032)	0.238* (0.093)	0.202** (0.041)	0.134 (0.104)	0.029 (0.144)
First-Year Math	-0.237** (0.002)		-0.268** (0.002)		-0.339** (0.002)		-0.387** (0.003)		-0.401** (0.002)	
First-Year Reading		-0.279** (0.002)		-0.306** (0.002)		-0.351** (0.002)		-0.381** (0.002)		-0.393** (0.002)
Gender (Female = 1)	0.020** (0.002)	0.051** (0.002)	0.011** (0.002)	0.045** (0.002)	0.048** (0.003)	0.070** (0.003)	0.041** (0.003)	0.071** (0.003)	0.009** (0.002)	0.060** (0.002)
Black	-0.132** (0.009)	-0.113** (0.007)	-0.137** (0.011)	-0.118** (0.010)	-0.184** (0.013)	-0.159** (0.011)	-0.213** (0.013)	-0.209** (0.010)	-0.253** (0.013)	-0.210** (0.011)
Hispanic	0.025* (0.010)	0.011 (0.009)	0.047** (0.013)	0.068** (0.011)	0.044** (0.015)	0.080** (0.013)	0.028 (0.016)	0.077** (0.013)	0.040* (0.016)	0.122** (0.013)
White	-0.027** (0.008)	0.023** (0.006)	-0.020 (0.010)	0.048** (0.010)	-0.038** (0.013)	0.024* (0.010)	-0.046** (0.014)	-0.018 (0.010)	-0.050** (0.013)	0.025* (0.011)
High school or some col.	0.086** (0.004)	0.116** (0.004)	0.099** (0.004)	0.130** (0.004)	0.125** (0.004)	0.152** (0.004)	0.124** (0.004)	0.169** (0.005)	0.137** (0.003)	0.122** (0.013)
2-year college	0.140** (0.005)	0.187** (0.0050	0.170** (0.005)	0.212** (0.005)	0.217** (0.006)	0.245** (0.005)	0.223** (0.005)	0.267** (0.006)	0.236** (0.004)	0.284** (0.004)
4-year college	0.187** (0.005)	0.252** (0.005)	0.240** (0.006)	0.284** (0.005)	0.337** (0.007)	0.338** (0.005)	0.358** (0.007)	0.362** (0.006)	0.339** (0.006)	0.371** (0.005)
Graduate school	0.228** (0.007)	0.303** (0.006)	0.292** (0.007)	0.341** (0.007)	0.424** (0.009)	0.412** (0.007)	0.488** (0.009)	0.430** (0.007)	0.446** (0.008)	0.440** (0.006)
Gifted	0.196** (0.004)	0.231** (0.004)	0.251** (0.004)	0.226** (0.004)	0.358** (0.005)	0.252** (0.004)	0.468** (0.006)	0.250** (0.004)	0.461** (0.005)	0.266** (0.003)
At-risk charter	-0.044 (0.062)	-0.064 (0.054)	-0.131 (0.090)	-0.097 (0.077)	-0.146 (0.112)	-0.229 (0.132)	-0.535** (0.115)	-0.388** (0.073)		
Change schools	-0.030** (0.006)	-0.029** (0.004)	-0.035** (0.005)	-0.027** (0.004)	-0.047** (0.005)	-0.037** (0.004)	-0.044** (0.006)	-0.047** (0.004)	-0.055** (0.005)	-0.045** (0.003)
Structural change	-0.115** (0.019)	-0.088** (0.013)	-0.103** (0.013)	-0.066** (0.009)	-0.043** (0.009)	-0.047** (0.006)	-0.017 (0.009)	-0.027** (0.007)	-0.011 (0.008)	-0.016** (0.005)
R-squared	.1357	.1562	.1436	.1648	.1967	.1953	.2273	.2082	.2113	.2118
Total observations	492,710	490,852	643,473	641,140	499,746	497,342	559,939	557,626	1,492,789	1,488,327

** Indicates statistical significance at .01 level, two-tail test. * Indicates statistical significance at .05 level, two-tail test.

The dependent variable is the change in the normalized End-of-Grade test score, from the student's first year in either a regular public school or a charter school to the student's last year in either a regular public school or a charter school. All models include grade-by-year fixed effects. The figures in parentheses are robust standard errors, using a generalized Huber/White/Sandwich estimator, and are robust to clustering within schools.

TABLE 11A: EFFECT OF CHARTER SCHOOL AGE

	LEVELS MODEL		GAINS MODEL		FIXED EFFECTS MODEL	
	Math	**Reading**	**Math**	**Reading**	**Math**	**Reading**
1st Year Charter	-0.230** (0.055)	-0.073 (0.038)	-0.206** (0.042)	-0.132** (0.024)	-0.308** (0.044)	-0.192** (0.027)
2nd Year Charter	-0.087** (0.027)	-0.025 (0.018)	-0.022* (0.012)	-0.021* (0.008)	-0.055** (0.013)	-0.033** (0.009)
3rd Year Charter	-0.045* (0.019)	-0.016 (0.011)	-0.005 (0.007)	-0.015** (0.004)	-0.027** (0.008)	-0.023** (0.005)
4th Year Charter	-0.016 (0.018)	0.004 (0.010)	-0.008 (0.007)	-0.003 (0.004)	-0.023** (0.008)	-0.008 (0.005)
5th Year Charter	-0.007 (0.018)	0.003 (0.010)	-0.006 (0.004)	-0.008* (0.004)	-0.019** (0.007)	-0.012* (0.005)
6th Year Charter	-0.011 (0.014)	0.007 (0.008)	-0.005 (0.007)	0.006 (0.006)	-0.010 (0.010)	0.012 (0.008)
R-squared	.3812	.3463	.0233	.0171	.1241	.1172
Total observations	4,711,977	4,699,046	2,910,903	2,898,512	2,674,927	2,661,677
Total students	1,800,671	1,798,491	1,001,917	998,704	743,363	741,268

** Indicates statistical significance at .01 level, two-tail test. * Indicates statistical significance at .05 level, two-tail test.

For the levels model, the dependent variable is the End-of-Grade test score, normalized. For the gains and fixed effects models, the dependent variable is the change in the normalized End-of-Grade test score. The levels and gains models include grade-by-year fixed effects and the other independent variables shown in the "Plus Gifted and At Risk" columns of Tables 3 and 6. The fixed effects model includes three other independent variables: "at-risk charter", "nonstructural change of school," and "structural change of school." For all models, the figures in parentheses are robust standard errors, using a generalized Huber/White/Sandwich estimator, and are robust to clustering within schools.

TABLE 11B: LEVELS, GAINS, AND FIXED EFFECTS MODELS WITH 1ST AND 2ND YEAR CHARTER SCHOOLS OMITTED

	LEVELS MODEL		GAINS MODEL		FIXED EFFECTS MODEL	
	Math	**Reading**	**Math**	**Reading**	**Math**	**Reading**
Charter School	-0.104 (0.056)	-0.007 (0.005)	-0.020 (0.021)	-0.010 (0.009)	-0.104** (0.023)	-0.050** (0.009)
Gender (Female=1)	0.026** (0.002)	0.181** (0.001)	0.007** (0.001)	0.001 (0.001)		
Black	-0.419** (0.017)	-0.295** (0.002)	-0.022** (0.004)	-0.028** (0.076)		
Hispanic	-0.057** (0.018)	-0.036** (0.003)	0.030* (0.005)	0.040** (0.004)		
White	0.137** (0.075)	0.210** (0.002)	-0.031** (0.004)	-0.019** (0.003)		
High school or some college	0.381** (0.004)	0.447** (0.001)	-0.008** (0.002)	0.001 (0.002)		
2-year college	0.609** (0.005)	0.699** (0.002)	-0.010** (0.003)	-0.000 (0.002)		
4-year college	0.808** (0.007)	0.896** (0.002)	0.005* (0.003)	0.012** (0.002)		
Graduate school	0.952** (0.011)	1.021** (0.002)	0.021** (0.003)	0.020** (0.002)		
Change schools	-0.115 (0.004)	-0.113** (0.002)	0.016** (0.003)	0.015** (0.002)	0.025** (0.003)	0.024** (0.002)
Gifted	1.063** (0.005)	0.912** (0.001)	-0.013** (0.002)	-0.045** (0.001)		
At-risk Charter	-0.331** (0.082)	-0.342** (0.014)	0.003 (0.043)	0.023 (0.036)	0.048 (0.053)	0.045* (0.020)
R-squared	.3786	.3441	.0008	.0012	.1240	.1165
Total observations	4,704,894	4,691,980	2,901,458	2,889,098	2,664,289	2,651,745

** Indicates statistical significance at .01 level, two-tail test. * Indicates statistical significance at .05 level, two-tail test.

For the levels model, the dependent variable is the End-of-Grade test score, normalized. For the gains and fixed effects models, the dependent variable is the change in the normalized End-of-Grade test score. The figures in parentheses are robust standard errors, using a generalized Huber/White/Sandwich estimator, and are robust to clustering within schools.

TABLE 12: FIRST-YEAR CHARTER EFFECT BY YEAR

	Levels Model		Gains Model	
	Math	Reading	Math	Reading
1998	-0.138 (0.110)	0.000 (0.068)	-0.178* (0.076)	-0.115** (0.026)
1999	-0.272* (0.118)	-0.090 (0.070)	-0.256** (0.075)	-0.146** (0.039)
2000	-0.316** (0.082)	-0.148* (0.061)	-0.288** (0.066)	-0.211** (0.037)
2001	-0.149* (0.079)	-0.008 (0.073)	-0.046 (0.106)	-0.089 (0.081)
2002	-0.272* (0.117)	-0.132 (0.119)	-0.231* (0.117)	-0.078 (0.168)
2003	-0.394** (0.102)	-0.195* (0.093)	-0.339** (0.054)	-0.156** (0.043)

** Indicates statistical significance at .01 level, two-tail test. * Indicates statistical significance at .05 level, two-tail test.

For the levels model, the dependent variable is the End-of-Grade test score, normalized. For the gains model, the dependent variable is the change in the normalized End-of-Grade test score. All models include grade-by-year fixed effects and the other independent variables shown in the "Plus Gifted and At Risk" columns of Tables 3 and 6. The figures in parentheses are robust standard errors, using a generalized Huber/White/Sandwich estimator, and are robust to clustering within schools.

	Number of First Year Schools
	[Number of Students at Those Schools]
1998	31 [2651]
1999	22 [1669]
2000	18 [1318]
2001	16 [1118]
2002	6 [614]
2003	5 [401]

TABLE 13: CHARTER EFFECT DISAGGREGATED BY GRADE

	LEVELS MODEL		GAINS MODEL	
	Math	**Reading**	**Math**	**Reading**
3rd Grade	-0.419** (0.035)	-0.248** (0.028)	-0.205** (0.049)	0.011 (0.043)
4th Grade	-0.207** (0.063)	-0.046 (0.042)	-0.047 (0.034)	-0.006 (0.018)
5th Grade	-0.130 (0.072)	-0.032 (0.041)	-0.080** (0.026)	-0.080** (0.022)
6th Grade	-0.017 (0.053)	0.054 (0.035)	-0.096* (0.041)	-0.078** (0.018)
7th Grade	0.044 (0.063)	0.106** (0.037)	-0.013 (0.034)	-0.017 (0.015)
8th Grade	0.025 (0.080)	0.082* (0.033)	-0.009 (0.037)	-0.006 (0.018)
3rd through 5th Grades	-0.255** (0.048)	-0.111** (0.030)	-0.061** (0.021)	-0.038** (0.011)
6th through 8th Grades	0.016 (0.060)	0.081* (0.034)	-0.026 (0.029)	-0.019 (0.014)

** Indicates statistical significance at .01 level, two-tail test. * Indicates statistical significance at .05 level, two-tail test.

For the levels model,s the dependent variable is the End-of-Grade test score, normalized. For the gains models, the dependent variable is the change in the normalized End-of-Grade test score. The levels and gains models also contain all the independent variables shown in the "Plus Gifted and At Risk" columns of Tables 3 and 6—but without grade-by-year effects—and the figures in parentheses are robust standard errors, using a generalized Huber/White/Sandwich estimator, and are robust to clustering within schools. Fixed effects models are not shown because they cannot be estimated on each grade individually.

TABLE 14: EFFECT OF CHARTER SCHOOL PERFORMANCE ON STUDENT EXIT

	MATH PERFORMANCE	READING PERFORMANCE	BOTH MATH AND READING PERFORMANCE
School math average	0.151** (0.025)		0.339** (0.069)
School reading average		0.112** (0.022)	-0.232** (0.067)
Student math score	-0.036** (0.003)		-0.040** (0.003)
Student reading score		-0.023** (0.002)	0.003 (0.002)
Student gender (female = 1)	-0.005* (0.002)	-0.002 (0.002)	-0.004 (0.002)
Student black	0.032** (0.009)	0.037** (0.009)	0.030** (0.008)
Student Hispanic	0.007 (0.014)	0.007 (0.015)	0.014 (0.0140
Student white	0.021* (0.008)	0.022** (0.008)	0.023** (0.008)
Student high school or some college	-0.013 (0.007)	-0.017* (0.007)	-0.009 (0.007)
Student 2-year college	-0.025** (0.008)	-0.031** (0.008)	-0.018* (0.008)
Student 4-year college	-0.044** (0.008)	-0.050** (0.008)	-0.036** (0.008)
Student graduate school	-0.035** (0.012)	-0.043** (0.010)	-0.023* (0.012)
Student gifted	0.080** (0.013)	0.073** (0.013)	0.076** (0.013)
Charter age	0.051** (0.010)	0.052** (0.010)	0.050** (0.010)
At-risk charter	-0.111** (0.024)	-0.134** (0.024)	-0.124** (0.025)
School gender average (female = 1)	-0.552** (0.147)	-0.563** (0.150)	-0.514** (0.147)
School black average	0.126** (0.047)	0.102* (0.051)	0.176** (0.049)
School Hispanic avg.	-0.583** (0.185)	-0.622** (0.196)	-0.505** (0.177)
School white average	-0.077 (0.041)	-0.084 (0.045)	-0.020 (0.049)
School high school or some college avg.	0.227** (0.087)	0.199* (0.092)	0.282** (0.092)
School 2-year college average	0.350** (0.115)	0.288* (0.121)	0.421** (0.115)
School 4-year college average	0.137* (0.069)	0.127 (0.073)	0.208** (0.072)
School graduate school average	0.101 (0.145)	0.068 (0.140)	0.223 (0.150)
School gifted average	-0.434** (0.124)	-0.380** (0.118)	-0.452** (0.129)
R-squared	.2154	.1983	.2249
Observations	54,176	54,013	53,925

** Indicates statistical significance at .01 level, two-tail test. * Indicates statistical significance at .05 level, two-tail test.

The dependent variable is a binary variable indicating whether a student attending a charter school in year t – 1 attended a charter school in year t (one = "no": exit). All the independent variables are measured at time t – 1. The figures in parentheses are robust standard errors, using a generalized Huber/White/Sandwich estimator, and are robust to clustering within schools.

References

Agostini, Michael. "Veteran Charter Schools Outperform Non-Charters on API." http://www.cacharterschools.org/pdf_files/veteran_API.pdf. Charter Schools Development Center, 2003.

Allen, Jeanne, and Autumn Cooper, editors. National Charter School Directory, 2004. Washington, DC: The Center for Education Reform, 2004.

Anderson, Lee, et al. A Decade of Public Charter Schools. Washington, DC: SRI International, November 2002.

Ballou, Dale. "Sizing Up Test Scores." Education Next, Summer 2002.

Ballou, Dale, William Sanders, and Paul Wright. "Controlling for Student Background in Value-Added Assessment of Teachers." Journal of Educational and Behavioral Statistics, Spring 2004.

Barnett, Adrian G., Jolieke C. van der Pols, and Annette J. Dobson. "Regression to the Mean: What It Is and How to Deal with It." International Journal of Epidemiology, Advance Access, August 27, 2004.

Bast, Joseph L., and Herbert J. Walberg. "Can Parents Choose the Best Schools for Their Children?" Economics of Education Review, August 2004.

Bettinger, Eric P. "The Effect of Charter Schools on Charter Students and Public Schools." Economics of Education Review, forthcoming.

Bifulco, Robert, and Helen F. Ladd. "The Impacts of Charter Schools on Student Achievement: Evidence from North Carolina." Working paper, undated.

———. "The Impacts of Charter Schools on Student Achievement: Evidence from North Carolina." Terry Sanford Institute of Public Policy, Duke University, Working paper, SAN04-01, August 2004.

Buddin, Richard, and Ron Zimmer. "Academic Outcomes." Charter School Operations and Performance: Evidence from California, chapter 3. Rand Corporation, 2003.

———. "The Performance of Charter Schools: A Complex Picture." Presented at the 2004 Annual Meeting of the American Educational Research Association, 2004.

Bulkley, Katrina, and Jennifer Fisler. "A Decade of Charter Schools: From Theory to Practice." CPRE (Consortium for Policy Research in Education) Policy Briefs, April 2002.

Chicago Tribune. "The Facts about Charter Schools." Chicago Tribune, August 18, 2004.

Dickman, Anneliese, et al. "Charter Schools in Wisconsin: Assessing Form and Performance." Paper prepared for the 2003 Annual Meeting, American Political Science Association, August 28–31, 2003.

Dominello, Amy. "School Seeks to Expand and Relocate." Greensboro News & Record, February 24, 2004.

Duflo, Esther. "Empirical Methods." For MIT course 14.771, http://ocw.mit.edu/NR/rdonlyres/Economics/14-771Development-Economics--Microeconomic-Issues-and-Policy-ModelsFall2002/2494CA2C-D025-40A6-B167-F8A5662520DB/0/emp_handout.pdf. Unpublished paper, Fall 2002.

Fernandes, Deirdre. "East Winston Primary Celebrates Scores." Winston-Salem Journal http://www.journalnow.com, September 16, 2002.

Finn, Chester E., Jr. "Defaming Charters." New York Post, August 19, 2004.

Freedman, Samuel G. "Report Offers No Clear Victory for Charter School Opponents." New York Times, August 25, 2004.

Gill, Brian P., et al. Rhetoric Versus Reality: What We Know and What We Need to Know about Vouchers and Charter Schools. Santa Monica, CA: Rand Corp., 2001.

Greene, Jay P., Greg Forster, and Marcus A. Winters. "Apples to Apples: An Evaluation of Charter Schools Serving General Student Populations." Manhattan Institute, Center for Civic Innovation, Education Working Paper No. 1, July 2003.

Greene, William H. Econometric Analysis, Fourth Edition. Upper Saddle River, NJ: Prentice-Hall, Inc., 2000.

Gronberg, Timothy J., and Dennis W. Jansen. "An Analysis of Texas Charter School Performance." Texas Public Policy Foundation, April 2001.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. "Disruption Versus Tiebout Improvement: The Costs and Benefits of Switching Schools." Working paper, October 2001.

———. "The Impact of Charter Schools on Academic Achievement." Working paper, December 2002.

Hassel, Bryan C., and Michelle Godard Terrell. The Rugged Frontier: A Decade of Public Charter Schools in Arizona. Washington, DC: The Progressive Policy Institute, June 2004.

Haynes, Doug. "Let Charter Schools Work." Business Leader Online, September 1998.

Herszenhorn, David M. "City Plans to Eliminate Most Middle Schools." New York Times, March 3, 2004.

Hill, Paul T. "Assessing Student Performance in Charter Schools." Education Week, January 12, 2005.

Houtenville, Andrew J., and Karen Smith Conway. "Parental Effort, School Resources, and Student Achievement: Why Money May Not 'Matter'." Working paper, June 2003.

Hoxby, Caroline M. "School Choice and School Competition: Evidence from the United States." http://post.economics.harvard.edu/faculty/hoxby/papers/hoxby_2.pdf, 2003.

———. "A Straightforward Comparison of Charter Schools and Regular Public Schools in the United States." Working paper, September 2004a.

———. "Achievement in Charter Schools and Regular Public Schools in the United States: Understanding the Differences." Working paper, December 2004b.

Hoxby, Caroline M., and Jonah E. Rockoff. "The Impact of Charter Schools on Student Achievement." Working paper, May 2004.

Jones, Sherry. "Tuition-Free Charter School in Brunswick Is Expanding." Wilmington Star News, January 19, 2004.

Kane, Thomas J., and Douglas O. Staiger. "The Promise and Pitfalls of Using Imprecise School Accountability Measures." Journal of Economic Perspectives, Fall 2002.

Looney, Melanie, editor. Charter School Closures: The Opportunity for Accountability. Washington, DC: Center for Education Reform, October 2002.

Loveless, Tom. The 2002 Brown Center Report on American Education: How Well Are American Students Learning? Washington, DC: The Brookings Institution, September 2002.

———. The 2003 Brown Center Report on American Education: How Well Are American Students Learning? Washington, DC: The Brookings Institution, October 2003.

Manual, John, and Mike McLaughlin. "The Charter School Experience in North Carolina." North Carolina Insight, July 2002.

McCaffrey, Daniel F., et al. "Models for Value-Added Modeling of Teacher Effects." Journal of Educational and Behavioral Statistics, Spring 2004.

Meyer, Bruce D. "Natural and Quasi-Experiments in Economics." Journal of Business & Economic Statistics, April 1995.

Miron, Gary. "Charter Schools." State Education Standard, Autumn 2003.

Miron, Gary, et al. "Strengthening Pennsylvania's Charter School Reform." The Evaluation Center, Western Michigan University, October 2002.

Miron, Gary, and Christopher Nelson. "Student Academic Achievement in Charter Schools: What We Know and Why We Know So Little." Working paper, December 2001.

Newmark, Craig M. "A Note on Hoxby's Finding for North Carolina Charter Schools." Unpublished paper, January 2005.

Noblit, George W., and Dickson Corbett. "North Carolina Charter School Evaluation Report." Submitted to the North Carolina State Board of Education, November 2001.

Rotherham, Andrew J. "Increasing the Supply of Public Schools." Progressive Policy Institute Briefing, http://www.ppionline.org/ppi_ci.cfm?knlgAreaID=110&subsecid=134&contentid=251463, April 9, 2003.

Sass, Tim R. "Charter Schools and Student Achievement in Florida." Working paper, October 18, 2004.

Sherman, Elizabeth B. "Study Questions Charter Schools." The Digital Courier, http://www.thedigitalcourier.com/articles/2004/08/24/news/news02.prt, August 24, 2004.

Solmon, Lewis. "Does Charter School Attendance Improve Test Scores? The Arizona Results." http://www.ncsc.info/newsletter/conference/solmon.html.

Solmon, Lewis, and Pete Goldschmidt. "Comparison of Traditional Public Schools and Charter Schools on Retention, School Switching, and Achievement Growth." Working paper, March 15, 2004.

Solmon, Lewis, Kern Paark, and David Garcia. "Does Charter School Attendance Improve Test Scores? The Arizona Results." Working paper, March 2001.

Todd, Petra E., and Kenneth I. Wolpin. "On the Specification and Estimation of the Production Function for Cognitive Achievement." Economic Journal, February 2003.

———. "The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps." Working paper, April 26, 2004.

Witte, John F., et al. "The Performance of Charter Schools in Wisconsin." Wisconsin Charter Schools Study, Robert La Follette School of Public Affairs, University of Wisconsin–Madison. Working paper, August 2004.

Wolf, Patrick J. "Democracy in Education, America's Latest Voluntary Organization: The Charter School." Education Next, Spring 2002.

About the Author

Craig M. Newmark is Associate Professor of Economics at North Carolina State University. He has a B.A. from George Washington University and a Ph.D. from U.C.L.A. His research has focused on American antitrust policy. Dr. Newmark has been invited to speak to economists at the U. S. Department of Justice and the Federal Trade Commission. He has published in leading economics journals, including the Journal of Political Economy, the Journal of Law and Economics, and the Review of Economics and Statistics. Dr. Newmark has taught a wide variety of courses that includes Managerial Economics, Introduction to Economic Research, Economics of Information Goods, Industrial Organization, and Politics and Markets, and he has been consultant to a number of organizations, including the American Petroleum Institute and the North Carolina Health Products Manufacturers Association. He and his wife, Betsy, a teacher at Raleigh Charter High School, live in Raleigh and they have two daughters.

About the John Locke Foundation

The John Locke Foundation is a nonprofit, nonpartisan public policy institute based in Raleigh. Its mission is to develop and promote solutions to North Carolina's most critical challenges. The Locke Foundation seeks to transform state and local government through the principles of competition, innovation, personal freedom, and personal responsibility in order to strike a better balance between the public sector and private institutions of family, faith, community, and enterprise.

To pursue these goals, the Locke Foundation operates a number of programs and services to provide information and observations to legislators, policymakers, business executives, citizen activists, civic and community leaders, and the news media. These include the foundation's monthly newspaper, *Carolina Journal*; its daily online news service, *CarolinaJournal.com*; the *Locke Letter*, a quarterly newsletter for donors; regular events and conferences on important public policy issues; and research reports of varying lengths on topics facing state and local governments.

The Foundation is a 501(c)(3) research institute and is funded solely from voluntary contributions from individuals, corporations, and charitable foundations. It was founded in 1990. For more information, visit www.JohnLocke.org.

"To prejudge other men's notions before we have looked into them is not to show their darkness but to put out our own eyes."

JOHN LOCKE (1632–1704)

Author, ***Two Treatises of Government*** **and** ***Fundamental Constitutions of Carolina***

200 West Morgan St.
Raleigh, NC 27601
V: 919-828-3876
F: 919-821-5117
www.johnlocke.org
info@johnlocks.org