A Comparison of Charter Schools and Traditional Public Schools in Idaho

Dale Ballou

Bettie Teasley

Tim Zeidner

Vanderbilt University

August, 2006

Abstract

We investigate the effectiveness of Idaho charter schools relative to traditional public schools, using the average difference in test score gains in the two sectors as well as the student fixed effects estimator favored in the literature. Our findings are quite sensitive to the choice of estimator. When student fixed effects are included, charter schools appear more effective than traditional public schools in the elementary grades. When student fixed effects are omitted, this is no longer true. We attribute the difference to biases associated with heterogeneity in schools and in the quality of school-student matches when the fixed effects estimator is used. We find much less evidence of selection bias, the standard rationale for the fixed effects estimator.

Several recent studies have investigated the effectiveness of charter schools vis-à-vis traditional public schools. Although they have reached varying conclusions, they share some methodological concerns. The students who enroll in charter schools are self-selected. As such, they may be atypical of the larger population of traditional public school students. This raises the danger that comparisons of student achievement in charter schools and traditional public schools will reflect pre-existing differences in the students rather than school quality. This concern is partly alleviated when schools are compared on a value-added basis, that is, on progress rather than the level of achievement. However, if students differ not only with respect to their initial level of achievement but also in their rate of growth, the use of value-added measures alone will not remove all selection bias. For this reason, researchers with access to longitudinal data at the student level have exploited the repeated observations on each student to control for unobservable, time-invariant factors that cause some students to learn more rapidly than others, regardless of the school they attend. The usual approach is to introduce a student fixed effect into a model of achievement gains. The fixed effect is picked up in the between-student variation in gains, leaving the within-student variation for estimating the effectiveness of charter schools relative to traditional public schools. Studies that have followed this approach include Bifulco and Ladd (2004), Hanushek et al. (2005), Sass (2004), and Solomon, Paark and Garcia (2001), examining charter schools in North Carolina, Texas, Florida, and Arizona, respectively.

The introduction of student fixed effects means that only those students who have moved between traditional public and charter schools within the sample period contribute information about the relative effectiveness of the two types of schools. This raises the prospect of new biases. Just as charter school students may be atypical of the total student population, so the students who move back and forth between charter and traditional public schools may be a nonrepresentative subset of all those who enroll in charter schools.

We illustrate this with a study of the effectiveness of charter schools in Idaho. Although Idaho is not a populous state, with only 19 charter schools, our investigation is facilitated by the fact that Idaho tests students in grades 2 through 10 in both fall and spring semesters.[1] This gives us cleaner measures of value-added as well as an opportunity to compare gains of a single student across adjacent years without being troubled by the negative correlation induced when the same test furnishes the end-of-period score for one value-added calculation and the start-of-period score for next year's value-added.[2] As we will see, Idaho test data clearly show that charter school effects can differ considerably when estimated on the population of all charter school students as opposed to movers only.

Figure 1 depicts mean levels of mathematics achievement in traditional public schools and charter schools across the nine tested grades. While the data shown are from spring of 2005, the pattern is the same in all semesters. Achievement levels are higher in every grade in charter schools. In several instances, charter school students even outscore traditional public school students at the next grade level. This does not mean charter schools in Idaho provide a superior education. Rather, Idaho charter schools attract above-average students.

For this reason, gain scores afford a more reasonable basis for comparing charter schools with traditional public schools. Figure 2 depicts mean gains between fall and spring testing in

---

[1] An additional six charter schools started operation in 2005-06, bringing the total to 24. However, the opening of these schools falls outside our sample period.

[2] We are also able to include students in our fixed effects specification who have as few as two years of data, provided one year was in a charter school and one year in a traditional public school. This is in contrast to studies relying on data from states that test only at the end of the school year, where a minimum of three years' data are needed to obtain two measures of gain. This increases our effective sample size, particularly for students in the lower grades. Because Idaho begins testing in grade 2, we can obtain fixed effects estimates for students as soon as they reach third grade, by contrast with states where the youngest students with two years of gain scores are fifth graders. Our samples are therefore more fully representative of the population served by charter schools, the more so given the decline in entry rates in higher grades (Hoxby and Rockoff, 2005).

the 2004-05 school year. In every grade, traditional public school students gain more than charter school students. (Again, the same pattern is evident in other years.) By this simple test, traditional public schools outperform charter schools. However, this conclusion rests on the implicit assumption that the students enrolled in charter schools do not differ from the students enrolled in traditional public schools in any other way that affects gains. As this may not be true, we consider a second simple test of the comparative effectiveness of charter schools: the change in gain scores as students move between sectors, a difference-in-differences estimator.

Figure 3a shows gain scores for students enrolled in fifth grade or lower in the 2002-03 year. Within-year gains (fall to spring) are shown for 2002-03 and 2003-04. The data are longitudinal: the students whose gains are depicted for 2002-03 are shown one year later in 2003-04. Because gains tend to diminish with advancing grade level (compare Figure 2), 2002-03 gains generally exceed gains in 2003-04. The middle columns in the graph depict students who changed sectors between these two academic years. Students who went from traditional public schools in the first year to charter schools in the second year are the only group that experienced greater gains after the move. By contrast, those who moved from charter schools to public schools saw the greatest decline in gains. Figure 3b shows gains among elementary students for 2003-04 and 2004-05. With rare exceptions, these students are different individuals from those in Figure 3a. However, the same pattern generally holds. The smallest drop in gain scores occurred among students who moved from the public to the charter sector. The largest drop occurred among students who moved in the opposite direction.

Figure 3c depicts outcomes at the secondary level in 2003-03 and 2003-04. Both groups of movers experienced a decline in gain scores, but the drop was greater among those who moved from the charter to the traditional public sector. This pattern does not hold for changes between the 2003-04 and 2004-05 years (Figure 3d). Students moving from charter to traditional public saw essentially no change in gains, while those moving the other way experienced a decline.

These two tests lead to very different conclusions about charter school effectiveness vis-à-vis traditional public schools. Whereas charter schools appear less effective with respect to mean gains, the difference-in-differences estimator favors charter schools in the elementary grades, with the situation less clear in the secondary grades. Of course, none of the graphical evidence we have considered so far should be regarded as conclusive. We have not introduced controls for any covariates, nor have we made efficient use of the data in our second test, as we have examined outcomes only in the years immediately before and after a move. Nonetheless, the evidence in these figures alerts us to the possibility that even with a more sophisticated analysis, our findings may depend on which of these two fundamental approaches we follow to estimate a charter school effect: one that uses all the cross-sectional variation in the data, or one that relies on within-student variation only.

A Model of Achievement Gains

We suppose that student achievement satisfies

(1) $y_{ijt} = X_{ijt}\,\beta + \alpha_i + \alpha_j + \alpha_{ij} + u_{ij}$,

where the dependent variable, y_{ijt}, represents the gain between fall and spring testing for student i in school j during year t. X is a set of covariates such as researchers typically have for students and schools (e.g., race, eligibility for the free and reduced-price lunch program). The remaining terms represent the contribution of unobservable factors. α_i represents time invariant factors such as family income and parental education that affect the achievement of student i, regardless where he attends school. α_j represent differences in school quality affecting all students alike. α_{ij} is an interaction term representing how well school j matches the learning style of student i. u_{it} represents the influence of transitory unobservables. We examine this model not because we expect to distinguish the separate effects of α_i, α_j, and α_{ij} in practice, but in order to understand what happens when researchers estimate simpler models.

Although still more elaborate specifications are possible, including those with time-varying effects, there is sufficient heterogeneity in equation (1) for our purposes. First, note that there are multiple ways of specifying a "charter school effect." One is the mean change in gain scores were all students (or a representative microcosm) to attend a charter school rather than a traditional public school, a quantity known as the average treatment effect (ATE). Given school and match heterogeneity, it matters which students attend which schools. If we make these assignments randomly (which is not likely to be the case in practice), the intended estimand is

$$(2) \quad N^{-1} \Sigma_i \Sigma_{j \epsilon C} w_j (\alpha_i + \alpha_j + \alpha_{ij}) - N^{-1} \Sigma_i \Sigma_{j \epsilon P} w_j (\alpha_i + \alpha_j + \alpha_{ij})$$

where C denotes the charter school sector and P the traditional public school sector, w_j are weights equal to the share of total enrollment in each sector accounted for by school j, and N is the total number of students.[3] The first term in (2) represents the average achievement gain if all students were to attend charter schools, with the chance that a given student attends school j equal to w_j. The second term is the corresponding average over traditional public schools. If the question of interest is whether charter schools are as effective as traditional public schools, averaging potential outcomes over all students gives us the most comprehensive answer. Thus, an estimate of ATE might be desired if the policy under consideration were whether to permit

[3] If we want to generalize beyond the population of current students, the empirical distribution, represented by N^{-1}, needs to be replaced by the probability distribution function defined over types of students. Nothing in the analysis depends on this refinement, which we ignore henceforth.

the establishment of charter schools in the first place, knowing nothing about where they might be located, or to raise the ceiling on the number allowed in the expectation that schools like those already in existence would spread to all types of communities in proportion to the number of students in each location.

The model in (1) is underidentified: even if we could observe all students in all schools, we cannot estimate unique values of all of the α_{ij} (or the α_i, if X contains a constant term). Conventional normalizations include $\Sigma_i \alpha_i = 0$ (the mean student ability is subsumed in the intercept) and $\Sigma_i \alpha_{ij} = 0$ for all j (the average match over all students is subsumed in α_j). With these normalizations the ATE becomes the enrollment-weighted difference in school quality, $\Sigma_{j\epsilon C} w_j \alpha_j - \Sigma_{j \epsilon P} w_j \alpha_j$, as one would expect.

Given the arbitrary way (2) distributes students over schools (that is, at random), the ATE may hold less interest than an estimate of the improvement in outcomes for students who have actually attended charter schools, a quantity known as the effect of treatment on the treated (ETT). Let δ_{ijt} be an indicator variable taking the value 1 if student i attends school j in year t. Let $\delta_{ijt}(a)$ be the school student i would attend under some counterfactual enrollment pattern, a. Among the possible counterfactuals, let a* represent the distribution of students over traditional public schools if there were no charter schools. Then the corresponding ETT would be

$$(3) \qquad \frac{\Sigma_t \Sigma_{i\epsilon C} \Sigma_{j\epsilon C} \delta_{ijt} (\alpha_i + \alpha_j + \alpha_{ij})}{\Sigma_t \Sigma_{i\epsilon C} \Sigma_{j\epsilon C} \delta_{ijt}} - \frac{\Sigma_t \Sigma_{i\epsilon C} \Sigma_{j\epsilon P} \delta_{ijt}(a^*) (\alpha_i + \alpha_j + \alpha_{ij})}{\Sigma_t \Sigma_{i\epsilon C} \Sigma_{j\epsilon P} \delta_{ijt}(a^*)}$$

where $\Sigma_{i\epsilon C}$ indicates that summation is over students enrolled in charter schools and we explicitly recognize the time period in δ_{ijt}.

Both (2) and (3) depend on counterfactuals. Alternatively, one might simply want to know whether charter school students are learning as much, on average, as students in traditional public schools, controlling for observable characteristics (X): that is, whether

$$(4) \qquad \frac{\Sigma_t \Sigma_i \Sigma_{j\epsilon C}\, \delta_{ijt}\,(\alpha_i + \alpha_j + \alpha_{ij})}{\Sigma_t \Sigma_i \Sigma_{j\epsilon C}\, \delta_{ijt}} \quad - \quad \frac{\Sigma_t \Sigma_i \Sigma_{j\epsilon P}\, \delta_{ijt}\,(\alpha_i + \alpha_j + \alpha_{ij})}{\Sigma_t \Sigma_i \Sigma_{j\epsilon P}\, \delta_{ijt}}$$

is greater than or equal to zero. We will refer to (4) as the average difference (AD). AD differs from ATE and ETT in that the students in the two terms are not the same. While the α_i drop out from ATE and ETT, this is not true of (4), where the comparison of charter to traditional public schools reflects pre-existing unobservable differences in the students enrolled in the two sectors. Thus, of the three expressions, AD is the least satisfactory representation of a charter school "effect," although it can be improved by restricting the set P to schools serving a population similar to that served by the charter sector, as in Hoxby (2004).

In practice, of course, it is AD or a similar expression that is estimated, given the unobservable counterfactuals in ATE and ETT. The question, then, is how close AD is to one of the average treatment effects in which we are truly interested. If the students enrolled in charter schools are a random subset of the student population (after controlling for X), and if their distribution over the set of charter schools is likewise random, then an unbiased estimate of AD—such as the difference between mean gains in the two sectors—is also an unbiased estimate of ATE. If we do not want to make such strong assumptions, we can remove the α_i from (1) by first-differencing or introducing student fixed effects. In this case, it is only the students who move between sectors who contribute to our estimate of the charter school effect. If the public schools from which they come or for which they leave are their a* schools, the result comes nearer being an estimate of ETT than of ATE. However, it is an estimate based only on those who actually move. If movers are not a random subset of those who enroll in charter schools, we will be estimating not ETT but

$$(5)\qquad \frac{\Sigma_t \Sigma_{i\epsilon M} \Sigma_{j\epsilon C}\, \delta_{ijt}\, (\alpha_j + \alpha_{ij})}{\Sigma_t \Sigma_{i\epsilon M} \Sigma_{j\epsilon C}\, \delta_{ijt}} \; - \; \frac{\Sigma_t \Sigma_{i\,\epsilon\, M} \Sigma_{j\,\epsilon P}\, \delta_{ijt}\, (\alpha_j + \alpha_{ij})}{\Sigma_t \Sigma_{i\,\epsilon M} \Sigma_{j\epsilon\, P}\, \delta_{ijt}}$$

where M denotes the subset of students who move between sectors within the sample period. We will refer to (5) as the fixed effects estimand, FE.

It is instructive to derive explicit expressions for the bias of the fixed effects and average difference estimators, FE and AD, on the assumption that we want to estimate the ETT in (3). There are three groups of students: movers between charter schools and traditional public schools (denoted with a subscript "M"), stayers who attend charter schools the entire period ("Sc") and stayers who attend traditional public schools ("Sp"). Let P_i denote the mean gains group i actually experiences while attending public schools, and P_i* the mean counterfactual public school gains for students in the charter sector. C_i is the mean gain of group i while enrolled in charter schools. All of these can be expressed in terms of the model developed above. For example, P_{Sc}* is given by

$$\frac{\Sigma_t \Sigma_{i\epsilon Sc} \Sigma_{j\,\epsilon P}\, \delta_{ijt}(a^*)\, (\alpha_i + \alpha_j + \alpha_{ij})\ .}{\Sigma_t \Sigma_{i\epsilon Sc} \Sigma_{j\epsilon P}\, \delta_{ijt}(a^*)}$$

Finally, let w_{Sc} denote the proportion of charter school observations accounted for by stayers and w_M the proportion by movers, with $w_{Sc} + w_M = 1$. Then we have

(6a) $ETT = [C_M\text{-}P_M^*]w_M + [C_{Sc}\text{-}P_{Sc}^*]w_{Sc}$

(6b) $FE = [C_M - P_M]$

(6c) $AD = [C_M - P_M]\, w_M + [C_{Sc} - P_{Sp}]\, w_{Sc}$.

The two biases are therefore

(7a) $\text{bias(FE)} = [P_M^* - P_M]w_M + [P_{Sc}^* - P_M]w_{Sc} + [C_M - C_{Sc}]w_{Sc}$

and

(7b) $\text{bias(AD)} = [P_M^* - P_M]w_M + [P_{Sc}^* - P_{Sp}]w_{Sc}$.

The first term, common to both biases, represents the difference between movers' counterfactual gains (had they remained in the traditional public sector) and their actual gains when enrolled there. Within the framework developed above, these terms differ if the traditional public schools they would attend under the counterfactual are not the schools they did at one time attend. In a more elaborate model that allowed for varying time subscripts on $\underline{\alpha}_j$ and α_{ij}, these terms could differ even if the identity of the schools remained the same.

The second term in (7b) represents the difference between the counterfactual achievement of charter school stayers and the actual achievement of public school stayers who make up their comparison group in the AD estimator. This term captures selection bias (differences in the student effects) as well as differences between the schools attended by the comparison group and charter school students' a* schools (although this can be mitigated by restricting the comparison group to students enrolled in the schools we think the charter school students would have attended) and differences in school-student matches.

The second and third terms in (7a) reflect possible differences between movers and charter school stayers. They include the following: (i) Movers are atypical of charter school students as a whole with respect to match-heterogeneity, α_{ij}. (ii) Movers may select on school quality, α_j, with the result that some values of α_j are over-represented among movers. (iii) The student-school match may be a function of a student's prior educational history. For example, Hoxby and Rockoff (2005) argue that charter schools are more effective with students who have attended them from the earliest grades than they are with students who come to them after attending traditional public schools. More generally, lagged effects of prior schooling can cause the value of α_{ij} a student encounters on moving to a new school to differ from the value of α_{ij} that would obtain had the student attended that school from the first.

Most of these problems have been recognized in the literature, though efforts to quantify these biases have been limited to a few specific sources (for example, whether the schools left by movers are particularly ineffective). There has been no discussion of the relative magnitude of the biases in (7a) and (7b); rather, there has been a presumption that (7a) is smaller. This seems unwarranted. Except for the first term common to both, the biases arise from different sources. The AD estimator is biased when charter school stayers differ from public school stayers. The FE estimator is biased when charter school movers differ from charter school stayers. Using FE rather than AD removes one source of bias but only at the price of introducing another.

In deriving (7a) and (7b), we have ignored problems occasioned by the correlation of u_{it} with other variables in empirical models. As these have attracted some attention in the literature, we note them here. First, students (and their parents) may make mistakes about α_{ij}. If they erroneously interpret u_{it} as information about school quality or the student-school match and move between sectors accordingly, the return of the error to a more normal level the following year induces a positive correlation between u_{it+1} and a charter school dummy variable (if they have moved into the charter sector), a negative correlation (if they have moved out of the charter school sector). Second, mobility per se can have a deleterious effect on student performance, which will be confounded with school quality if mobility is not appropriately controlled for in the vector of student characteristics, X. Third, serial correlation in u_{it} (trends in achievement) can be confounded with differences in school quality when movers switch sectors. All of these phenomena represent additional reasons to mistrust estimators based on samples of movers.

Data

We explore these issues further by using the AD and FE estimators to compare charter schools to traditional public schools in Idaho. Data for this study have been furnished by Northwest Evaluation Association (NWEA). The state of Idaho has contracted with NWEA to provide tests for its statewide assessments in grades 2 through 10. Participation rates are over 90 percent. Unique student identification numbers allow us to link records longitudinally. Although tests are administered in reading, language arts, and mathematics, for this study we use only the mathematics results.

NWEA uses the one-parameter IRT model to place all students on a single developmental scale. Though the meaningfulness of such scales for students at very different levels of achievement has come into question, there is no alternative of greater validity and we use the fall to spring change in the scale score as our dependent variable.[4] Our empirical model takes one of two forms, corresponding to the FE and AD estimators:

(8) $y_{ijt} = \mathbf{X}_{ijt} \boldsymbol{\beta} + \Sigma_g C_{it} \gamma_g \psi_g + \alpha_i \varphi_i + e_{it}$,

or

(9) $y_{ijt} = \mathbf{X}_{ijt} \boldsymbol{\beta} + \Sigma_g C_{it} \gamma_g \psi_g + v_{it}$.

[4] The basic problem is that a 5-point gain at one point on the scale may not represent the same increment of "true achievement" as a 5-point gain elsewhere on the scale. This problem is not solved by standardizing scores to have zero mean and standard deviation of one within each grade, by expressing results as percentiles or normal curve equivalents, or by any of the other devices commonly used to make test results "comparable." Indeed, it is not clear that "an increment of true achievement" is a meaningful notion.

(8) includes student dummy variables φ_i while (9) does not. Thus in (9) this form of heterogeneity is incorporated in the error term along with the other types of heterogeneity that are not recognized explicitly in either equation. C_{it} is an indicator variable for charter schools. Because the effect of attending a charter school is not constant across all grade levels, C_{it} is interacted with an indicator of grade, γ_g. The coefficients on these interactions, ψ_g, represent our estimates of (5) or (4) as the model includes or omits student fixed effects. Neither, of course, may correspond closely to the ETT in (3).

We have a limited number of covariates for inclusion in the model. $\mathbf{X_{ijt}}$ includes indicators of race (white = 1) and special education. Preliminary analysis has shown that relationships between these variables and achievement vary with grade level. Accordingly, both covariates are interacted with grade.[5] We also included dummy variables for year by grade interactions, to control for possible changes in the difficulty of the tests. Although students in both the traditional public and charter sectors take the same examinations, the size of the charter sector has been growing, thereby placing greater weight on results of more recent exams.

Testing dates vary in Idaho. The average time elapsed between fall and spring testing is about 135 school days, with a standard deviation of slightly more than a week. However, the range is quite wide. In 2004-05, for example, the maximum time between fall and spring testing was 175 school days, while the minimum was only 70. To control for the impact on learning,

---

[5] Because all covariates are interacted with grade, there are no time invariant regressors in the model. Thus inclusion of student fixed effects does not cause any other variables to drop out. However, it is still the case that only movers directly furnish information about the effectiveness of charter schools relative to traditional public schools. Observations on non-movers furnish information about relative effectiveness of instruction at different grade levels within sector (charter or traditional public), but only through this channel do they have any influence on estimates of the difference in average effectiveness between sectors. Accordingly, we will continue to refer to the FE estimator as an estimator based on a sample of movers, which, though not strictly true here, correctly characterizes the rest of the literature.

we normalize gain scores by dividing by the approximate number of school days between fall and spring testing, multiplied by 180 to represent a "standard year's" gain.[6]

The model was estimated using fall-to-spring gain scores from 2002-03, 2003-04, and 2004-05 for students in grades 2 through 10, when testing is mandatory. The sample was restricted to students enrolled in charter schools or traditional public schools: if all or a portion of a year was spent in a private school or a state other than Idaho, data for that year were dropped. The sample was further restricted to students whose fall and spring schools were the same, thereby excluding students who switched schools in mid-year.[7] We did this to avoid having to apportion the year's gains between the two schools when we do not know when the move occurred, as well as out of concern that mid-year transitions, because they are disruptive, have an independent effect on gains. We also dropped observations from Idaho's five virtual (on-line) charter schools. Although of interest in their own right, these schools are so distinctive (and their combined enrollment so large) that their inclusion in the estimation sample would skew the comparison.[8]

Of the remaining observations, testing dates are missing for about 10.5 percent of the charter school observations and 4.5 percent of the public school observations. All but a handful of these missing values occur in the 2002-03 data. Students with missing testing dates tend to have higher scores than other students in the same grade and sector. There are also some statistically significant differences in gain scores, though of mixed signs. Because it was not possible to adjust for the time elapsed between fall and spring testing, we do not know whether

[6] The normalization formula is [5/7(Elapsed Calendar Days)-15](180), where 15 is an estimate of the number of vacation days (Thanksgiving, Christmas, winter or spring vacation).
[7] Of the 8251 observations on students who spent all or part of any year in a charter school, 921 or 11.2% were dropped for this reason, as was a much smaller share of the observations on public school students (3.7%).
[8] Students enrolled in virtual schools account for 35% of the charter school observations.

these differences would remain were the scores normalized.[9] These observations are also dropped from the final sample.

Characteristics of the estimation sample are displayed in Table 1. Students are classified into three groups: those who remained in the traditional public sector throughout the sample period; those who attended only charter schools; and those who moved between sectors. Interestingly, there are more of the latter than of charter school stayers. Observations on movers peak during grades 5-8, suggesting that students are more likely to move in or out of charter schools at the same time they transition between elementary school and middle or junior high school.

Compared to students in the traditional public sector, students who spent at least part of this period in a charter school are more likely to be non-hispanic whites. There are also fewer special education students attending charter schools, though the percentage among movers is virtually the same as among traditional public school students. Both of these differences are, of course, consistent with the evidence in Figure 1 that charter schools tend to attract higher-achieving students.[10]

Results

Estimates of the charter school effect are displayed in Table 2. (A full set of results with coefficients on all the variables in the model is available from the authors.) Models were estimated by ordinary least squares. Standard errors have been adjusted for clustering at the

---

[9] In almost all cases, if the testing date is missing for one student in a school, it is missing for all of them, leaving no basis for imputing the missing dates.

[10] Because several of Idaho's charter schools do not participate in the free and reduced-price lunch program, data on student eligibility are spotty. Accordingly we have not included this variable as a control.

school level using a robust asymptotic covariance matrix.[11] As anticipated, there is a pronounced difference between the models with and without fixed effects. The charter school effect is negative at every grade in the average difference estimator (column 1a). By contrast, when the model includes student fixed effects, we find gains are greater in charter schools in the elementary grades, with the estimates for grades three, four, and five significant at the 5% level or better. Between columns 1a and 1b there is a discrepancy of about seven points for third graders, four and a half points for fourth graders. These are sizeable differences, given the mean annual gain in these grades is 12 points.

We have explored a variety of alternative specifications. The fixed effects estimator depends on students moving between sectors to identify charter school effects. This raises the possibility that the charter school effects are confounded with the impact of movement per se. We therefore added a binary indicator for all students who changed schools, whether within or between sectors. This made virtually no difference to the estimates of charter school effectiveness (columns 2a and 2b), and the effect of mobility itself was small and statistically insignificant in most grades. We also tried a variant in which the mobility indicator picked out students who transitioned between two districts, on the grounds that moving between a traditional public school and a charter school might more nearly resemble a change of district than a move to a new school in the same district. In this specification (column 3b), the FE charter school effect for fifth grade is no longer statistically significant, though it remains positive. The third and fourth grade charter effects remain large and significant at the 1% level. In all of these specifications, the charter school effects obtained with the average difference estimator remain negative in every grade.

[11] In order to compute the robust standard errors, it was sometimes necessary to drop observations where the number of students in a particular grade and school was quite small (e.g., fewer than four). This accounts for some of the discrepancies in sample sizes in Table 2. The impact on the coefficient estimates was trivial.

As noted in the introduction, charter school students have higher test scores than students in traditional public schools, often by as much as a full grade level. Because gains diminish on average with higher level scores (recall Figure 2), the negative charter school effect on gains reported in column 1a may simply be a function of the difference in level scores. This turns out not the case: when scores from fall, 2002, are included as a baseline regressor (column 4), the charter school effects remain negative in every grade except 10, where the positive coefficient is statistically insignificant.[12]

Studies of charter schools in Texas, North Carolina, and Florida have found that new charter schools are less effective than schools that have been in operation longer, though the relationship between effectiveness and years of operation is not always monotonic (Hanushek et al., 2005 ; Bifulco and Ladd, 2004 ; Sass, 2004). Although this finding is consistent with the anecdotal evidence on the difficulties faced by new charter schools, it does not appear to hold in Idaho. Interactions of charter school age with grade level are negative and usually statistically significant. This is true in models with and without student fixed effects. Controlling for charter school age diminishes the FE estimates of the charter school effect when the latter is evaluated at the mean charter school age for each grade level. However, there continue to be differences of 2 points or more between the AD and FE estimates for grades two through five (columns 5a and 5b).

Because our data do not go back to the earliest years of operation of the oldest charter schools, age effects may be confounded with the disadvantages of leading the way. Benefiting from the experience of the schools that preceded them, newer charter schools in Idaho may be founded on superior educational models. Or, given the relatively small number of charter

[12] The dependent variable in this model is the normalized gain score from 2003-04 or 2004-05.

schools in this state, it may simply be an accident that some of the better schools are among the more recent to open. However, when we include a fixed effect for each charter school in the model, we continue to obtain negative and statistically significant coefficients on charter school age (not shown). This suggests that either the performance of charter schools deteriorates the longer they are in operation, or the match between charter schools and students is better when a school first opens than it is in later years. Thus, schools founded to serve students with particular learning styles may attract other kinds of students over time. Unfortunately, we are unable to test this hypothesis with the short panel available to us.

Finally, our use of fall to spring gains as a measure of school effectiveness may be unfair to schools with distinctive programs that reduce summer learning loss. Such programs could include extended school years, summer tutoring, summer reading lists that students are induced to take seriously, or simply superior instruction. To explore these possibilities, we examine fall to fall gains as a function of the type of school attended the previous year and the other student characteristics in the model.[13] We also include binary indicators for students who have switched sectors over the summer as the fall to fall gain may be influenced by instructional practices during the weeks prior to testing (e.g., more intensive review).

Results are shown in columns (6a) and (6b) of Table 2. AD estimates are no longer negative in the elementary grades, though the positive coefficients are not statistically significant. The coefficients in other grades, though negative, are smaller in magnitude than was the case when the dependent variable represented fall to spring gains. However, we still

[13] The dependent variable was calculated as [(Change in Scale Score)/(Elapsed Calendar Days)](180).

obtain substantially more positive estimates of the charter school effect from the FE model in the early elementary grades and now, surprisingly, grade 10.[14]

In summary, the choice of estimator makes a considerable difference to the inference one is apt to draw about the effectiveness of Idaho charter schools. According to the fixed effects estimator, children in the early elementary grades benefit from attending charter schools. In most specifications, the average difference estimator leads to the opposite conclusion. Even the AD estimates most favorable to charter schools suggest only that charter schools are no worse in these grades than traditional public schools.

We argued above against the presumption that the fixed effects estimator is superior to the average difference estimator, as both can be biased. In which is the bias worse?

a. Bias in the average difference estimator.

The fixed effects estimator is motivated by the concern that charter school students differ from traditional public school students in unobservable ways and that these differences should not be confounded with the quality of their schooling. There is no question that this is an important consideration when comparing achievement levels. But the case is much weaker when comparing gains. The inter-year correlation in a student's level scores is approximately .9. By contrast, the correlation between gains in any two years never rises above .13. Moreover, not all of this is attributable to student-level unobservables. Some of it reflects the contribution of schools. If we restrict the comparison to students who changed districts between

[14] These estimates rely on only two years of data: gains between fall of 2002 and fall of 2003, and between fall of 2003 and fall of 2004. To verify that the differences between columns (6a) and (6b) and the rest of Table 2 are not due to this change in the sample, we have estimated the baseline model using the column 6 sample. The results (not shown) are very similar to those in columns 1a and 1b.

one year and the next, the correlation of level scores continues to be .88 or higher, but the correlation between gains falls as low as .06, depending on the years chosen.

One way to shed light on this question is to calculate the degree of self-selection needed to generate the results in column 1a of Table 2, under the null hypothesis that charter schools are just as effective as traditional public schools. Under this null, the negative coefficients in column 1a imply that charter schools attract students whose underlying rate of gain is below average. For concreteness, suppose that charter schools attract students with above average values of α_i at one uniform rate, below average values at another uniform rate. Assume also that the distribution of underlying growth rates in the student population is normal with a standard deviation of 8.55, equal to the standard deviation of the estimated student fixed effect in the model reported in column 1b.[15] Then the apparent charter school effect will equal -1.63 (the median effect for the elementary grades in column 1a of Table 2) when charter schools attract 64 percent more students with below average rates of growth than students who are above average.[16] This is a pronounced bias. Moreover, it was calculated under the null hypothesis that charter schools and traditional public schools are in reality equally effective. If instead we take the effects reported in column 1b of Table 2 as the truth, it requires an extraordinary propensity on the part of slow gainers to select charter schools to generate the estimates in column 1. For example, to turn a true charter school effect of 3.05 (the median elementary effect in column 1b) into an apparent effect of -1.63, charter schools would need to attract more

[15] This overstates the variation in true student ability, as the estimated fixed effects are affected by sampling error. A better estimate of the variability of true ability would be the standard deviation of the estimated fixed effects less their average standard error. However, the relevant consideration for our purposes is less whether charter schools recruit students whose true rate of gain is below average than whether the students recruited in this sample had below average estimated effects, as the latter cause negative charter school coefficients when student fixed effects are omitted from the model. Of course, had we used an estimate of the true variance in our calculations, an even more pronounced selection bias would be required to reconcile the AD and FE estimates than we have reported here.

[16] Found by solving ${}_{-\infty}\int^{0}\ \varphi(x)\,dx\,[1+k]\ +\ {}_{0}\int^{\infty}\ \varphi(x)\,dx\,[1-k]\ = -1.48$ for k.

than five students with below average rates of gain for every one above average. It is very hard for discrepancies of this magnitude to arise either by chance or by policy, given the difficulty of predicting gains of individual students.

We have also estimated a variant of our initial model in which student fixed effects are replaced by level scores and gains from 2002-03. The model is fit using fall to spring gains from 2003-04 and 2004-05 as the dependent variable. The results look very similar to those in column 1a of Table 2. The point estimate of the charter school effect is negative at every grade level (and significant at the ten percent level or better in grades four, five, seven and eight). Evidently the AD and the FE estimates differ not because the latter controls for student-specific rates of gain while the former does not, for when we use baseline achievement and gains from 2002-03 as an alternative to student fixed effects, we obtain results very similar to the original AD estimates. This suggests we need to look elsewhere for the explanation, notably to the fact that the FE estimates are based on the experiences of students who moved within the sample period rather than the full sample.

To sum up, the discrepancies between the AD and FE results at the elementary level are so great that self-selection on α_i is not a plausible explanation. Selection would need to be highly skewed in favor of slow gainers to produce the AD estimates in column 1a, if in fact the FE estimates in column 1b represent the truth. Moreover, we find no corroborating evidence for so skewed a selection in a variant specification of the model that controls for prior level scores and gains.

b. Selection bias among movers

Fixed effects estimates of ETT are biased if movers differ in important respects from the set of charter school students as a whole. Differences can arise due to school heterogeneity, if movers select charter schools either better or worse than the sector average, and match heterogeneity, if movers tend to have better or worse matches than non-movers. Unfortunately, direct tests of these hypotheses are not feasible: among non-movers, we cannot distinguish student heterogeneity from match heterogeneity. As a result, our investigation of selection bias among movers is at best suggestive.

We begin by asking whether movers tend to select non-representative schools. We have already noted that in Idaho, the effectiveness of charter schools appears to decline the longer they are in operation (or at least the quality of the match is not as good). Because students who move between sectors (particularly those who move into the charter sector) tend to choose schools that are newer, charter schools will appear to be more effective among a sample of movers than among a representative sample of all charter school students. We quantify this effect in Table 3. As shown in columns 1 and 2, elementary school movers are especially likely to select newer charter schools. In columns 3 and 4 we multiply mean school age by the coefficient on charter school age from the AD estimates, and in columns 6 and 7 by the coefficient from the FE estimates. Regardless of which estimates we use, the pattern is the same: the difference between movers and all charter school students is greatest (and most positive) in the early elementary grades (except grade 2). This, of course, is what we observed in columns 5a and 5b of Table 2. When we controlled for charter school age, the FE charter school coefficients declined compared to their values in the baseline model. No such decline was evident in the AD coefficients.[17] An appreciable part of the FE estimates of charter school effectiveness appears to be a consequence of movers' tendency to select newer charter schools.

[17] In fact, they are necessarily the same. The effect of charter school age, evaluated at the within-grade mean, is already subsumed in the grade-level charter school effects in column 1a.

As a more general measure of school quality, we use mean fall to spring gains of non-movers after controlling for race and special education. By using non-movers' gains, we abstract from the effect that movers themselves have on measured school quality. Nonetheless, this is only a rough measure of school quality, ignoring both individual and match heterogeneity. We then ask two questions: (1) is there positive (or negative) selection on quality of the destination school among students moving across sectors; (2) is there positive (or negative) selection on quality of movers' school of origin? To answer these questions, we regress our measure of school quality on the number of students moving into the school (or, in the case of question two, the number of students leaving the school). In both cases we count only those students changing sectors as movers. In these regressions we control for the total enrollment at the school so that the number of movers is not confounded with school size. To improve statistical power, we control for two levels, elementary and secondary, rather than each grade level. For question one, which is forward-looking, we use mean gains from 2003-04 or 2004-05 as the dependent variable. As indicated above, these are gains among students who are not new to the school. For question two, which is backward-looking, the dependent variable is the mean gain from 2002-03 or 2003-04 among students who remained in the school the following year.[18] Regression equations are weighted by school size, but results for unweighted regressions are very similar.

Results are displayed in Table 4. Rows 1 and 2 show results for school of destination and rows 3 and 4 for school of origin. Students moving to a charter school tend to select better than average schools, as measured by next year's gains among students already enrolled at the

[18] We have also re-run the models using the mean gain among all students as a measure of the quality of a school of origin, so that the decision of movers to depart does not affect measured quality. Very similar results are obtained. Among charter schools of origin, the coefficients on number of movers are .24 at the elementary level and -.08 at the secondary level. Among traditional public schools of origin, the coefficients are -.05 and -.001.

school. The coefficient on the number of movers is approximately .1 at both the elementary and secondary level. The annual number of movers into charter schools has a standard deviation of 30 (elementary) or 25 (secondary). Thus, an increase of one standard deviation in the number of movers into a charter school is associated with an increase of 2.5 to 3 points in gain scores, controlling for student characteristics. By contrast, students moving from charter schools to traditional public schools are more likely to select destination schools where next year gains among students already enrolled in the school will be below average. The coefficient at the elementary level is -.13 and at the secondary level -.06. However, a one standard deviation increase in the number of movers into these schools is only about three students. As a result, bias from this source is slight.

In row three we see that elementary students quitting the charter sector for traditional public schools leave schools where non-movers' previous year gains were better than average for the sector, though the coefficient is small and far from statistically significant. Given that the standard deviation in the number of students exiting a charter school is about 4, this is not an important factor, but its direction is consistent with the hypothesis that the fixed effects estimator, by relying on data of movers, exhibits a positive bias at the elementary level. Among traditional elementary public schools, an increase in the number of students leaving for a charter school is associated with below-average school performance (column 4). However, the coefficient again is small, so that a one standard deviation in the number of exiting students (approximately 3) causes only a minor bias in the estimate of charter school effectiveness.

To summarize, movers appear to select on school quality, both with respect to schools of origin and schools of destination. Of the two, selection on schools of destination appears more important. We quantify the impact of selection on estimates of charter school effectiveness in

Table 5, where we have introduced residual gains by non-movers as an additional regressor (interacted with indicators of elementary or secondary level). This necessarily limits the sample to observations from 2003-04 and 2004-05, as there are no identifiable destination schools in 2002-03. Column 1 of Table 5 displays estimates of the baseline model for this reduced sample. The charter school coefficients are quite similar to those of the FE estimates in column 1b in Table 2 obtained from the full sample for all three years. Column 2 displays results when our measure of school quality is included. There is an appreciable drop in the magnitude of the charter school effects at the elementary level, suggesting that much of the estimated charter school "effect" reflects the selection of a non-representative sample of schools by movers.

In addition to choosing schools that are atypical, movers are a self-selected group that may differ from other students in the same schools with respect to the quality of the student-school match. To investigate this issue, we compare school leavers to charter school stayers, asking whether their gains in the year prior to the move differed from gains of students who remained in the charter sector. Our regression equation includes controls for race and special education by grade. The equation also includes dummy variables for each charter school, so that the comparison is based on within-school differences between leavers and stayers. Results are displayed in Table 6, column 1. Coefficients on mover are positive in grades 2 through 5, negative thereafter. The standard errors are large, suggesting that with repeated sampling different patterns of selection might well emerge. However, for purposes of understanding the discrepancies between AD and FE estimates in *this* sample, it is the point estimates that are important. Clearly leavers were not a representative sample of all charter school students.

The differences between leavers and stayers could be a reflection of student heterogeneity rather than match heterogeneity. However, post-move differences between leavers and stayers, displayed in column 2, do not show the same pattern. There is remarkably little consistency between the two columns. Only the students who left charter schools after grade 5 appear to have done equally well the following year, relative to charter school stayers. The FE estimator takes this decline as evidence that traditional public schools are not as effective as charter schools; however, the pre-move differences between leavers and stayers suggest that leavers at the elementary level may have quit schools where they had good matches for schools where the match was not as favorable.[19]

To summarize, we find evidence of selection on school quality. Movers are more likely to select new charter schools, which in Idaho appear to be more effective. Alternatively, when we use the gains of non-movers as an indicator of school quality, we find that movers' schools of origin and of destination are atypical, with the charter schools at the elementary level tending to be better than average and the traditional public schools at that level worse. Not all of these effects are large, but they all point in the same direction: a positive bias at the elementary level in estimated charter school effectiveness when the estimation sample is restricted to movers. There are also indications that students who leave the charter sector are apt to move from schools where the quality of the match is high for schools where it is not as favorable. However, all these findings must be taken as suggestive only, as we are unable to disentangle fully student heterogeneity from school and match heterogeneity.

[19] This behavior appears perverse: why are students with above average gains moving out of this sector? It appears that exit decisions may have less to do with gains than level scores. Charter school leavers have lower level scores on average than their schoolmates, a circumstance that may lead them to feel out of place. These may also be students with fewer advantages at home, including parents less able or willing to ensure their children can continue at a charter school (e.g., by supplying transportation).

Implications for research

Estimates of charter school ETT that control for student fixed effects are biased when movers systematically differ from the population of all charter school students with respect to the schools they select or the quality of the school-student match. We are not the first to make this observation or investigate the magnitude of the biases.

In a study of Texas charter schools, Hanushek et al. (2005) note the following sources of possible bias when relying on a sample of movers to estimate charter school effectiveness: (1) a temporary negative shock to achievement the year before moving to a charter school; (2) bad school-student matches leading to exit from charter schools; (3) higher-than-average rates of exit from the least effective charter schools. In terms of our model, these three sources of bias correspond to selection on u_{it}, on α_{ij} and on α_j, respectively.

To investigate the first of these sources, the authors omit the year prior to a move from the estimation sample. (Their panel is sufficiently long to allow this.) The resulting charter school effects are very close to those obtained from the full sample, suggesting that temporary declines in achievement are not driving decisions to switch sectors or their estimates. However, there is evidence of bias from the second and third sources. The authors distinguish charter school entrants from students exiting the charter sector and estimate separate charter school effects for the two groups. The estimates are substantially more negative for those leaving charter schools, suggesting a bad match makes exit more likely. However, charter school effects continue to be negative (though smaller in magnitude) in the subsample composed of entrants.

Hanushek et al. also investigate whether school quality (as measured by the school's estimated value added for a fixed-weight "representative" mix of students) is associated with subsequent exits.[20] They find a strong inverse relationship between quality and the probability of exiting the charter sector for the public sector. The effect of public school quality on transfers in the other direction is also statistically significant, though somewhat weaker.

Bifulco and Ladd (2004) use a student fixed effects estimator to study North Carolina charter schools. They find the magnitude of the (negative) charter school effect is two to three times larger among students exiting the charter school sector than among entrants. The other sources of potential bias they examine do not appear to be important.[21]

Both of these studies also include models that correspond to our average difference estimator. In each case, the average differences favor traditional public schools over charter schools. However, in these studies, the choice between average differences and fixed effects estimates is not crucial to the qualitative conclusion that charter schools are less effective than traditional public schools.[22] Clearly this is not the case in our Idaho sample. An uncritical use of the fixed effects estimator on the grounds that it controls for selection bias would lead to the conclusion that charter schools in Idaho are more effective than the traditional public schools serving the same student population.

In long data panels spanning most, if not all, of the years that charter schools have been in operation, the biases discussed in this paper will be diminished. For example, the Texas data used in Hanushek et al. cover 1997 to 2002. As most charter schools were founded after 1997, a

[20] The weights reflect the statewide distribution of student characteristics. The result is something like an estimated ATE for each school, though it may be a poor indicator of what a school's value added would be if such a mix of students were actually enrolled there. Nor is it clear why this is the relevant measure of school quality for parents contemplating withdrawing their child.

[21] These include the effect of charter school competition on public school quality, the possibility that students moving from traditional public to charter schools were experiencing a downward trend in achievement that would have affected their gains anyway, and the possibility that the performance of charter school students declined due to changes in the impact across grade levels of observable student characteristics.

[22] In Hanushek et al. (2005) this conclusion is qualified: charter schools that have been in operation several years appear to be as effective as traditional public schools. They obtain this result in models with and without student fixed effects.

majority of the charter school students in their sample are observed first in a traditional public school, then in a charter school (and perhaps again in a regular school, if they leave the charter sector). To the extent that the majority of students who enroll in a charter school are movers, it is obviously less worrisome that movers may differ from stayers.

However, having a panel of this length does not remove all concerns. The Texas sample covers grades 4-7. About 25 percent of the observations are of students who were enrolled in charter schools as fourth graders. Because testing starts in grade 3, the first gain score for any student is the fourth grade-third grade difference. Unless these students subsequently leave the charter sector (in which case the evidence shows that they are apt to be atypical in some important respect), they do not contribute to the fixed effects estimator of charter school quality.[23] Likewise, the data set used by Bifulco and Ladd covers the entire period charter schools operated in North Carolina through 2001-2002. Nonetheless, of the 8745 students in their sample observed to attend a charter school at least once, test score gains in math can be computed in both sectors for only 5741, leaving 3004 charter school students who will not contribute to a fixed effects estimator.

Summary

We have investigated the effectiveness of Idaho charter schools relative to traditional public schools, using two estimators: a comparison of average test score gains in the two sectors, controlling for a limited number of student characteristics; and a student fixed effects estimator. Unlike other studies examining charter school performance, our findings are quite sensitive to the choice of estimator. When student fixed effects are included, charter schools

[23] To this must be added several hundred observations on students who were enrolled from 1997 forward in the charter schools already in existence in that year.

appear more effective than traditional public schools in the elementary grades. When student fixed effects are omitted, this is no longer true: charter school effects are negative or statistically insignificant at every grade.

The fixed effects estimator has been favored in the literature as a way of removing selection bias. However, the inclusion of student fixed effects can exacerbate biases stemming from heterogeneity in schools and in the quality of school-student matches, as the fixed effects estimator uses only those students who move between sectors to identify the charter school effect. As these students may not be representative of all charter school students, there should be no presumption that the fixed effects estimator provides the more accurate answer. Selection on individual heterogeneity does not appear to be a plausible explanation for the difference between estimates of charter school effectiveness obtained with and without student fixed effects in Idaho. Biases associated with school and school-student match heterogeneity appear to pose the greater problem.

References

Bifulco, Robert and Helen F. Ladd. 2004 The Impacts of Charter Schools on Student Achievement: Evidence from North Carolina. Durham NC: Sanford Institute of Public Policy, Working Papers Series SAN04-01.

Hanushek, Eric A. et al. 2005. Charter School Quality and Parental Decision-Making with School Choice. Cambridge, MA: NBER Working Paper 11252.

Hoxby, Caroline M. 2004. A Straightforward Comparison of Charter Schools and Regular Public Schools in the United States. Unpublished.

Hoxby, Caroline M. and Jonah E. Rockoff. 2005. The Impact of Charter Schools on Student Achievement. Unpublished.

Sass. Tim R. 2004. Charter Schools and Student Achievement in Florida. Unpublished.

Solomon, Lewis, Kern Paark and David Garcia. 2001. Does Charter School Attendance Improve Test Scores? The Arizona Results. Goldwater Institute.

Figure 1: Comparison of Achievement Levels: Spring 2005

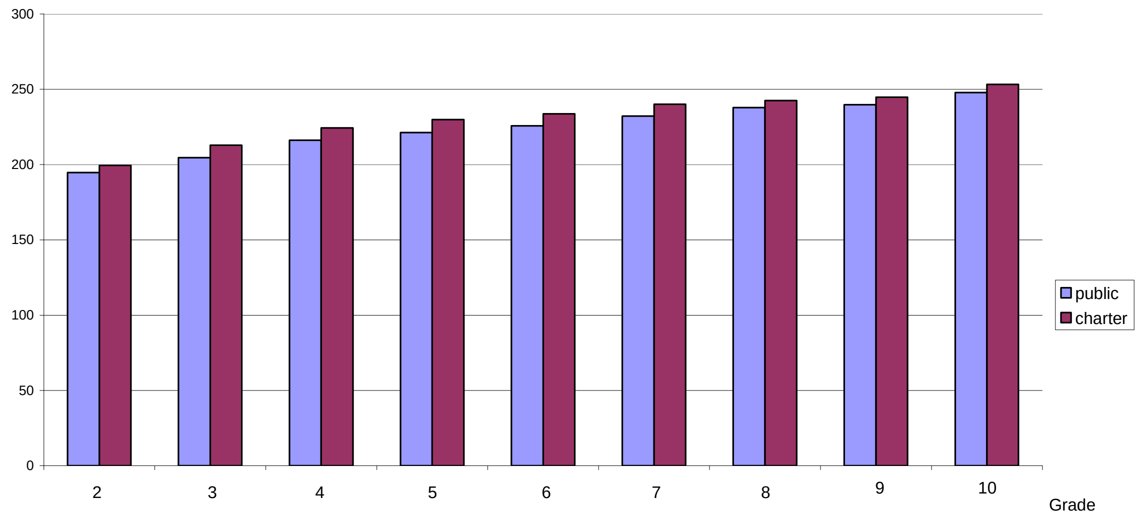

Figure 2: Comparison of Achievement Gains: Fall 2004 to Spring 2005

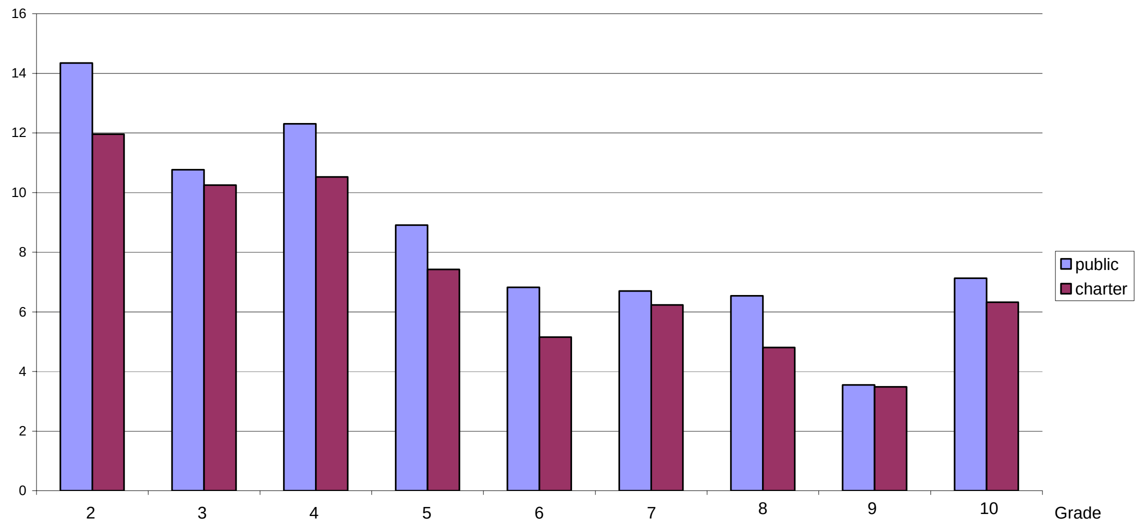

Figure 3a: Achievement Gains of Movers and Stayers: Elementary, 02-

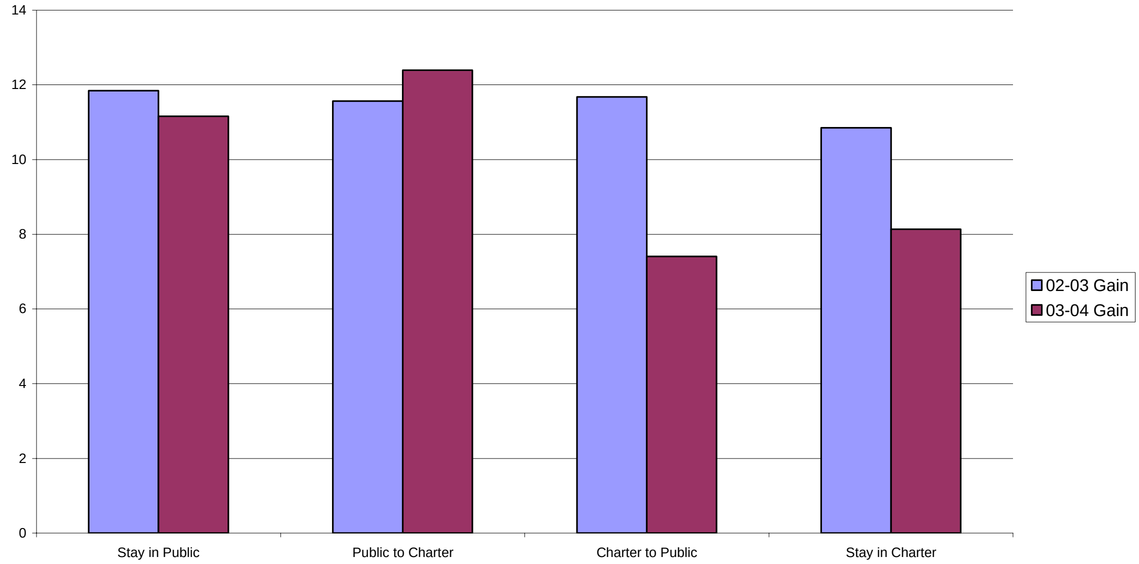

Figure 3b: Achievement Gains of Movers and Stayers: Elementary, 03-04 & 04-05

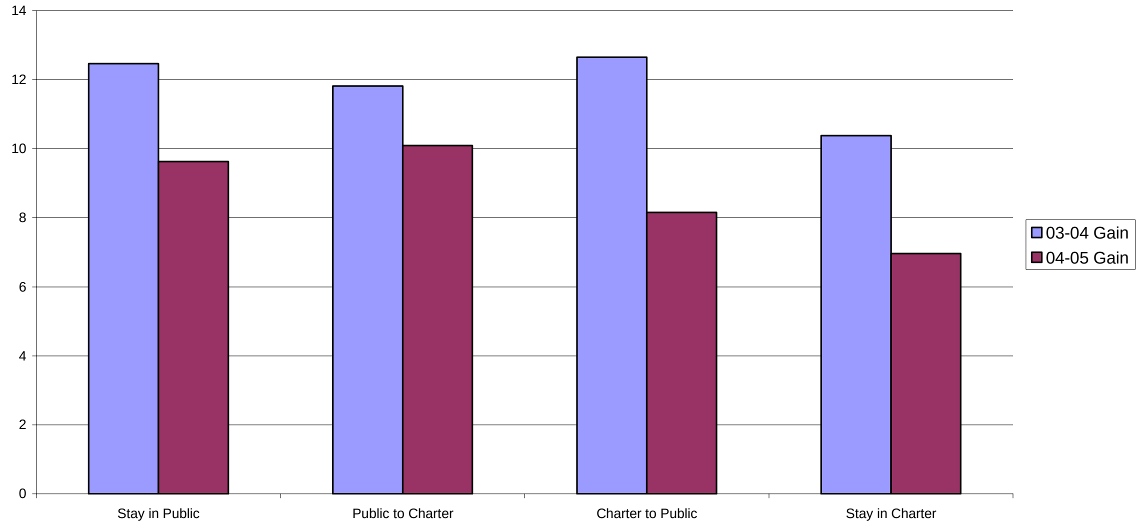

Figure 3c: Achievement Gains of Movers and Stayers: Secondary, 02-03 & 03-04

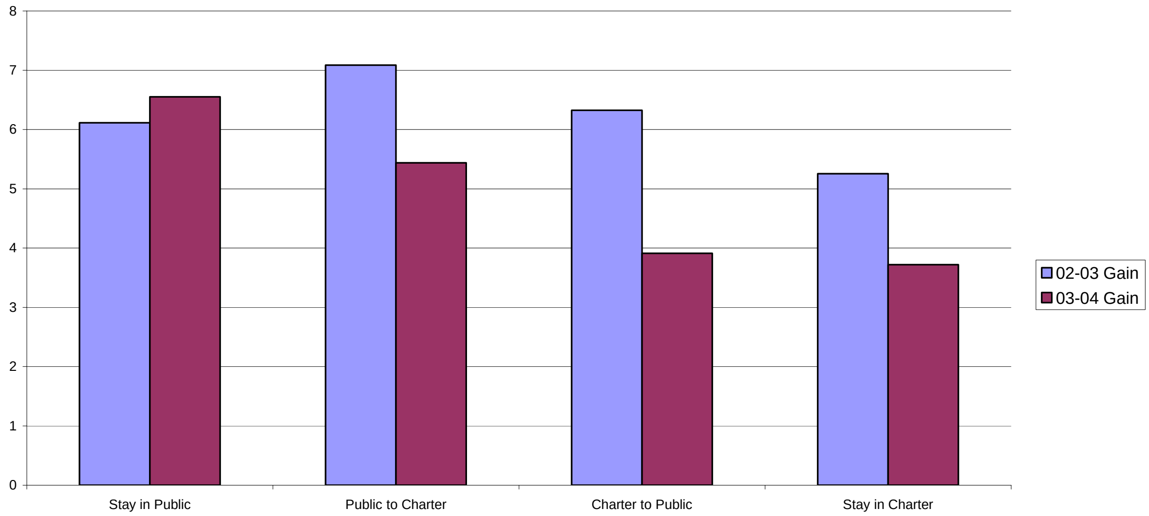

Figure 3d: Achievement Gains of Movers and Stayers: Secondary, 03-

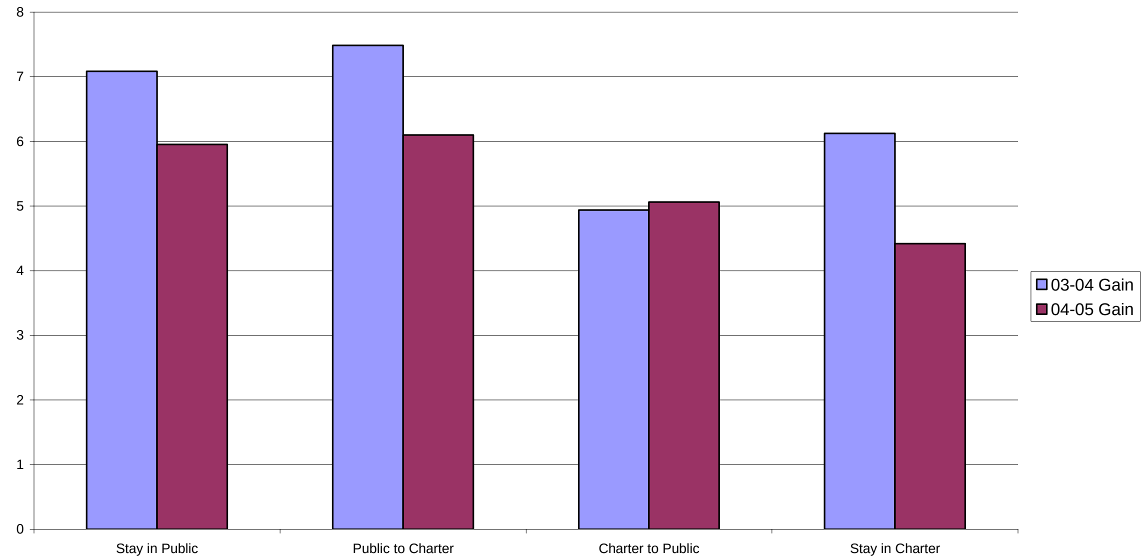

Table 1: Sample Characteristics

	Stayers: Traditional Public	Stayers: Charter	Movers
Percentage or mean			
White	84	94	94
Hispanic	12	3	3
Free or reduced-price lunch[a]	49	13	32
Special education	10	7	6
School days between fall and spring tests	136	134	136
Percentage of students in:			
Second Grade	11	17	7
Third Grade	11	13	10
Fourth Grade	11	12	11
Fifth Grade	11	11	14
Sixth Grade	11	11	15
Seventh Grade	12	12	15
Eighth Grade	12	9	14
Ninth Grade	11	8	10
Tenth Grade	10	8	5
No. of observations			
2002-03	123209	411	1247
2003-04	152886	1110	1221
2004-05	155255	1432	1232

a. Eligibility for free and reduced-price lunch is understated in charter schools, as some do not participate in the lunch program.

Table 2: Estimates of Charter School Effectiveness

	(1a)	(1b)	(2a)	(2b)	(3)	(4a)	(4b)	(5a)	(5b)
Estimator	AD	FE	AD	FE	AD	AD	FE	AD	FE
Grade									
2	-3.57	4.28	-3.56	4.19	-.69	-3.57	.21	.25	4.66
	(0.82)	(3.17)	(0.81)	(3.20)	(3.86)	(0.71)	(1.87)	(0.55)	(1.29)
3	-1.44	5.49	-1.52	5.31	-1.94	-1.44	3.26	1.12	4.26
	(1.05)	(1.73)	(1.01)	(1.79)	(1.82)	(0.90)	(1.11)	(0.65)	(0.73)
4	-1.97	3.06	-1.99	3.09	-2.63	-1.97	1.89	.59	2.11
	(1.34)	(0.78)	(1.33)	(0.78)	(1.62)	(1.24)	(0.76)	(0.57)	(0.60)
5	-1.63	1.50	-1.66	1.65	-1.75	-1.63	.66	.12	.75
	(1.21)	(0.86)	(1.18)	(0.87)	(1.34)	(1.18)	(0.97)	(0.54)	(0.81)
6	-0.55	.44	-.61	.43	-1.27	-.55	-.23	.66	.74
	(1.76)	(1.12)	(1.91)	(1.25)	(1.37)	(1.82)	(1.09)	(0.52)	(0.42)
7	-2.02	-2.11	-2.03	-2.09	-1.63	-2.02	-2.48	-.57	.84
	(1.04)	(1.05)	(1.03)	(1.07)	(1.02)	(0.80)	(1.03)	(0.37)	(0.46)
8	-1.87	-.68	-1.97	-.72	-1.36	-1.87	-.82	-.79	.82
	(0.57)	(0.67)	(0.58)	(0.68)	(0.61)	(0.38)	(0.72)	(0.38)	(0.45)
9	-0.74	-.49	-.69	-.37	-.72	-.74	-.48	-.29	.31
	(0.82)	(0.54)	(0.89)	(0.59)	(1.05)	(0.81)	(0.48)	(0.34)	(0.50)
10	-1.07	-.24	-1.10	-.27	.26	-1.07	-.29	.29	4.66
	(1.03)	(0.91)	(0.99)	(0.98)	(0.98)	(1.00)	(0.84)	(0.34)	(1.29)
Other regressors:									
Race x Grade	yes	yes	yes	yes	yes	yes	yes	yes	yes
Special education x Grade	yes	yes	yes	yes	yes	yes	yes	yes	yes
Year x Grade	yes	yes	yes	yes	yes	yes	yes	yes	yes
Moved between schools	no	no	yes	yes	no	no	no	no	no
Base year (fall, 2002) score	no	no	no	no	yes	no	no	no	no
Years of charter operation	no	no	no	no	no	yes	yes	no	no
N	438516	438505	438516	438388	230727	438516	438516	241192	241192

Table 3: Effect of Charter School Years of Operation, Movers vs. All Students

	Average Charter Age		Effect of School Age on Student Gain Scores					
			Using AD Coefficient Estimates			*Using FE Coefficient Estimates*		
Grade	All Students	Movers	All Students	Movers	Difference	All Students	Movers	Difference
2	3.6	4.2	-2.98	-3.46	-0.48	-13.10	-15.20	-2.10
3	3.6	2.0	-3.51	-1.99	1.52	-6.14	-3.48	2.66
4	3.6	2.4	-4.42	-2.86	1.56	-3.47	-2.24	1.23
5	3.6	2.2	-2.58	-1.58	1.00	-0.94	-0.58	0.36
6	4.1	3.7	-1.31	-1.19	0.12	-1.82	-1.66	0.16
7	3.8	3.3	-3.80	-3.26	0.44	-4.22	-3.61	0.61
8	4.0	3.4	-2.58	-2.18	0.40	-2.45	-2.07	0.38
9	4.6	4.5	-0.69	-0.68	-0.01	-2.60	-2.55	-0.05
10	4.8	4.3	-3.06	-2.75	0.31	-1.62	-1.45	0.17

Column 3 represents the effect of an additional year of operation on mean gains multiplied by the average number of years of operation over all students at the indicated grade level. Column 4 represents the same coefficient multiplied by the average years of operation over movers (typically fewer).

Columns 6 and 7 are defined likewise, using coefficients from the fixed effects model rather than the average difference model.

Table 4: Selection on School Quality

	Change in Mean Gain Among Non-Movers Associated with an Additional Mover: Elementary	Secondary
Destination Schools For		
Charter Entrants	.10	.11
	(.02)	(.02)
Traditional Public Entrants	-0.13	-0.06
	(.09)	(.02)
Schools of Origin for		
Charter Leavers	.21	-.10
	(.26)	(.13)
Traditional Public Leavers	-.05	-.000
	(.02)	(.02)

Mean gains for destination schools are calculated for the post-move year.
Mean gains for schools of origin are calculated for the pre-move year.

Regression equation controlled for race, special education, and school size at elementary and secondary levels. Standard errors in parentheses.

Table 5: Confounding of School Selection and Charter School Effectiveness

Grade	Baseline Model, Reduced Sample	Controlling for School Quality
2	1.75	-2.09
	(1.48)	(1.46)
3	4.41	1.95
	(1.09)	(1.08)
4	3.91	2.37
	(0.99)	(0.97)
5	3.12	1.75
	(0.91)	(0.90)
6	1.26	-1.25
	(0.82)	(0.81)
7	-2.23	-1.55
	(0.81)	(0.80)
8	-0.42	-0.04
	(0.86)	(0.85)
9	0.28	-0.47
	(0.90)	(0.89)
10	0.93	3.45
	(1.38)	(1.36)

Sample: 2003-04 and 2004-05. Column 1 displays estimated charter school effects from baseline model. Model in column 2 controls for destination school quality, as measured by residual gains of non-movers. Standard errors in parentheses.

Table 6: Gains of Charter School Leavers Relative to Stayers

Pre-move grade	Pre-Move Difference	Post-Move Difference
2	2.59	-4.76
	(3.38)	(3.17)
3	6.61	-1.06
	(3.38)	(3.17)
4	0.56	-2.48
	(2.74)	(2.56)
5	3.37	4.92
	(2.33)	(2.18)
6	-1.86	0.56
	(1.65)	(1.56)
7	-0.21	-2.64
	(2.00)	(1.98)
8	-1.52	1.47
	(1.33)	(1.28)
9	-1.54	1.05
	(1.67)	(1.63)
N	1849	1777

Standard errors in parentheses.
Sample consists of all charter school students in 2002-03 and 2003-04 not enrolled in virtual schools.
Leavers are students who enrolled in traditional public schools the following year.
Pre-move difference equals leavers' gains minus stayers' gains for the year prior to the move.
Post-move difference equals leavers' gains minus stayers' gains for the year following the move.
Differences derived from regression equation controlling for race & special education by grade.
Post-move sample size is smaller due to students observed in neither the charter nor the traditional public sector following the year of move.