Charter, Private, Public Schools and Academic Achievement:

New Evidence from NAEP Mathematics Data[1]

Christopher Lubienski
University of Illinois
club@uiuc.edu

Sarah Theule Lubienski
University of Illinois
STL@uiuc.edu

(January 2006)

National Center for the Study of Privatization in Education
Teachers College, Columbia University
525 West 120th Street, Box 181
New York, NY 10027
ncspe@columbia.edu

[1] This project was funded through a National Assessment of Educational Progress Secondary Analysis Grant (#R902B05017) from the National Center for Education Statistics, Institute of Education Sciences. The authors would like to thank Jane Loeb, Eric Camburn, and Jay Verkuilen for their helpful comments on this work, and also Corinna Crane, who provided valuable assistance on this project. Of course, only the authors are responsible for the analysis and interpretations presented in this report.

EXECUTIVE SUMMARY

This analysis of US mathematics achievement finds that, after accounting for the fact that private schools serve more advantaged populations, public schools perform remarkably well, often outscoring private and charter schools.

Recent debates have highlighted the issue of school sector as an important consideration in student academic achievement. In 2004, a report contending that charter school students scored lower than students in public schools was fiercely contested. Other studies were then released to demonstrate that charter schools produce greater gains in student learning. Similarly, studies of students using vouchers to attend private schools have ignited heated debates about whether or not these programs boost student achievement, especially for poor and minority students.

Common wisdom and past research holds that private schools achieve better academic results. Assumptions of the superiority of private-style organizational models are reflected in voucher and charter programs, and in the choice provisions of the No Child Left Behind Act. According to this thinking, schools in the choice-based independent sector are the best model for improving achievement in public schools. Market-oriented school choice reforms are premised on the idea that, by positioning parents as the driving force in the quest for quality, schools will be forced to improve when faced with competition from higher performing rivals.

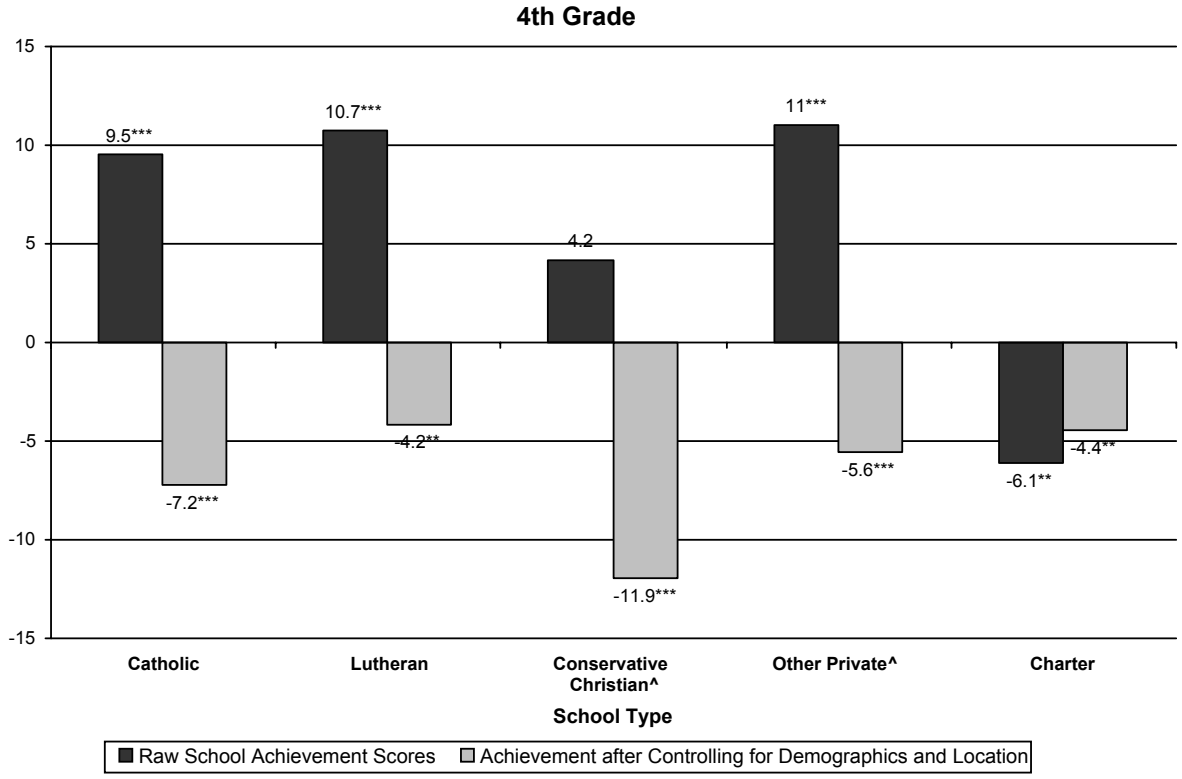
However, new results from a study of a large, comprehensive dataset on US student achievement seriously challenge assumptions of private school superiority overall, and find substantial differences between different types of private schools. Based on the 2003 National Assessment of Educational Progress (NAEP) mathematics exam, this analysis compares achievement in public, charter, and different types of private schools. When compared with other subjects (like reading, for instance), math is more heavily influenced by school than home experiences, so studying math achievement provides clearer insights into the relative performance of different types of schools. The 2003 NAEP samples are over ten times larger than in any previous NAEP administration, providing achievement and student, teacher, and administrator survey data on over 190,000 4^{th} graders (up from 13,855 in 2000) in 7485 schools, and more than 153,000 8^{th} graders (up from 15,930 in 2000) in 6092 schools. Earlier studies of charter schools based on these data were disputed because researchers had to rely on a web tool that did not allow for simultaneous analysis of multiple student- and school-level variables. This new analysis of the complete raw data employs advanced statistical techniques (hierarchical linear modeling) to study the relationship between school type and mathematics achievement while controlling for demographic differences in the populations served by the schools.

Major Findings

Without controlling for student background differences, private schools scored higher than non-charter public schools, as would be expected. However, this study examines these patterns further, determining whether they are due simply to the fact that higher proportions of disadvantaged students are enrolled in public schools, and the extent to which the gaps persist

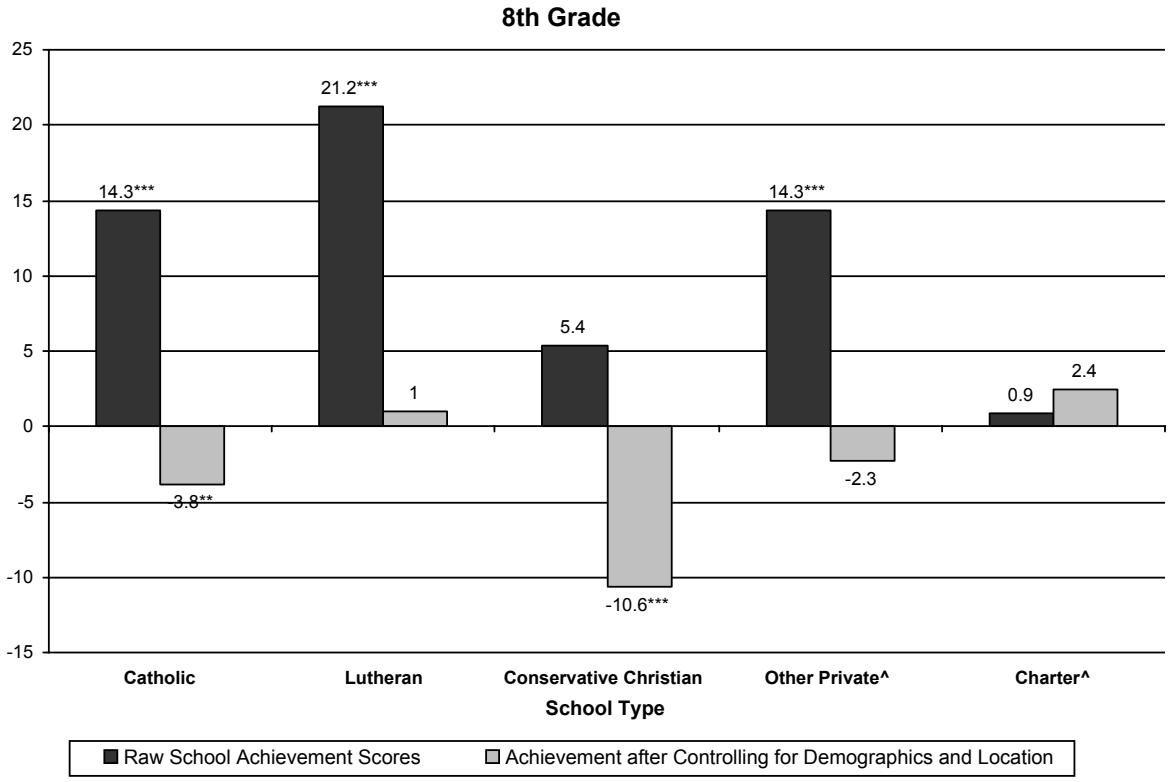
after controlling for potential student- and school-level confounding variables, including measures of socioeconomic status, race/ethnicity, gender, disability, limited English proficiency, and school location. ***Overall, the study demonstrates that demographic differences between students in public and private schools more than account for the relatively high raw scores of private schools. Indeed, after controlling for these differences, the presumably advantageous "private school effect" disappears, and even reverses in most cases.***

These graphs show the disparities in school achievement before and after controlling for differences among schools. The 0-axis represents average public school achievement, and the bars show the difference between the average achievement of the major school types and that of public schools, with and without controls for demographics and location.



** p<.05 ** p<.01 *** p<.001 (Indicates means are significantly different than corresponding public school means)*

^ Participation rates did not meet NCES reporting requirements. Results for these subsamples should be interpreted with caution.



** p<.05 ** p<.01 *** p<.001 (Indicates means are significantly different than corresponding public school means)*
^ Participation rates did not meet NCES reporting requirements. Results for these subsamples should be interpreted with caution.

To summarize the most important findings once demographic and location differences were controlled:

- Public schools significantly out-scored Catholic schools (by over 7 points in 4th grade, and almost 4 points in 8th grade).
- Of all private school types studied, Lutheran schools performed the best. 4th grade scores in Lutheran schools were roughly 4 points lower than in comparable public schools, but were (a statistically insignificant) 1 point higher at the 8th grade.
- The fastest growing segment of the private school sector, conservative Christian schools, were also the lowest performing, trailing public schools by more than 10 points at grades 4 and 8.
- Charter schools scored a significant 4.4 points lower than non-charter public schools in 4th grade, but scored (a statistically insignificant) 2.4 points higher in 8th grade.

These notable findings regarding the remarkable performance of public schools are significant, not just statistically, but also in terms of their policy implications. The presumed panacea of private-style organizational models — the private-school advantage — is not supported by this comprehensive dataset on mathematics achievement. These data suggest significant reasons to be suspicious of claims of general failure in the public schools, and raise substantial questions

regarding a basic premise of the current generation of school reforms based on mechanisms such as choice and competition drawn from the private sector. Furthermore, assumptions that academic quality will be driven by parental choice need to be re-examined in view of the fact that conservative Christian schools, the fastest growing segment of the private school market, were also the lowest performing.

Explanations for these noteworthy findings offer considerable insights into the state of American education, and into the remedies proposed to change US schools. Further results from these data will be released soon, including analyses of:

- the relationship between school size and academic achievement.
- the relationship between teacher certification and student achievement.
- interactions among school sector, school climate, and achievement.

Technical Notes on NAEP Data

NAEP mathematics results are reported on a 0-500 scale, with the 2003 mean being 235 at grade 4 and 278 at grade 8. The NAEP mathematics scale was originally designed to allow cross-grade comparisons, indicating that the 43-point difference between grade 4 and grade 8 would mean that a gap of 10 or 11 points represents roughly one grade level. NAEP no longer maintains that the scale is consistent across grades. However the scales have not changed markedly, and therefore the idea that "10-11 points is approximately 1 grade level" is still often used as a helpful, albeit rough, guide for interpreting score differences. Another way to interpret the magnitude of these disparities is in terms of effect sizes. Standard deviations for school achievement means were 14.7 at grade 4 and 19.3 at grade 8. Hence, a 10-11 point difference represents a moderate to large effect (roughly .5-.7).

Sample sizes for some private school types were relatively small (ranging from 78-224 schools), and reduced slightly further in the full HLM models (57-203 schools). The participation rate for 4th grade conservative Christian schools, for 8th-grade charter schools, and for "other private" schools at both grades did not meet NAEP's stringent reporting requirements; hence, results for these particular sub-samples should be viewed as only suggestive of patterns that may exist in the entire population of such schools.

Finally, NAEP data are cross-sectional, not longitudinal, so they do not allow for analyses of value-added effects. However, HLM techniques allow researchers to account for the primary possible confounding variables that could explain patterns in these data (making it less likely that longitudinal data would tell a different story).

Charter, Private, Public Schools and Academic Achievement:
New Evidence from NAEP Mathematics Data

A government report on mathematics results from the 2003 National Assessment of Educational Progress (NAEP) highlighted this finding: "Public-school students scored lower on average than non-public-school students at both grades 4 and 8."[2] This finding is nothing new — private school students in the United States have typically scored higher than public school students on standardized tests, confirming the perception among the US public and policymakers that private schools are inherently more effective than public schools.

However, the real question for researchers and policymakers is whether differences in test scores between various school types — public schools, charter schools, or Catholic and other private schools — are primarily due to differences in the student populations served by these different sectors. Earlier research on school sector effects indicated that private schools score higher even after adjusting for the fact that these schools tend to serve students with fewer "risk" factors. And yet, a few more recent studies have found instances where, after employing statistical controls to account for the differences in student populations, achievement in public schools has been equal to or, in some cases, higher than those of private schools (as described below).

Furthermore, this debate re-ignited recently around the issue of charter schools — publicly funded but independently managed schools of choice. A report comparing 2003 NAEP achievement levels in public schools and charter schools, based on data publicly available on the internet, concluded that students in public schools outscore their peers in charter schools. However, the analysis was criticized for the limited information on student background, and for controlling for student characteristics individually, instead of multiple characteristics simultaneously — limitations of the web-based NAEP data tool.

Now, these NAEP data are available to researchers for use in raw form, offering the opportunity to examine simultaneously relationships among school organizational type, demographic characteristics and mathematics achievement in private, charter, and non-charter public schools. NAEP is often referred to as "The Nation's Report Card" due to being the only nationally representative, ongoing assessment of US academic achievement in various subject

[2] http://nces.ed.gov/nationsreportcard/mathematics/results2003/schooltype.asp

areas. The 2003 sample is over ten times larger than in any previous NAEP administration, making it possible to closely examine school characteristics and achievement within public, charter,[3] and several major types of private schools.

This is a particularly salient issue now. Policymakers' belief in the superior organizational attributes of private schools serves as the primary premise for the current generation of school reform — exemplified, for instance, in voucher and charter programs, and in the choice provisions of the No Child Left Behind (NCLB) Act of 2001. Although beliefs about what is known as the "private school effect" have been widespread for some time now, public and private school sectors have since undergone significant changes. Furthermore, such generalized assumptions about school sectors may miss important differences not only between, but also within, sectors.

While it is commonly known that average achievement in private schools tends to be higher than in public and charter schools, this study examines this pattern further, examining whether relationships between school type and achievement persist after controlling for demographic and other differences among schools.

There are reasons to expect that private schools — and schools such as charter schools, which borrow important attributes from private school models — would have higher achievement than public schools, even after accounting for demographic differences in their clientele. Private schools are free of much of the bureaucracy that plagues schools in the public sector, and are able to avoid political entanglements around controversies such as textbook adoption or graduation prayers in focusing on a core academic curriculum. Furthermore, in private and charter school sectors, parents are positioned to select a school based on academic quality, and to choose another if their school fails to meet expectations. In light of such factors, it seems reasonable to anticipate that schools such as these have structural incentives to excel. Additionally, private and charter school parents, through the act of choosing, demonstrate a commitment to their children's education — a characteristic that goes beyond typical SES measures and is associated with higher student achievement.

[3] We are aware that charter schools are, legally, also "public" schools. Yet they also resemble private schools in some important respects, particularly in terms of autonomy from local education authorities — as reformers intended for these schools. For the sake of clarity in categories in the analysis, we use the term "public school" as short-hand for non-charter public schools; the "independent sectors" include private and charter schools.

On the other hand, private and charter schools might emphasize non-academic considerations in attracting families (especially since it can be difficult for parents, and researchers, to make accurate judgments about academic quality). For instance, they may hire teachers based more on religious beliefs than academic training, or choose curricula that put a particular worldview ahead of other considerations such as boosting test scores. It is also possible that some parents choose private or charter schools for their children who have struggled academically in traditional public schools, so that these children enter private schools with relatively low test scores. Moreover, public schools are thought to be better resourced than independent (private and charter) schools, on average, and public school teachers and administrators are required to receive a certain level of training. If such factors are linked to higher student achievement, then we might expect that public schools could perform relatively well. Indeed, in our preliminary work on this topic, using the much smaller 2000 NAEP dataset, we found that public schools were, in fact, performing as well as private schools on mathematics achievement measures (Lubienski & Lubienski, 2005).[4]

Consistent with what is typically reported, the raw 2003 NAEP mathematics achievement means for private schools were indeed higher than public schools means, which were similar to or higher than charter school means. This study employs hierarchical linear modeling to examine the extent to which these gaps persist after controlling for SES, race/ethnicity, gender, disability, limited English proficiency (LEP), and school location.

In the following section, we outline the theory and past research on the role of school sector in academic achievement, focusing on the implications for current reform policies. We then describe the present analysis, including the variables and controls used to compare different types of schools that serve different student populations. The findings from this study include some consistent patterns that challenge much of the conventional wisdom on this topic. Specifically, when controlling for differences in student populations, public school achievement is roughly equal to or higher than that of other school types. The study also illuminates some important differences in school performance within the independent school sectors. We conclude with a discussion of the implications of these findings with regard to both future research and present policy.

[4] This result was uncovered accidentally when a binary school sector variable was employed as a control in a study of reform-oriented instruction and mathematics achievement. Our surprise at the negative coefficient for private schools prompted us to conduct a more in-depth study on the issue.

Theory & Research on the Impact of School Sector on Achievement

Over the last quarter-century, much research has been conducted on the relative effects of school organizational type on academic achievement. This line of inquiry was colored by theories of the relative merits of public and private sectors in education and other endeavors, and had obvious implications for the allocation of government resources — that is, if the private sector can educate children more effectively (and for less money), then it is difficult to justify the exclusive hold that public schools have on public funding for education (Lubienski, 2001, 2003c). Indeed, the private sector is known to offer advantages over the public sector in the production and distribution of a range of goods and services. So the question becomes, Why is it necessary that the government operate schools for the public if the private sector can better serve that function (Friedman, 1955, 1962)?

In this line of thinking — based in neoclassical economic analyses as applied to institutions like public education (and known as Public Choice theory or, in education, "market theory") — a school's institutional environment, or sector, shapes its organizational structure and the external incentives that drive its internal productive processes.[5] Public schools are input-oriented organizations, accountable to bureaucracies, not to consumers, so they lack structural incentives to innovate, improve, or respond to demands for quality from the groups that they serve (Chubb & Moe, 1990; Coleman, 1997). Thus, it is believed, students stuck in the public sector should be given the opportunity to switch to the higher performing schools in the private sector. Indeed, not only are private schools free of much of the bureaucracy and regulation that inhibit performance in the public sector, but they are not shielded from competition as public schools are. They must demonstrate greater effectiveness in terms of their outputs in order to attract families willing to pay tuition. So, while private schools tend to draw more advantaged families that can afford the added costs, if such schools can be shown to achieve superior results with the same types of students who attend public schools, then there is a stronger argument for policies that encourage students to leave government-run schools for schools in the private sector. Indeed, not only would this be a more efficient and effective use of public resources in educating the public, but there is a serious equity concern about trapping poorer families in the underperforming public sector.

[5] For more about Public Choice in education, see Chubb and Moe (1990) and Lubienski (2003b, 2005).

While initially a hypothetical argument emerging out of economic theory applied to public schooling, this issue assumed a more immediate tone when researchers began to collect actual evidence on the relative performance of schools in public and private sectors. This research goes back almost three decades, but has become more pointed in recent years in its implications for education reform.

In addition to the seminal work on achievement in public and private schools, two recent debates — on charter schools and voucher programs — have highlighted the question of the role of school organizational type in promoting academic achievement. These research literatures on public and independent schools point to the presumed importance of public and private sectors for understanding the organizational behavior of schools and, ultimately, the propensity of different models of schooling to effect superior academic performance.

Earlier Research on Public & Private Schools

Charter schools and voucher programs are current manifestations of policymakers' assumptions regarding the superior performance of private or independent schools. These assumptions are based, at least in part, on solid, well-respected precedents in the research literature comparing representative samples of schools in public and private sectors. A number of influential studies were based on the High School and Beyond (HSB) dataset, a longitudinal study of over 58,000 students in approximately 1,000 schools. These data allowed researchers to examine achievement in public and private schools while controlling for a number of school and family variables. The initial report by James Coleman and his colleagues found a notable private school effect — inherent advantages for schools in the private sector that resulted in greater academic achievement even after controlling for differences in student populations (Coleman & Hoffer, 1987; Coleman *et al.*, 1981, 1982). Even as these findings were being published, the results were challenged by a number of scholars who contested the methods, disputed the implications the authors drew regarding school choice, or found little or no evidence of a private school effect in the data (see, e.g., Alexander & Pallas, 1985; Bryk, 1981; Catterall & Levin, 1982; Goldberger & Cain, 1982; Willms, 1985). Yet in another set of influential studies, Anthony Bryk and colleagues followed this line of inquiry with the HSB data, using the representative sample of Catholic schools, and a comparable subsample of public schools. This work, highly regarded in the research community, advanced from a recognition of substantially

greater academic achievement independent of student background in Catholic schools, and then focused primarily on *why* such schools were more effective (Bryk *et al.*, 1984; Bryk *et al.*, 1993; Lee & Bryk, 1989). These researchers found the Catholic school effect to be substantial for disadvantaged students, and highlighted the unique social and academic organizational characteristics, communities based on shared values, a curriculum that engages students in core academic subjects, and distinctive forms of school governance (see also Greeley, 2002). In perhaps the most provocative work with the HSB data, Chubb and Moe (1990) tied such findings to Public Choice theory's emphasis on inherent pathologies in public sector institutions — providing empirical justification for programs that would enable students to switch to private sector schools.

More recent work on another dataset, the National Education Longitudinal Study of 1988 (NELS), offers additional insights into achievement in public and private secondary schools. This random sample of 25 8th graders in each of a thousand schools was supplemented by several subsequent rounds of data collection on these same students, providing multi-point data for comparison of students who attended public, Catholic, and other private schools. In the subsequent follow-up two years later, researchers noted that Catholic and independent private schools outscored public schools, although there were no statistically significant differences in the two-year growth in mathematics achievement between public schools and these other schools; there was, however, a significant difference between types of private schools, with Catholic school gains outpacing those of independent private schools (Scott *et al.*, 1995). Dan Goldhaber (1996) examined a subsample from the NELS data of over 3,000 students each in mathematics and reading. After controlling for the fact that the private school students come from more affluent and educated families, he found no achievement advantage in private schools. In another NELS study of 4,000 students in public, magnet (public), Catholic, and secular private schools in urban areas, Adam Gamoran (1996) found advantages for students in magnet schools in several subjects. Catholic schools were shown to have a positive effect in mathematics, while secular private schools offered no advantage over public schools. Drawing also on the ensuing wave of NELS data collection (through 12th grade), Figlio and Stone (1997) reported on student achievement for over 5,000 students in public, private-religious, and secular private schools. Accounting for selection effects, they found evidence of a slight but significant negative private school effect for math and science achievement in religious schools relative to

public schools, except for urban minorities who were found to benefit from religious schools, while secular private schools were shown to offer a substantive advantage in these subjects. Grogger and Neal (2000) examined NELS data through the 1994 wave of data collection, finding no significant Catholic school effect on mathematics achievement for suburban students, but a moderate effect for urban White students, and larger gains for urban minority students (as well as a substantial Catholic school impact on other outcomes such as graduation rates and college attendance). Yet they also found that independent private schools did not enhance student achievement any more than did public schools. More recently, Kim and Placier (2004) found significant differences in a subsample of 144 private schools in the NELS data, with non-Catholic schools outperforming Catholic schools in reading, but not in mathematics.

Obviously, these studies together present a rather blurred picture of the impact of different school sectors on student achievement — much more mixed than is commonly assumed in popular and policy wisdom on the superiority of the private school sector. The findings have often varied by subject area, sometimes supporting a somewhat modest private or Catholic sector effect, but often depending on factors such as school location and student ethnicity. Indeed, findings from both HSB and NELS suggest that results are quite sensitive to methodological and sampling issues (Grogger & Neal, 2000). Additionally, the studies have focused on grades 8-12, raising the unaddressed question of how patterns in earlier grades might compare. However, the research focus has turned increasingly away from sector effects on achievement to more specific questions — for example, course-taking and retention in urban Catholic high schools.[6] Furthermore, the literature on sector effects on achievement is becoming dated. Students represented in the NELS data graduated in 1992; students in HSB in 1982. In the meantime, substantial changes have continued to re-shape private schools and the populations they serve — for instance, many urban Catholic schools are closing or enrolling higher proportions of minority (and non-Catholic) students, and homeschooling increasingly draws students from both public and private sectors (Broughman & Pugh, 2004; Grogger & Neal, 2000; Lubienski, 2003a; Perie *et al.*, 2005).

Schools are now operating in a different policy context. In the HSB literature, researchers were writing at a time when vouchers were essentially just an idea. The NELS literature came to

[6] See, for instance, work by Lee and associates (Lee & Burkam, 2003; Lee *et al.*, 1998). This research trend could be due to the belief that the work by Bryk *et al.* on HSB largely settled the issue of the academic superiority of Catholic schools (Greeley, 2002).

the fore when charter schools were just beginning their rapid proliferation, and some researchers also noted the potential significance of their analyses of NELS for the nascent publicly funded voucher programs, which had not yet been found to be constitutional (e.g., Figlio & Stone, 1997; Goldhaber, 1996; Grogger & Neal, 2000). Now, policymakers' assumptions about sector effects are re-shaping the education landscape in the US. Voucher and charter school programs premised on the assumption of private or independent school effects are increasingly popular with policymakers. And NCLB endorses school choice and conversion to charter status as sanctions for under-performing public schools, which are thought to be lacking internal incentives to succeed.

School Voucher Programs

Policymakers' belief in the superiority of private schools is evident in the emergence of voucher programs in the US — programs that give students the opportunity to leave public schools and attend private schools. One of the tenets of the thinking undergirding voucher plans holds that the relative autonomy of schools in the private or independent sector, combined with the lack of a captive clientele, creates the opportunities and competitive incentives for private schools to achieve superior levels of academic achievement; consequently, vouchers allow students otherwise trapped in the public schools the chance to learn more at these higher performing private schools (Chubb & Moe, 1990; Friedman, 1955).

Although a famous early experiment in the 1970s was much studied, few researchers look to the results now because the actual implementation of the Alum Rock program compromised the integrity of the voucher plan as a true source of insight on choice across sectors (Bridge & Blackman, 1978; Lines, 1993). The more recent creation of a voucher program in Milwaukee in the 1990s drew much attention in allowing public school students to attend private schools, eventually including schools run by religious groups. After accounting for differences between the public school population and the students using vouchers to attend private schools, the official evaluation of the outcomes indicated no significant private school effect on student performance in math and reading (Witte, 2000). However, in another analysis of this data, researchers using different control groups in a randomized model found substantial gains in mathematics and reading achievement for public school students who used vouchers to switch from public to private schools (Greene *et al.*, 1996; Greene *et al.*, 1998b). In a later study,

Cecelia Rouse found significant gains in math (but not in reading) for students using vouchers to attend private schools, but suggested that these gains were similar to (and, in reading, outpaced by) those of public school students in smaller classes (1998a, 1998b). (The legislature subsequently ended the evaluation of the "experiment," so participating schools are no longer required to provide data on student achievement.) A second city-wide voucher program in Cleveland generated similarly conflicting findings (Greene *et al.*, 1998a; Metcalf *et al.*, 2003).

While the Milwaukee and Cleveland programs were publicly financed, vouchers funded by private groups have also been examined in a number of cities. In perhaps the most publicized research, Paul Peterson and colleagues examined programs in New York City, Dayton, and the District of Columbia. The researchers used experimental models in comparing students who were randomly selected to receive a voucher to those who applied but were (randomly) denied — thereby allowing researchers to study students essentially equal on both observable and unobservable (i.e., motivation) characteristics. This approach is thought to distinguish the institutional effects of schools from other confounding variables such as family background. The researchers found no boost in math or reading achievement for White and Hispanic students, but significant gains for African-American students using vouchers to switch from public to private schools (Howell *et al.*, 2000; Mayer *et al.*, 2002; Myers *et al.*, 2000). A review by the General Accounting Office (2002) noted that these gains were largely relegated to African-American students in early grades in New York City. A secondary analysis of the data disputed the initial findings, arguing that a substantial proportion of students were excluded from the study, and that an inappropriate method of racial categorization was used — and that these two factors accounted for the finding of a private school effect, which was otherwise insignificant (Krueger & Zhu, 2004a, 2004b). This re-analysis of the data has been contested, and the original findings defended by the authors (Howell & Peterson, 2004; Peterson & Howell, 2003, 2004).

As with the research literature on achievement from larger-scale data in HSB and NELS, it is important to note the central position of school sector for researchers trying to understand relative gains in academic achievement. The institutional location of a school in the private or independent sector is assumed to generate incentives that will cause that school to produce greater gains in academic achievement than will public schools, even with the same students when they switch from public to private schools. Yet these studies of voucher programs typically deal with only a discrete number of schools involved in a local program, or focus on a

select number of students who cross sector boundaries, rather than with samples of students representing whole sectors. Local policy and contextual factors become a predominant consideration for these studies, and the over-riding assumption behind these programs — that, in general, private schools outperform public schools — is partially obscured from examination behind the fact that only a proportion of the private schools in these areas actually participate in these programs and are therefore included in these studies.

Charter Schools

In August of 2004, the American Federation of Teachers (AFT) released an analysis of 2003 NAEP data available from a web-based data tool, indicating that students in charter schools lagged behind comparable public schools (Nelson *et al.*, 2004; Schemo, 2004). The authors found that charter school students trailed their public school counterparts by about a half-year in mathematics and reading in grades 4 and 8, and proportionately fewer charter school students reached proficiency levels. Furthermore, the analysis noted that public school students whose family incomes made them eligible for free or reduced-priced lunches outscored comparable students in charter schools. This study was immediately criticized by charter school advocates because of its methodology, due largely to the inherent limitations of the NAEP web-tool: for instance, its reliance upon free/reduced lunch eligibility as the only available SES measure, failure to control for multiple confounding variables simultaneously (e.g., race/ethnicity and SES), and lack of attention to the data's multi-level structure (students nested within schools). (For a more detailed description of this debate, see Carnoy *et al.*, 2005). A few months later, a similar US Department of Education (2004) analysis largely concurred with the AFT report, finding that — accounting for the race/ethnicity of students — proportionately fewer 4th and 8th grade charter school students were likely to reach proficiency levels on NAEP math and reading than comparable students in other public schools. Differences in mathematics performance persisted even after controlling for differences between the charter and non-charter schools on other single variables such as location, teacher experience, subsidized lunch eligibility, or student disability.

However, in another set of studies, Caroline Hoxby found that charter school status had a positive effect on academic achievement, with these independent schools outperforming public schools in math and reading proficiency. In a study encompassing virtually all charter school

students in the US, Hoxby (2004a, 2004b) studied achievement on state exams for students in charter schools compared to their counterparts in the nearest public school (or to the nearest public school of similar racial/ethnic composition). Students in charter schools were found to be more likely to achieve proficiency on state math and reading exams than their counterparts in the schools they presumably would have attended without the charter school option. Thus, it was argued that school type *causes* differences in achievement: "the average student attending a charter school has higher achievement than he or she otherwise would" (Hoxby, 2004b, p. 3). The findings suggested that charter schools were particularly beneficial for poor and Hispanic students. In another study of the effects of charter status on academic achievement, Hoxby and Rockoff (2004) used an experimental model based on students randomly selected from all applicants (by lottery) to attend three charter schools in Chicago. These students, particularly those entering in early grades, were found to outscore the control group (those in the lottery who were not selected) in math and reading. Based on these findings, the authors drew some general lessons about the importance of school type in effecting higher achievement.

However, in addition to their conflicting findings, these studies of charter schools exhibit some significant limitations that point to the need for further study. These reports often involve basic comparisons controlling for one variable at a time (e.g., student race/ethnicity), rather than statistical techniques that could control for multiple student- and school-level confounding variables simultaneously. Many of these studies focused largely or even exclusively on the number of students who attained proficiency levels, which, arguably, offer a rather arbitrary and less precise measure of school achievement than scale scores (Rothstein, 2004). The studies focused primarily on grade 4, raising the question of whether the results in other grades would be similar. Additionally, the NAEP studies compared student achievement means, rather than school means, raising the question of whether a small number of large, low-performing charter schools might have unfairly biased the results. Hoxby's national study comparing charter schools to neighboring schools was criticized for errors in matching schools which may have skewed the outcome (Carnoy *et al.*, 2005). Furthermore, that analysis assumes that charter schools draw from nearby public schools. While this may be true, there is reason to believe that charters also draw from a range of public schools and districts, private schools, and even other charter schools — suggesting the need for a larger scope of comparison (Hoxby & Rockoff, 2004; Lubienski, in press).

Research Question

Thus, despite the notable contributions of the research on charter schools, vouchers and public and private schools, there is good reason at this point for researchers to take another look at achievement across and within sectors, giving special attention to students in elementary and middle grades — where students are more likely to attend schools in the independent sectors (Broughman & Pugh, 2004; Miron & Nelson, 2002). Indeed, a new, nationally representative cross-sectional dataset has recently become available, and is larger than those previously brought to this question, including the substantial HSB and NELS datasets that served as the basis for earlier attempts to study this issue.

In the context of past academic debates and current policies pertaining to public and private forms of schooling, this paper reports on a study of mathematics achievement in public, charter, and different types of private schools across the country. This analysis addresses the following central question: ***After controlling for differences in demographics and location, how does achievement in public schools compare with that in charter, Catholic, and other types of private schools?***

In examining this issue, we focus on student performance in mathematics. This focus allows us to further isolate school effects from family background, since mathematics is a subject that is learned primarily in school relative to other subjects, such as reading, which tend to be more heavily influenced by students' experiences at home (Bryk *et al.*, 1993; Heyneman, 2005).

Data & Methods

The 2003 Main NAEP data were the source for this study (accessed from a restricted-access CD). A major strength of NAEP is that its mathematics assessment is based on a combination of multiple-choice, short-answer, and extended constructed response items from the five areas of mathematics recognized by the National Council of Teachers of Mathematics (1989, 2000): Number, geometry, measurement, algebra, and statistics/probability. Another important strength of NAEP is that it offers detailed student, teacher, and administrator survey data regarding a variety of student and school characteristics. Finally, compared with previous datasets that have been used to study school sector and student achievement, the 2003 NAEP dataset contains the

largest nationally representative sample of public and private schools, allowing a more detailed comparison of various school types.

However, it is important to note that NAEP data are cross-sectional, not longitudinal. Longitudinal data have strengths when it comes to making claims about school effectiveness, because initial achievement differences among students can be controlled. However, longitudinal datasets are not without difficulties. For example, the necessity to measure groups of students over time makes it difficult to maintain the analytical integrity of those groups for comparison when students change schools or disappear from the experiment.[7] Differential dropouts can bias a study's results, particularly when the dropout process is related to variables that are not observed. For instance, in the case of private schools, this can result in an even greater over-representation of students from families who are stable and/or highly committed to their children's schooling than what would occur in cross-sectional studies.

Yet, when using cross-sectional data, it is possible that there are differences in participants' initial achievement that are unknown. The use of statistical models is useful in such cases, as such models allow researchers to account appropriately for the primary potential student- and school-level variables that could shape initial achievement differences, making it less likely that longitudinal data would tell a different story.

Overall, due to the complexities of the issues involved, no single study can provide a definitive determination of the effectiveness of various forms of schooling. We view this study as providing up-to-date, nationally representative evidence that can inform current debates about the relationship between school sector and achievement.

NAEP Samples

Given the scope and richness of the 2003 NAEP data, and given that there are no other nationally representative datasets that provide the numbers of schools and the detailed achievement/survey information that NAEP offers, we chose to use NAEP for this study. In 2003, NAEP measured student performance in 4th and 8th grades only. Although NAEP's main focus is academic achievement, the students, their teachers, and school administrators complete detailed questionnaires pertaining to a variety of factors, including students' attitudes and

[7] For instance, Grogger and Neal (2000) report that missing data on dropouts represented a substantial problem for researchers using the NELS data.

classroom experiences, teachers' educational backgrounds, and school climate. The 2003 Main NAEP mathematics samples are much larger than previous NAEP samples, as they encompassed "State NAEP" samples for the first time. Specifically, the samples include 190,147 fourth graders (up from 13,855 in 2000), and 153,189 eighth graders (up from 15,930 in 2000), from representative samples of public and private schools (7485 schools at grade 4, and 6092 schools at grade 8).

Unweighted 4th- and 8th-grade samples sizes by school type are shown in Tables 1 and 2.

Table 1: Sample Sizes, Achievement, Student Demographics, and School Location by School Type, Full Grade 4 NAEP Reporting Sample (190,147 students from 7,485 schools)

	Public (non-charter)	*Catholic*	*Lutheran*	*Conservative Christian*	*Other Private*	*Charter*
	n=182,328 students 6,797 schools	n=2,285 students 216 schools	n=555 students 88 schools	n=651 students 78 schools	n=1,227 students 157 schools	n=3,101 students 149 schools
Student Factors (data analyzed at the student level)						
Mean student achievement	234	244	245	240	248	228
(standard error)	(.3)	(.9)	(1.5)	(1.5)	(1.8)	(2.3)
Pct black	17	7	12	12	10	30
Pct Hispanic	19	12	3	8	3	20
Pct American Indian	1	0	0	0	0	1
Pct Asian	4	3	1	6	6	2
Pct LEP	10	1	1	0	0	8
Pct with IEPs	11	3	2	4	3	8
Pct free/ reduced lunch	46	9	7	3	3	45
Home resources (of 6 items)	3.6	4.4	4.2	4.2	4.4	3.7
School Factors (data analyzed at the school level)						
Sch lunch (scale 1-6)^	4.1	2.1	1.9	1.6	1.6	4.2
Mean home resources	3.6	4.3	4.2	4.3	4.0	3.7
Pct minority	32	20	16	23	22	43
Pct LEP (scale 1-6)^	1.75	1.07	1.04	1.01	1.15	1.63
Pct large city	14	27	19	16	23	31
Pct rural/small town	39	16	22	26	17	23
Pct schools in Northeast	17	32	4	16	16	6
Pct schools in South	33	17	15	43	40	29
Pct schools in West	23	14	15	20	18	42
Pct schools in Midwest	27	37	66	21	26	23

^Scale was as follows: 1=0-5%, 2=6-10%, 3=11-25%, 4=26-50%, 5=51-75%, 6=76-100%

Note: Sample sizes are the unweighted NAEP reporting samples, however the means and percentages reported are for the samples that are weighted to represent U.S. students and schools.

Table 2: Sample Sizes, Achievement, Student Demographics, and School Location by School Type, Full Grade 8 NAEP Reporting Sample (153,189 students from 6092 schools)

	Public (non-charter)	*Catholic*	*Lutheran*	*Conservative Christian*	*Other Private*	*Charter*
	n=146,512 students 5,449 schools	n=2,463 students 224 schools	n=605 students 96 schools	n=659 students 90 schools	n=1,346 students 148 schools	n=1,604 students 85 schools
Student Factors (analyzed at the student level)						
Mean Student Achievement (standard error)	276 (.3)	289 (1.5)	296 (1.7)	286 (2.7)	298 (2.6)	271 (3.1)
Pct black	17	10	8	13	8	28
Pct Hispanic	15	12	5	4	4	19
Pct American Indian	1.2	.2	.5	.5	.5	2.6
Pct Asian	4	5	3	3	7	3
Pct LEP	5	.4	1	.0	.2	6
Pct with IEPs	11	3	2	2	3	11
Pct free/ reduced lunch	38	15	11	.2	14	50
Home resources (of 6 items)	4.3	5.0	5.0	4.8	5.1	4.3
School Factors (analyzed at the school level)						
Sch Lunch (scale 1-6)^	4.0	2.3	1.7	1.4	1.8	4.1
Mean Home resources	4.2	4.9	5.0	4.6	4.8	4.1
Pct Minority	29	22	12	23	21	40
Pct LEP (scale 1-6)^	1.4	1.1	1.0	1.1	1.1	1.4
Pct large city	12	29	26	17	15	24
Pct rural/small town	51	10	21	28	32	30
Pct schools in Northeast	17	30	3	10	19	7
Pct schools in South	32	18	15	43	44	28
Pct schools in West	21	16	17	23	14	50
Pct schools in Midwest	30	36	65	25	24	15

^Scale was as follows: 1=0-5%, 2=6-10%, 3=11-25%, 4=26-50%, 5=51-75%, 6=76-100%

Note: Sample sizes are the unweighted NAEP reporting samples, however the means and percentages reported are for the samples that are weighted to represent U.S. students and schools.

Although the full NAEP samples were used for descriptive comparisons, there were some missing data which reduced the samples used in the HLM analyses. Overall, the HLM samples contained 166,736 students across 6,664 schools at grade 4, and 131,497 students across 5,377

schools at grade 8 (see Table A1 in Appendix). Missing data impacted the samples for conservative Christian and other private schools (73%-79% schools included in the HLM analysis) more than the samples for the remaining school types (82%-91%). However, the demographics of the reduced HLM samples were remarkably similar to the demographics of the entire dataset, and in our varied analyses we saw no indication that missing data biased the study's results.[8]

Additionally, there were relatively low numbers of Lutheran and conservative Christian schools at both grades, as well as charter schools at grade 8. Also, the participation rates of conservative Christian schools did not meet NCES reporting standards in grade 4, and these rates were not met for charter schools in grade 8, nor for "other private schools" in grades 4 or 8.[9] Hence, readers are cautioned to view the results for these sub-samples as suggestive, rather than definitive, as to patterns that would hold for the full populations of these schools.

Variables Utilized in the HLM Analysis

NAEP's student- and school-level variables that have been found to correlate with mathematics achievement and that could confound the relationship between school sector and achievement were included in this analysis (Lubienski *et al.*, 2004). In selecting variables to include, the intent was to focus strictly on those that account for the differences in students/communities served by the school, as opposed to those factors that could be influenced by the school, such as the school discipline climate, teacher qualities, and even parent involvement. These other variables will be included in subsequent studies focusing on possible causes of the patterns identified in this paper.

[8] For example, we re-ran the HLM models with an expanded sample by deleting student lunch eligibility at the student and school level, as well as by deleting the school-level LEP indicator (given that these variables contributed the majority of missing data). The samples were then 185,681 students in 7,321 schools at grade 4, and 149,127 students across 5,965 schools at grade 8. The general finding that private school coefficients were statistically equal to or less than 0 after controlling for demographics and school location did not change.

[9] A 70% participation rate (before alternate schools are substituted) is required to meet NCES reporting standards. However, the use of NAEP's school weight helps diminish the bias resulting from school nonresponse, making this less of a concern. We decided to include all of the data in our analyses in order to provide as much information as possible on patterns in public, charter, and private school achievement. For example, given that prior studies gave much attention to the grade 4 charter school results, it seems important to note that the grade 8 results we obtained are different than those at grade 4, raising the question for further study as to whether this disparity is due to sample limitations at grade 8, or whether there might be an actual grade level difference in charter school achievement as compared with that in other public schools.

School Type.

Binary variables were used to distinguish among Catholic, Lutheran, conservative Christian, other private schools, and charter schools (which by default allowed comparisons with non-charter public schools).

Student Demographics.

Binary variables were used for Black, Hispanic, American Indian, Asian/Pacific Islander[10] and female students, as well as for students with "limited English proficiency" (LEP) and those with an "Individualized Education Plan" (IEP).[11]

A "Home Resources" composite was created by summing the number of the following items that students reported having in their home:[12]

- Magazines (received regularly)
- Newspaper (received at least four times weekly)
- Computer
- Encyclopedia (in book form or on computer)
- Atlas (a "book of maps" or on computer)
- Books (0-10 coded as 0, 11-25 books coded as .33, 26-100 books coded as .67, and more than 100 coded as 1)[13]

Survey information about individual students' eligibility for free or reduced lunch posed a challenge, given that a disproportionate number of private school administrators reported that the school did not participate in the program (the percentage was over half for some types of private schools). Although these schools generally appeared to be of high SES on other measures,[14] we did not assume that every student in such schools was ineligible for lunch. To preserve data

[10] We recognize the disagreements about the appropriate categories and terms to use when discussing racial and ethnic subgroups. To be consistent with the NAEP data, we use the designations employed by NAEP.

[11] There may be differences in IEP use in public and private schools. Hence, we ran models both with and without this variable. Although removing this variable slightly elevated the private school coefficients, it did not change the general finding that those coefficients were statistically equal to or less than 0 after controlling for demographics and school location.

[12] To preserve data, students who reported that they did not know if they had a particular resource were recoded to "no," with the logic that even if the resource was present in his/her home, it was not (directly) enriching the student's home experiences.

[13] These were the four categories used on the NAEP survey, and the intent was to simply convert this "books" variable to a 0-1 scale to match the remaining 5 home resources variables.

[14] For example, 95% of private school students in schools "not participating" in the lunch program had higher than average scores on the home resources composite

while also being cautious in imputing eligibility status, such students were recoded to "ineligible" only if their school administrator also reported (on a separate question) that less than 5% of the school was eligible for lunch, or if the student reported having at least 5 of the 6 resources that made up the home resources composite (5 was higher than the mean of 4.1 at grade 4, and 4.7 for grade 8, for *ineligible* students). Overall, the recoded students' mean achievement was higher than that of the other lunch ineligible students within their school type, providing further evidence that the recoding was warranted and did not negatively bias the performance of private schools. The final "lunch" variable was binary, with 0=ineligible, and 1=eligible for free or reduced lunch.[15]

School Demographics.

School administrators reported the percentage of students eligible for free or reduced lunch in their schools, as well as the percentage of LEP students, using the categories: 0%, 1-5%, 6-10%, 11-25%, 26-50%, 51-75%, 76-90%, and above 90%. The two top and bottom categories were collapsed to create a more evenly spaced 6-point scale. The percentage of students with IEPs was also included in initial models but was ultimately deleted due to lack of significance.

The home resources composite was aggregated to the school level to provide a school-level "mean home resources" measure. Additionally, school administrators reported the percentage of students who were Black, Hispanic and American Indian (on a 0%-100% scale), and these percentages were summed and then divided by 100 to create a "percent minority" variable for each school.[16]

School Location.

A binary variable denoted schools located within large cities, while another indicated if schools were in a rural or small-town locale. Similarly, binary variables were used to distinguish

[15] Another important socioeconomic indicator, parent education level, was in the 2003 NAEP data, but roughly one third of fourth graders and one fifth of 8th graders responded "I don't know" to the question. To limit missing data, we ultimately decided not to include parent education level. However, the HLM models were run at grade 8 with parent education included, and the results were very similar to the final models reported here. Additionally, NAEP includes student- and school-level Title 1 information, but inconsistencies in the ways schools determine students' Title 1 eligibility prompted us to exclude these variables from analyses.

[16] To preserve data, race/ethnicity information regarding sampled students from the school was used as a proxy when administrator-reported school information was missing. Also, the resulting variable was divided by 100 so that the HLM coefficients would be more accurate when rounded to 1 decimal place (e.g., 1.2 instead of 0.012).

among schools in the Northeastern, Southern, Western (and by default Midwestern) portions of the United States.

Data Analysis

Any statistical analysis of NAEP data is complicated by two methodological challenges: (1) NAEP's multi-stage cluster sampling design, and (2) measurement error associated with the matrix sampling scheme. These challenges and ways they were addressed are described briefly here (for more information, see Johnson, 1992; Johnson & Rust, 1992).

(1) Multi-stage cluster sampling. The 2003 NAEP national sample is a combination of data from pooled state samples along with a selection of additional schools to ensure that the sample represents the nation as a whole. In creating the samples, schools were stratified based on urbanicity, minority population, size, and area income, and then schools within each stratum were selected at random. Finally, students were selected randomly within schools. Deliberate oversampling of certain strata, such as private schools and those with high enrollments of minority students, resulted in more reliable estimates for the oversampled subgroups, but the resulting samples must be weighted to provide accurate estimates of the US school population. Weights are assigned to each student and school to adjust for both unequal probabilities of selection and nonresponse.

(2) Matrix sampling scheme. In order to reduce the test-taking burden on individual students, no student takes the entire NAEP battery of items, and individual students are not assigned a single "score." Instead, a distribution of plausible values for each student's proficiency is estimated, based on the student's responses to administered items and other student characteristics. Five "plausible values" are drawn at random from the conditional distribution of proficiency scores for each student. Hence, when analyzing NAEP achievement data, the results of separate analyses must be obtained for each of the five plausible values and then synthesized, following Rubin (1987) on the analysis of multiply-imputed data. This procedure accounts for the additional uncertainty arising from the use of five plausible values rather than a single observed outcome.

Throughout this project, the appropriate weights and statistical techniques were used to

address these special features of NAEP data.[17] Although most data management tasks were conducted within SPSS, most statistical analyses were conducted with the use of two software programs, AM and HLM, designed to address the needs of complex datasets such as NAEP. We also tested some of the mixed regression models using Stata 9.1 program xtmixed to confirm the results obtained within HLM; only minor differences were observed.

Specifically, student achievement means by school type were generated with the crosstabulation feature of AM Statistical Software, designed by the American Institutes for Research.[18] AM is designed to handle the weighting and jackknifing needs of NAEP, and was used to calculate basic achievement means and standard errors.[19]

However, the bulk of the analyses focused on the main research question about the relationship between school type and achievement. Because of the nested nature of the data (students within schools), we used HLM 6.0 to create 2-level hierarchical linear models to examine achievement by school type while controlling for potential student- and school-level confounding variables. A school-level weight was used at level 2; no level-1 weight was used because students were randomly selected within schools. The plausible values feature of HLM was used (prompting the program to run models for each of the five plausible values internally, and producing their average value and correct standard errors.) A detailed explanation of the data analysis methods used by the HLM software is available from Raudenbush and Bryk (2002).

A sequence of two-level HLM models were created at both 4th and 8th grades to examine the relationship between school type and mathematics achievement while controlling for demographic and other relevant variables. In all of these models, binary variables were entered uncentered, while continuous variables were entered "grand mean" centered, both at the student and school levels.

We began by running a traditional null model, followed by a model with school sector

[17] Given the many complexities of NAEP, when questions about variables or procedures arose, NAEP experts at NCES, ETS, and Westat were consulted.

[18] At annual NAEP training institutes, NCES experts have recommended the use of AM for running cross-tabulations of NAEP data. The standard errors obtained with AM were slightly more conservative than those on the NAEP Data Explorer (http://nces.ed.gov/nationsreportcard/nde). More information about AM is available at (http://am.air.org).

[19] Due to NAEP's cluster sampling, AM to calculates the sampling errors of data-based estimates using the jackknife repeated replication method.

variables only.[20] We then added student/school demographics to examine how much the demographics "explained" the school achievement disparities evident in the prior model. Finally, we added school location variables to examine the additional impact of rural/urban and US region on the school type coefficients.[21]

Given that previous studies indicated that some achievement inequities were smaller in Catholic schools than in public schools (e.g., Bryk *et al.*, 1993), interactions between school type and race- and SES-related achievement gaps were included in additional HLM models to determine if the gaps were larger or smaller in public, charter, Catholic, or other private schools.[22] No significant interactions were found. Hence, detailed results of this examination are not reported here. Further explorations of interactions, such as those involving school location in conjunction with minority status, will be considered further in follow-up studies.

Results

To help the reader interpret the results discussed here, some information about NAEP scores is necessary. NAEP mathematics results are reported with scale scores, with the 2003 mean scores being 235 at grade 4 and 278 at grade 8. The NAEP mathematics scale was originally designed to allow cross-grade comparisons, indicating that a 43-point difference between grade 4 and grade 8 would mean that a gap of 10 or 11 points represents a difference of roughly "one grade level." NAEP no longer maintains that the scale is consistent across grades, however the scales have not changed markedly. Therefore the idea that "10-11 points translates to approximately one grade level" is still often used as a helpful, albeit rough, guide for interpreting score differences. Another way to interpret the magnitude of these disparities is in terms of effect sizes. Standard deviations for school achievement means were 14.7 at grade 4 and 19.3 at

[20] We acknowledge that this "school sector only" model was rather unorthodox, containing no student-level variables. However, the inclusion of this model was important for allowing comparisons of achievement means by school type before and after demographic/location differences were controlled.

[21] We performed the usual regression diagnostics for the final 4th- and 8th-grade models, and no problems were identified. Specifically, we examined the means, standard deviations, and plots of fitted and residual values by school type, and no group appeared to be substantially better predicted than others. The means of residuals were close to 0 for each school type, and the standard deviations were similar.

[22] Only interactions for Black and Hispanic students were examined, given the sample size limitations for American Indian and Asian groups.

grade 8.[23] In these terms, a 10-11 point difference represents a moderate to large effect (roughly .5-.7 standard deviations).

Descriptive Comparisons

At both 4th and 8th grades, the mean mathematics achievement for charter school students tended to be lower than that of other public school students, and the mean among public school students was lower than the means for private school students (see Tables 1 and 2).[24] However, schools differed in their student demographics and location.

The percentage of minority students was highest in charter schools and lowest in private schools. The percentages of LEP, IEP, and low-SES (as measured by home resources and lunch eligibility) students were roughly similar in charter and other public schools, but were relatively low in private schools. Charter schools and some private schools were more likely to be located in larger cities, while public schools were disproportionately located in rural areas. Additionally, some school types were more concentrated in particular regions of the country. For example, conservative Christian schools were most likely found in the South, charter schools in the west and Lutheran schools in the Midwest.[25]

These patterns highlight the question of whether the relatively high student achievement in private schools and relatively low achievement in charter schools is due simply to differences in demographics or location, or if the achievement disparities would persist after controlling for such differences.

[23] These standard deviations were calculated by taking the square root of the variance between schools in the study's null HLM models (following the methods of Von Secker & Lissitz, 1999).

[24] In Tables 1 and 2, the unit of analysis for student-level data was individual students, while the unit of analysis for school-level data was schools. Hence, larger schools are weighted the same as smaller schools in school-level analyses, but are weighted more in student-level analyses because they have more students.

[25] Descriptive comparisons of school differences are not the primary focus of this article, and therefore statistical differences among the many measures in Tables 1 and 2 are not discussed here. Those interested in a more detailed descriptive comparison of NAEP data by private school type should consult a recent report published by NCES (Broughman & Pugh, 2004), which provides descriptive comparisons of mathematics, reading, science and writing data by school type. Interested readers can also create additional tables using the web-based NAEP Data Explorer (http://nces.ed.gov/nationsreportcard/nde).

Grade 4 HLM Models

At each grade level, four HLM models were used to examine achievement by school type. Raw, unweighted, descriptive statistics for the variables used in the HLM analyses are included in the appendix (Table A2).

Table 3 presents the grade 4 HLM results. Standard errors are included for the intercept and school type coefficients, given that these are of primary interest in this study. Model 1, the traditional HLM null model, indicates that the school mathematics mean for all of the schools averaged 235.7 points. It also reveals that 29% of the variance in achievement was between schools, and 71% of the variance involved students within schools. Model 2 indicates that the mean math achievement for non-charter public schools was 234.2 points. In comparison, without controlling for any demographic or other differences among schools, school achievement means were almost 10 points higher in Catholic schools, almost 11 points higher in Lutheran and "other private" schools, about 4 points higher in conservative Christian schools, and 6 points lower in charter schools.[26]

[26] Note that these are school-level means, not student-level.

Table 3: School Sector and Demographic Variables Predicting 4th Grade Mathematics Achievement (166,736 students, 6,664 schools)

	Model 1 Null Model	*Model 2 School Sector Only*	*Model 3 Sector + Demographics*	*Model 4 Sector + Demographics + Location*
Fixed Effects	*Coefficient (standard error)*	*Coefficient (standard error)*	*Coefficient (standard error)*	*Coefficient (standard error)*
Intercept – School Mean Achievement	235.7*** (0.3)	234.2*** (0.3)	247.0*** (0.2)	247.4*** (0.4)
School Level				
Catholic School		9.5*** (1.1)	-7.0*** (.9)	-7.2*** (.9)
Lutheran		10.7*** (1.9)	-4.1** (1.4)	-4.2** (1.4)
Cons. Christian^		4.2 (2.3)	-11.3*** (1.8)	-11.9*** (1.8)
Other Private^		11.0*** (2.1)	-4.6** (1.5)	-5.6*** (1.5)
Charter		-6.1** (2.3)	-4.7** (1.7)	-4.4** (1.6)
Pct Lunch			-2.0***	-2.1***
Mean Home Resources			5.2***	5.2***
Pct Minority			-1.1	-2.8**
Pct LEP			1.3***	1.7***
Large City				-0.2
Rural				-2.1***
Northeast				-0.6
South				2.7***
West				-2.4***
Student Level (slopes)				
Black			-15.6***	-15.8***
Hispanic			-6.6***	-6.5***
Am. Indian			-4.6***	-4.2***
Asian/White			6.4***	6.4***
Female			-4.4***	-4.4***
LEP			-11.3***	-11.3***
IEP			-20.9***	-20.9***
Lunch			-7.2***	-7.2***
Home Resources			2.4***	2.4***
Random Effects	*Variance Component*	*Variance Component*	*Variance Component*	*Variance Component*
Intercept (variance between schools)	215.2	201.0	55.4	51.5
Level-1 (variance within schools)	515.1	515.1	420.9	421.0
Intraclass Correlation	0.29	0.28	0.12	0.11

^Note: Participation rates did not meet NCES reporting requirements. Results for these subsamples should be interpreted with caution.

* p<.05 ** p<.01 *** p<.001

Model 3 adds student and school demographic controls, revealing that after adjusting for such differences among schools, the coefficients for charter schools and all private school types were significantly negative, ranging from -4.1 (Lutheran) to -11.3 (conservative Christian). To help the reader interpret the results, the intercept of 247 is the estimated mean achievement of a student who is 0 on all of the binary predictors and at the mean of all of the continuous predictors. As a specific example, 247 is the estimated mean achievement for White, non-IEP, non-LEP, lunch-ineligible males with average home resources, in a public school of average minority, LEP and SES populations. In a Catholic school of similar demographics, the estimated achievement of such a student would be 7 points lower, or 240. If the student was Black instead of White, the achievement would be an average of 15.6 points lower within the same school, regardless of type (given that no interaction effects were found).

Overall, Model 3 indicates that the private school advantage evident in Model 2 reverses after accounting for the higher proportions of advantaged students attending private schools. Given that 4th-grade charter schools had relatively high concentrations of minority students, their negative coefficient decreased in magnitude after controlling for these population differences, yet remained significantly negative.

Because the concentration of particular school types varied by US locale, we included location variables in our final model. Model 4 adds school location (urban and rural) and US region, yet the coefficients for the various school types are not markedly different than those in Model 3.[27] Specifically, the mean achievement was 247.4 for White, non-IEP, non-LEP lunch-ineligible males with average home resources, in a Midwestern public school of average minority, LEP and SES populations located in a small or medium-sized city or in the suburbs. The mean achievement in a school of similar demographics/location was an average of 7.2 points lower for Catholic schools, 4.2 points lower for Lutheran schools, 11.9 points lower for conservative Christian schools, 5.6 points lower for other private schools, and 4.4 points lower for charter schools.

[27] Despite the small changes in the coefficients and variance components, a multivariate hypothesis test revealed that the addition of location was significant ($p<.000001$), indicating that Model 4 is a better fit than Model 3. Similarly, although the change in variance was modest when moving from Model 1 to Model 2, the addition of school sector variables was also significant ($p<.000001$).

Model 4 also highlights some important inequities that persisted across all schools (again, race- and SES-related interactions with school type were examined but insignificant). Specifically, within schools, Black students scored an average of 15.8 points lower than their White peers of similar SES, LEP and disability status, while this gap was 6.5 points for Hispanic students and 4.2 points for American Indian students. Students of limited English proficiency scored 11.3 points lower than similar students within the same school who were not LEP, and this difference was 20.9 points for students with IEPs. Students who qualified for free/reduced lunch scored 7.2 points lower than their schoolmates who were ineligible (but demographically similar in every other way). For each of the 6 additional items students reported having at home, the mean achievement was 2.4 points higher, or nearly a 10-point gap between students with 6 items versus their demographically similar schoolmates who reported having only 2.

Also according to Model 4, in addition to student-level demographics correlating significantly with achievement, school SES and percentage of minority and LEP students were significantly correlated with achievement. Moreover, rural schools, as well as schools in the West, scored slightly but significantly (about 2-3 points) lower than Midwestern schools with equivalent demographics, while Southern schools scored significantly higher than others.

The proportion of variance between schools (intraclass correlation coefficient) was reduced from 29% in the null model to 11% in the final model. The final model explained 76% of the variance between schools and 18% of the variance within schools, largely due to the demographic controls, as opposed to school sector or location variables.

Grade 8 HLM Models

The same HLM models were created for 8th grade, as presented in Table 4. Model 1, the null model, shows that the overall mean school achievement was 279.7 points. It also reveals that 37% of the variance in achievement was between schools, and 63% of the variance involved students within schools.

Model 2 indicates that the 8th-grade school math achievement means for non-charter public schools averaged 275.5 points. In comparison, school achievement means were over 14 points higher in Catholic schools, 21 points higher in Lutheran, 14 points higher in "other private" schools, and more than 5 points higher in conservative Christian schools. Achievement in charter schools was roughly 1 point higher than in other public schools.

Model 3 shows that once student and school demographic differences are controlled, public school means are significantly higher than those of Catholic and conservative Christian schools, and statistically equal to means for the other schools.

As seen in Model 4, once school location is added, the results are similar to those of Model 3.[28] Overall, this full model reveals that, when compared with public schools, the mean mathematics achievement of schools with similar demographics/location was a statistically significant 3.8 points lower for Catholic schools, and a significant 10.6 points lower for conservative Christian schools. The remaining differences between public school means and those of others were not significant. It is worth noting that, unlike at grade 4, the coefficient for charter schools was a positive 2.4 points, but this was not statistically significant (p=.22).

As in grade 4, Model 4 also highlights important inequities that persist across all schools. Specifically, Black students scored an average of almost 20 points lower than their White schoolmates of similar SES, LEP and disability status within the same school. This gap was roughly 9 points for Hispanic students and 7.6 points for American Indian students. Students of limited English proficiency scored 15 points lower than similar students within the same school who were not LEP, and this difference was almost 35 points for students with IEPs. For each of the 6 additional items students reported having at home, the mean achievement was 4.5 points higher. With the exception of school lunch (for which the coefficient was roughly -7 at both grades), each of these demographic coefficients was larger at grade 8 than at grade 4.

Also according to Model 4, school-level SES measures were significant correlates of achievement. Additionally, the -7.5 coefficient for the percentage of minority students within a school was significant and greater in magnitude than the -2.8 coefficient at grade 4. Rural schools and those located in the West scored slightly but significantly lower than demographically similar schools in other locales.

The proportion of variance between schools was reduced from 37% in the null model to 18% in the final model. The final model explained 72% of the variance between schools and 24% of the variance within schools. As at grade 4, the majority of the variance explained was due to student- and school-level demographic factors.

[28] As in grade 4, a multivariate hypothesis test revealed that the addition of location was significant (p<.001), as was the addition of school sector variables when moving from Model 1 to Model 2 (p<.000001).

Table 4: School Sector and Demographic Variables Predicting 8th Grade Mathematics Achievement (131,497 students, 5,377 schools)

	Model 1 Null Model	*Model 2 School Sector Only*	*Model 3 Sector + Demographics*	Model 4 *Sector + Demographics + Location*
Fixed Effects	*Coefficient (standard error)*	*Coefficient (standard error)*	*Coefficient (standard error)*	*Coefficient (standard error)*
Intercept – School Mean Achievement	279.7*** (0.6)	275.5*** (.4)	290.6*** (.4)	291.9*** (.7)
School Level				
Catholic		14.3*** (1.5)	-3.0* (1.4)	-3.8** (1.4)
Lutheran		21.2*** (2.7)	1.8 (2.2)	1.0 (2.2)
Cons. Christian		5.4 (3.3)	-10.4*** (2.7)	-10.6*** (2.7)
Other Private^		14.3*** (3.7)	-2.0 (2.5)	-2.3 (2.6)
Charter^		.9 (4.0)	2.4 (2.0)	2.4 (2.0)
Pct Lunch			-1.9***	-1.8***
Mean Home Resources			4.5***	4.4***
Pct Minority			-5.9***	-7.5***
Pct LEP			0.8	1.0
Large City				1.3
Rural				-1.5*
Northeast				-.7
South				-.4
West				-2.2*
Student Level (slopes)				
Black			-19.6***	-19.7***
Hispanic			-9.3***	-9.3***
American Indian			-7.9***	-7.6***
Asian			7.2***	7.3***
Female			-4.0***	-4.0***
LEP			-15.1***	-15.1***
IEP			-34.7***	-34.7***
Lunch			-7.0***	-7.0***
Home Resources			4.5***	4.5***
Random Effects	*Variance Component*	*Variance Component*	*Variance Component*	*Variance Component*
Intercept (variance between schools)	372.6	325.8	104.4	103.1
Level-1 (variance within schools)	627.5	627.8	476.2	476.3
Intraclass Correlation	0.37	0.34	0.18	0.18

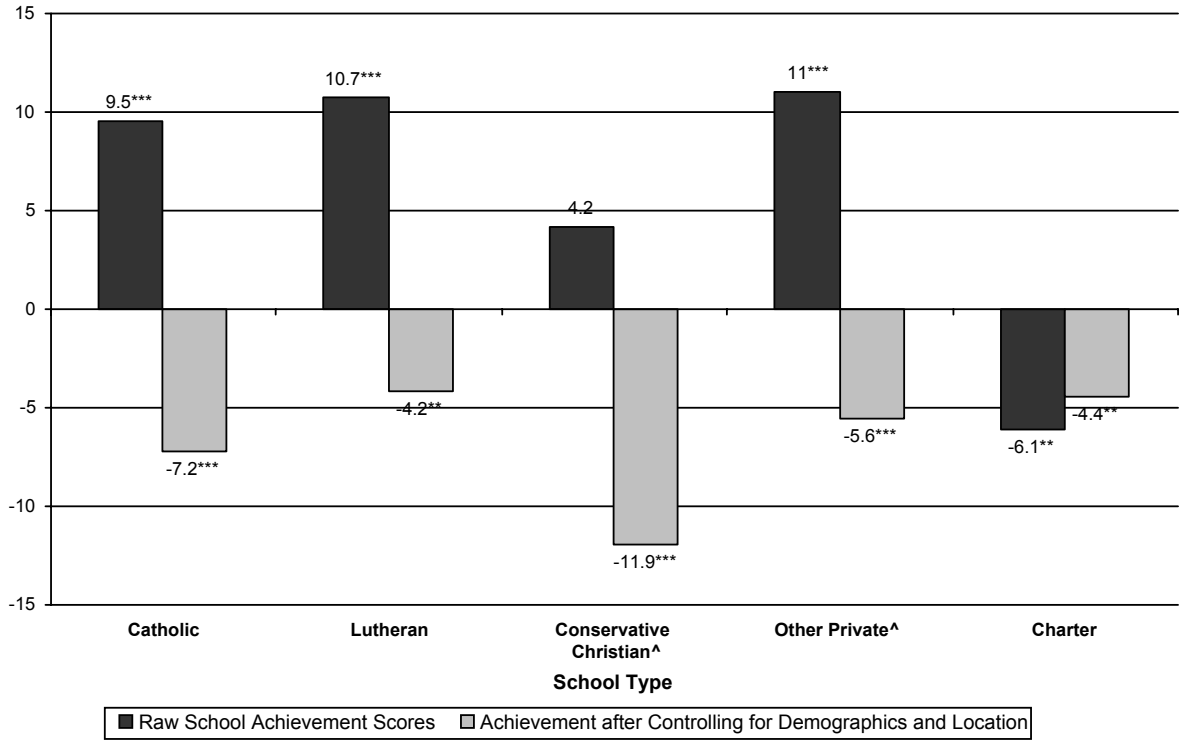
^Note: Participation rates did not meet NCES reporting requirements. Results for these subsamples should be interpreted with caution.

* p<.05 ** p<.01 *** p<.001

Summary of HLM Results

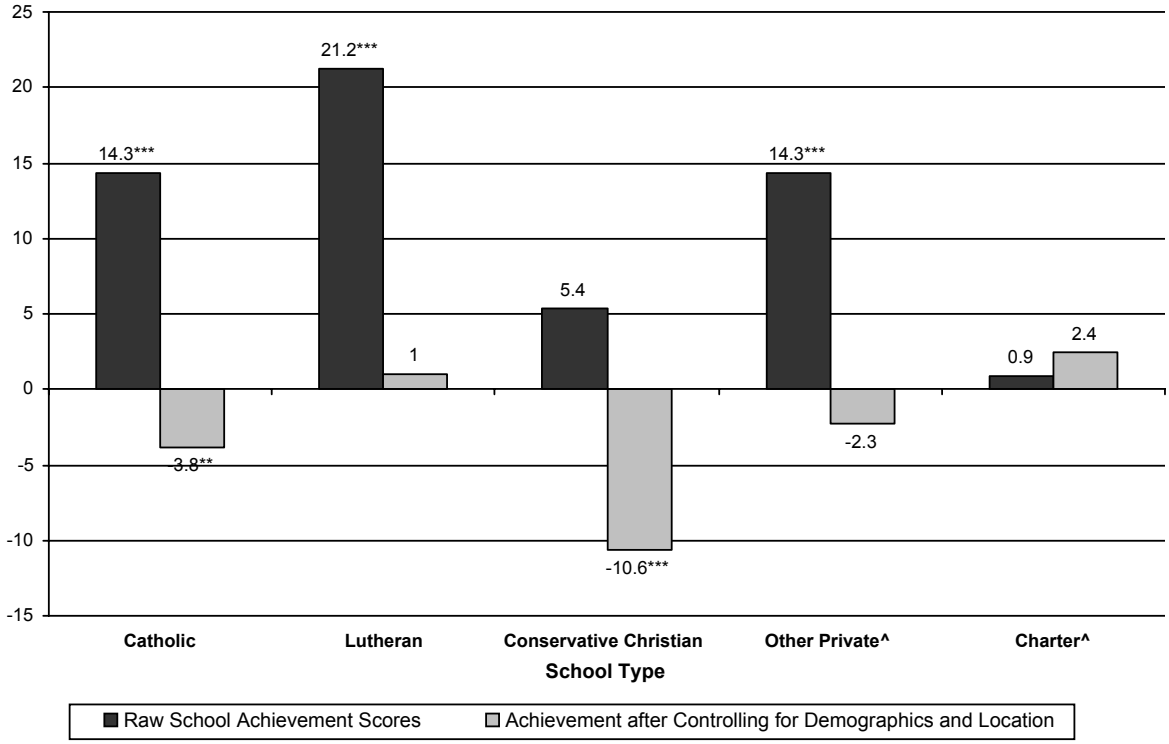
Figures 1 and 2 display the HLM results in a more visual way.

Figure 1: Predicted Differences Between Mean Achievement in Public Schools and Other School Types, Before and After Controlling for Demographics/Location, Grade 4



*NOTE: * p<.05 ** p<.01 *** p<.001 (Indicates means are significantly different than corresponding public school means.) ^ Participation rates did not meet NCES reporting requirements. Results for these subsamples should be interpreted with caution.*

Figure 2: Predicted Differences Between Mean Achievement in Public Schools and Other School Types, Before and After Controlling for Demographics/Location, Grade 8



*NOTE: * p<.05 ** p<.01 *** p<.001 (Indicates means are significantly different than corresponding public school means.) ^ Participation rates did not meet NCES reporting requirements. Results for these subsamples should be interpreted with caution.*

The bar graphs show the predicted differences in school achievement, before and after controlling for demographic and location differences among schools. The 0-axis represents public school mean achievement, and the bars show the HLM coefficients from Model 2 (school sector only) and Model 4 (with demographic and location controls), thereby displaying the predicted difference between the mean achievement of the five school types and those of public schools. In terms of ways to interpret the disparities between school achievement means, recall that one approach is to consider a 10-11 point disparity as very roughly representing a difference of one grade level. In that light, a difference of 5 or 10 points may seem substantial. In terms of "effect size," a 10-11 point difference in school means represents a moderate to large effect of roughly .5-.7 when compared with the standard deviations at grades 4 (14.7) and 8 (19.3).

Discussion

The results for fourth- and eighth-grade consistently indicated that demographic differences between public and private schools more than accounted for the relatively high raw NAEP mathematics scores of private schools. After controlling for demographic differences, no charter or private school means were higher than public school means to any statistically significant degree; moreover, particularly at grade 4, public schools actually scored significantly higher than did private and charter schools.

The analysis also points to some notable differences *within* school sectors. Charter schools scored significantly lower than non-charter public schools at grade 4, but scored slightly (statistically insignificantly) higher in grade 8 after accounting for differences in demographics and school location. Thus, in view of recent debates about the relative performance of charter schools (based on analyses of the 2003 NAEP data that were unable to control simultaneously for multiple variables), this more comprehensive analysis of the raw data indicates somewhat mixed results for charter schools. This analysis largely substantiates findings from the AFT (Nelson *et al.*, 2004) and the US Department of Education (2004) that, even after demographic differences are accounted for, public school achievement means are higher than charter school means at grade 4. However, similar results were not obtained at grade 8 (where the charter school sample was more limited and therefore not included in previous reports). Overall, this study suggests that charter schools are neither the unqualified failure that detractors claim, nor that there is something inherent in the independent structure of charter school organization that promotes greater student achievement, as choice enthusiasts would have us believe.

The data also highlight important differences within the private sector. Much of the research on private school effects has focused on the beneficial impact of Catholic schools, particularly for urban and disadvantaged minority students. Yet these results show that means for both public and Lutheran schools tended to be higher than for Catholic schools after accounting for differences in student populations. Additionally, as the Catholic school share of the private school sector has declined, conservative Christian schools have accelerated in growth (Broughman & Pugh, 2004). This is particularly interesting in view of our finding that conservative Christian schools were the lowest performing school type.

Our general findings regarding private and public school achievement, although notable, are not entirely without precedent. Despite common perceptions of a consensus around this issue, several studies have called into question claims of a general, positive private school effect in a

number of areas, including mathematics achievement (e.g., Alexander & Pallas, 1985; Figlio & Stone, 1997; Goldhaber, 1996; Krueger & Zhu, 2004a; Willms, 1985). It is also important to note that these present findings do not necessarily contradict much of the earlier work finding a private school effect in national-level datasets. This study examined achievement in 4th and 8th grade (and in fact, indicates less of a difference favoring public schools at grade 8 than at grade 4), while older HSB and NELS studies focused primarily on achievement in high school. Although these patterns between grades could be due to a number of explanations — teaching effectiveness, student attrition from or transfers between sectors, etc. — the current study does not necessarily refute, and may, in fact, be compatible with findings from earlier studies.

This nationally representative picture of achievement in US schools suggests that, in some ways, public schools may be doing a relatively good job compared to private schools once we account for differences in student populations. Still, these findings must be interpreted with care, due to limitations of this study.

Limitations

The size and representative nature of the NAEP samples and the general consistency and robustness of the results at grades 4 and 8 lend support for the reliability of this analysis. However, any analysis of NAEP has limits. The most important limitation of these data lies in the fact that NAEP data are cross-sectional, not longitudinal. Hence, NAEP data do not allow for examinations of individual student growth in achievement over time, nor do they include information about student movement between school sectors. Therefore, one cannot definitively conclude from this analysis that public schools are more effective at promoting student *growth* than other school types.

For example, some scholars might hypothesize that students who are lower achieving than others in their demographic peer group are more likely to enroll in private or charter schools. Given that the coefficients favoring public schools were less striking in grade 8 than in grade 4, such scholars might infer a long-term, positive effect of being in private or charter schools from these data. However, other observers might hypothesize that students enrolling in private or charter schools are likely to be educationally advantaged (by virtue of having parents who make an active investment in their schooling) in ways not accounted for in basic demographic controls. Those scholars might attribute the lessening of the coefficients in grade 8 to private school

students' home advantages having a greater effect in students' older years, when homework becomes a more integral part of learning, and students' inner beliefs/motivation could be a more important influence on achievement than particular school practices. Furthermore, attrition of low-performing private and charter school students in the later grades could occur because parents are no longer willing to make the sacrifice of tuition or transportation, and instead enroll these children in the local public schools. Clearly, longitudinal studies are needed to determine the reasons for the differences in the grade 4 and 8 results.

Another caution is that some missing data limited the HLM samples, particularly for conservative Christian and "other private" schools. The samples for some school types did not meet the stringent NCES reporting standards, namely conservative Christian and "other private" schools in grade 4, and charter and "other private" schools in grade 8. Hence, the results found for these groups at those grade-levels should be viewed as *suggestive* of relationships that may exist in the general population of all such schools, and indicate the need for further research to examine these relationships.

Additionally, national NAEP data tell us little about local differences among schools. For instance, it is becoming increasingly difficult to speak of "charter schools" because of the great variation between the authorizing legislation in different states. A wide variety of charter schools has emerged — some small-scale independent operations, others run by management companies, some with a particular social mission, others with a for-profit orientation. The present analysis treats charter schools as monolithic when they are not.

This analysis also raises, but does not answer the question of why there were such striking differences among private school types — for instance, how to account for the relatively strong performance by Lutheran schools. And yet these differences may be important considerations if we are ever to understand possible causes for variations in achievement in different sectors, and to identify potential processes, practices or attributes that can be replicated across sectors in order to boost student achievement.

Our ongoing analyses will examine some of the differences in school practices that could explain some of the achievement disparities by school type. Initial results from a follow-up study point to some intriguing differences. For example, the fraction of 4th and 8th graders with fully certified teachers averaged only about one half in conservative Christian schools, compared with over two thirds in Catholic and Lutheran schools, and over 80% in public schools.

Moreover, our initial results suggest that having a fully certified teacher correlates positively with achievement even after controlling for potential confounding variables. This follow-up study, as well as additional studies utilizing longitudinal data and encompassing other subject areas, can further our understanding of the relationship between school sector and student achievement.

Policy Implications

As a remarkably large, nationally representative dataset, NAEP offers important evidence regarding the state of US schools. The patterns in achievement reported from these data are significant for research and policy, especially in the context of recent efforts to reform the organization of schooling in the US. In particular, school choice and privatization movements have grown dramatically over the last generation. Two of the pillars of school choice — charter schools and voucher programs — are based specifically on the notion that private or independent schools engender a positive effect on student achievement, especially when compared to public schools. This is also reflected in the most sweeping federal reform of schooling in decades, the NCLB Act of 2001, which attempts to bring additional dimensions of accountability to public schools, using sanctions such as school choice and charter school status for public schools found to be failing at boosting achievement for different groups of students.

The gross inequities revealed in the final HLM models, such as the fact that Black eighth graders scored an average of almost 20 points (roughly 2 grade levels) lower than White students within the same school who were identical on all other demographic measures, suggests that the goals of "leaving no child behind," and monitoring and reducing achievement gaps within schools, are critically important. However, despite reformers' focus on school organizational type, school sector variables mattered relatively little in the HLM models when compared with the impact of demographic measures, and there was far more variation within schools than between them. This suggests that reform efforts targeting inequities among students could have more potential for closing achievement gaps and raising US test scores than reforms promoting private forms of schooling (S. Lubienski, 2003; Rothstein, 2004).

Indeed, the findings from this study, which suggest that higher private school mathematics achievement is more than accounted for when demographic differences are controlled, call into question a basic premise of such reforms. There are many reasons one could support school

choice, but evidence of inherently higher student achievement in private and independent schools may not be among them.

This issue is underscored in these data by the relative performance of different types of schools. In the economic logic underpinning market theories of school choice, liberating consumers to choose will allow families trapped in poorly performing schools to escape to more effective ones (Walberg & Bast, 2003). In this perspective, by positioning parents as the driving force in the quest for quality, schools will be forced to improve when faced with competition from higher performing rivals. But this logic assumes that parents will indeed choose schools of higher academic quality for their children. These data question that assumption. Most markedly, achievement in conservative Christian schools was below that of all other types of schools in this analysis, and significantly under that of tuition-free public schools (after controlling for demographics). Yet conservative Christian schools are the fastest growing segment of the private school sector, with growth outdistancing that of higher achieving private schools (Broughman & Pugh, 2004). Of course, there are many other reasons that parents may choose conservative Christian schools — from religious-based curricula to proximity. However, inasmuch as that is true it undercuts the argument that parents will choose schools primarily based on academic quality, and calls into question the idea of using parents as the driving force for general school improvements through competitive dynamics.

Overall, the results of this study suggest that, despite the many difficulties faced by public schools, they appear to be performing relatively well when compared to demographically similar private and charter schools, without the remedy of major, private-style structural reforms in their governance and management. These findings question the idea of an inherent superiority of the private sector in education. Furthermore, the data here suggest significant reasons to be suspicious of claims of general failure in the public schools, and raise substantial questions regarding a basic premise of the current generation of school reform.

REFERENCES

Alexander, K. L., & Pallas, A. M. (1985). School Sector and Cognitive Performance: When Is a Little a Little? *Sociology of Education, 58*(2), 115-128.

Bridge, R. G., & Blackman, J. (1978). *A Study of Alternatives in American Education: Family Choice in Schooling* (Vol. IV). Santa Monica, CA: Rand.

Broughman, S. P., & Pugh, K. W. (2004). *Characteristics of Private Schools in the United States: Results from the 2001-2002 Private School Universe Survey* (No. NCES 2005-305). Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Bryk, A. S. (1981). Disciplined Inquiry or Policy Argument? *Harvard Educational Review, 51*(4), 497-510.

Bryk, A. S., Holland, P. B., Lee, V. E., & Carriedo, R. A. (1984). *Effective Catholic Schools: An Exploration*. Washington, DC: National Center for Research in Total Catholic Education.

Bryk, A. S., Lee, V. E., & Holland, P. B. (1993). *Catholic Schools and the Common Good*. Cambridge, MA: Harvard University Press.

Carnoy, M., Jacobsen, R., Mishel, L., & Rothstein, R. (2005). *The Charter School Dust-Up: Examining the Evidence on Enrollment and Achievement*. Washington, DC: Economic Policy Institute.

Catterall, J. S., & Levin, H. M. (1982). Public and Private Schools: Evidence on Tuition Tax Credits. *Sociology of Education, 55*(2/3), 144-151.

Chubb, J. E., & Moe, T. M. (1990). *Politics, Markets, and America's Schools*. Washington, D.C.: Brookings Institution.

Coleman, J. S. (1997). The Design of Schools as Output-Driven Organizations. In R. Shapira & P. W. Cookson (Eds.), *Autonomy and Choice in Context: An International Perspective* (pp. 249-270). Oxford, UK: Pergamon.

Coleman, J. S., & Hoffer, T. (1987). *Public and Private High Schools: The Impact of Communities*. New York: Basic Books.

Coleman, J. S., Hoffer, T., & Kilgore, S. (1981). *Public and Private Schools: A Report to the National Center for Educational Statistics*. Chicago: National Opinion Research Center.

Coleman, J. S., Hoffer, T., & Kilgore, S. (1982). *High School Achievement: Public, Catholic, and Private Schools Compared*. New York: Basic Books.

Figlio, D. N., & Stone, J. A. (1997). *School Choice and School Performance: Are Private Schools Really Better?* (Discussion Paper No. 1141-97). Madison, WI: Institute for Research on Poverty.

Friedman, M. (1955). The Role of Government in Education. In R. A. Solo (Ed.), *Economics and the Public Interest* (pp. 127-134). New Brunswick, NJ: Rutgers University Press.

Friedman, M. (1962). *Capitalism and Freedom*. Chicago: University of Chicago Press.

Gamoran, A. (1996). Student Achievement in Public Magnet, Public Comprehensive, and Private City High Schools. *Educational Evaluation and Policy Analysis, 18*(1), 1-18.

General Accounting Office. (2002). *School Vouchers: Characteristics of Privately Funded Programs* (Report to the Honorable Judd Gregg, U.S. Senate No. GAO-02-752). Washington, DC: General Accounting Office.

Goldberger, A. S., & Cain, G. G. (1982). The Causal Analysis of Cognitive Outcomes in the Coleman, Hoffer and Kilgore Report. *Sociology of Education, 55*(2/3), 103-122.

Goldhaber, D. D. (1996). Public and Private High Schools: Is School Choice an Answer to the Productivity Problem? *Economics of Education Review, 15*(2), 93-109.

Greeley, A. M. (2002). *Catholic High Schools and Minority Students*. New Brunswick, NJ: Transaction Publishers.

Greene, J. P., Peterson, P., & Du, J. (1996). *The Effectiveness of School Choice in Milwaukee: A Secondary Analysis of Data from the Program's Evaluation* (Research Papers Series). Harvard Program on Education Policy and Governance.

Greene, J. P., Howell, W. G., & Peterson, P. E. (1998a). Lessons from the Cleveland Scholarship Program. In P. E. Peterson & B. C. Hassel (Eds.), *Learning from School Choice* (pp. 357-392). Washington, DC: Brookings Institution Press.

Greene, J. P., Peterson, P. E., & Du, J. (1998b). School Choice in Milwaukee: A Randomized Experiment. In P. E. Peterson & B. C. Hassel (Eds.), *Learning from School Choice* (pp. 335-356). Washington, DC: Brookings Institution Press.

Grogger, J., & Neal, D. (2000). Further Evidence on the Effects of Catholic Secondary Schooling. In W. G. Gale & J. R. Pack (Eds.), *Brookings-Wharton Papers on Urban Affairs, 2000* (pp. 151-202). Washington, DC: Brookings Institution Press.

Heyneman, S. P. (2005). Student Background and Student Achievement: What Is the Right Question? *American Journal of Education, 112*(1), 1-9.

Howell, W. G., & Peterson, P. E. (2004). Uses of Theory in Randomized Field Trials: Lessons from School Voucher Research on Disaggregation, Missing Data, and the Generalization of Findings. *American Behavioral Scientist, 47*(5), 634-657.

Howell, W. G., Wolf, P. J., Peterson, P. E., & Campbell, D. E. (2000). *Test-Score Effects of School Vouchers in Dayton, Ohio, New York City, and Washington, D.C.: Evidence from Randomized Field Trials* (Research Papers Series). Harvard Program on Education Policy and Governance.

Hoxby, C. M. (2004a). *Achievement in Charter Schools and Regular Public Schools in the United States: Understanding the Differences*. Cambridge, MA: Department of Economics, Harvard University.

Hoxby, C. M. (2004b). *A Straightforward Comparison of Charter Schools and Regular Public Schools in the United States*. Cambridge, MA: Department of Economics, Harvard University.

Hoxby, C. M., & Rockoff, J. (2004). *The Impact of Charter Schools on Student Achievement: A Study of Students Who Attend Schools Chartered by the Chicago Charter School Foundation*. Cambridge, MA: Department of Economics, Harvard University.

Johnson, E. G. (1992). The Design of the National Assessment of Educational Progress. *Journal of Educational Measurement, 29*(2), 95-110.

Johnson, E. G., & Rust, K. F. (1992). Population Inferences and Variance Estimation for NAEP Data. *Journal of Educational Statistics, 17*(2), 175-190.

Kim, M. M., & Placier, M. (2004). Comparison of Academic Development in Catholic Versus Non-Catholic Private Secondary Schools. *Education Policy Analysis Archives, 12*(5), Retrieved on June 24, 2004 from http://epaa.asu.edu/epaa/v2012n2005/.

Krueger, A. B., & Zhu, P. (2004a). Another Look at the New York City School Voucher Experiment. *American Behavioral Scientist, 47*(5), 658-698.

Krueger, A. B., & Zhu, P. (2004b). Inefficiency, Subsample Selection Bias, and Nonrobustness: A Response to Paul E. Peterson and William G. Howell. *American Behavioral Scientist, 47*(5), 718-728.

Lee, V. E., & Bryk, A. S. (1989). A Multilevel Model of the Social Distribution of High School Achievement. *Sociology of Education, 62*(3), 172-192.

Lee, V. E., & Burkam, D. T. (2003). Dropping Out of High School: The Role of School Organization and Structure. *American Educational Research Journal, 40*(2), 393-394.

Lee, V. E., Chow-Hoy, T. K., Burkam, D. T., Geverdt, D., & Smerdon, B. A. (1998). Sector Differences in High School Course Taking: A Private School or Catholic School Effect? *Sociology of Education, 71*(4), 314-335.

Lines, P. M. (1993). *Parents and Their Informational Resources: A Reassessment of Findings from Alum Rock*. Washington, DC: U.S. Dept. of Education, Office of Educational Research and Improvement, Educational Resources Information Center.

Lubienski, C. (2001). Redefining "Public" Education: Charter Schools, Common Schools, and the Rhetoric of Reform. *Teachers College Record, 103*(4), 634-666.

Lubienski, C. (2003a). A Critical View of Home Education. *Evaluation and Research in Education, 17*(2/3), 167-178.

Lubienski, C. (2003b). Innovation in Education Markets: Theory and Evidence on the Impact of Competition and Choice in Charter Schools. *American Educational Research Journal, 40*(2), 395-443.

Lubienski, C. (2003c). Instrumentalist Perspectives on the "Public" in Public Education: Incentives and Purposes. *Educational Policy, 17*(4), 478-502.

Lubienski, C. (2005). Public Schools in Marketized Environments: Shifting Incentives and Unintended Consequences of Competition-Based Educational Reforms. *American Journal of Education, 111*(4), 464-486.

Lubienski, C. (in press). Incentives for School Diversification: Competition and Promotional Patterns in Local Education Markets. *Journal of School Choice, 1*(2).

Lubienski, S. T. (2003). Celebrating Diversity and Denying Disparities: A Critical Assessment. *Educational Researcher, 32*(8), 30-38.

Lubienski, S. T., Camburn, E., & Shelley, M. C. (2004). *Reform-Oriented Mathematics Instruction, Achievement, and Equity: Examinations Of Race and SES in 2000 Main NAEP Data*. Report to the National Center for Education Statistics.

Lubienski, S. T., & Lubienski, C. (2005). A New Look at Public and Private Schools: Student Background and Mathematics Achievement. *Phi Delta Kappan, 86* (9) 696-699.

Mayer, D. P., Peterson, P. E., Myers, D. E., Tuttle, C. C., & William, H. G. (2002). *School Choice in New York City after Three Years: An Evaluation of the School Choice Scholarships Program* (Final Report). Princeton, NJ: Mathematica Policy Research Inc.

Metcalf, K., West, S. D., Legan, N. A., Paul, K. M., & Boone, W. J. (2003). *Evaluation of the Cleveland Scholarship and Tutoring Program: Summary Report 1998-2002*. Bloomington, IN: Indiana University.

Miron, G., & Nelson, C. (2002). *What's Public About Charter Schools? Lessons Learned About Choice and Accountability*. Thousand Oaks, CA: Corwin Press.

Myers, D. E., Peterson, P., Mayer, D., Chou, J., & Howell, W. G. (2000). *School Choice in New York City after Two Years: An Evaluation of the School Choice Scholarships Program* (No. 8404-036): Mathematica Policy Research Inc.

National Council of Teachers of Mathematics. (1989). *Curriculum and Evaluation Standards for School Mathematics*. Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (2000). *Principles and Standards for School Mathematics*. Reston, VA: NCTM.

Nelson, F. H., Rosenberg, B., & Van Meter, N. (2004). *Charter School Achievement on the 2003 National Assessment of Educational Progress*. Washington, DC: American Federation of Teachers.

Perie, M., Vanneman, A., & Goldstein, A. (2005). *Student Achievement in Private Schools: Results from NAEP 2000–2005* (NCES No. 2006-459). Washington, DC: National Center for Education Statistics.

Peterson, P. E., & Howell, W. G. (2003). *Efficiency, Bias, and Classification Schemes: Estimating Private-School Impacts on Test Scores in the New York City Voucher Experiment*. (Research Papers Series). Harvard Program on Education Policy and Governance.

Peterson, P. E., & Howell, W. G. (2004). Efficiency, Bias, and Classification Schemes: A Response to Alan B. Krueger and Pei Zhu. *American Behavioral Scientist, 47*(5), 699-717.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods* (2nd ed.). Thousand Oaks: Sage Publications.

Rothstein, R. (2004). *Class and Schools: Using Social, Economic, and Educational Reform to Close the Black–White Achievement Gap*. Washington, DC: Economic Policy Institute.

Rouse, C. (1998a). *Schools and Student Achievement: More Evidence from the Milwaukee Parental Choice Program*. Princeton, NJ: Industrial Relations Section, Princeton University.

Rouse, C. E. (1998b). Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program. *Quarterly Journal of Economics, 113*(2), 553-603.

Schemo, D. J. (2004, August 17). Charter Schools Trail in Results, U.S. Data Reveals. *New York Times,* p. A1.

Scott, L. A., Rock, D. A., Pollack, J. M., & Ingels, S. J. (1995). *Two Years Later: Cognitive Gains and School Transitions of NELS:88 Eight Graders* (Statistical Analysis Report No. 95-436). Washington, DC: National Center for Education Statistics.

U.S. Department of Education. (2004). *The Nation's Report Card: America's Charter Schools* (No. NCES 2005-456). Washington, DC: National Assessment of Educational Progress.

Von Secker, C. E., & Lissitz, R. W. (1999). Estimating the Impact of Instructional Practices on Student Achievement in Science. *Journal of Research in Science Teaching, 36*(10), 1110-1126.

Walberg, H. J., & Bast, J. L. (2003). *Education and Capitalism: How Overcoming Our Fear of Markets and Economics Can Improve America's Schools*. Stanford, CA: Hoover Institution Press.

Willms, J. D. (1985). Catholic-School Effects on Academic Achievement: New Evidence from the High School and Beyond Follow-up Study. *Sociology of Education, 58*(2), 98-114.

Witte, J. F. (2000). *The Market Approach to Education: An Analysis of America's First Voucher Program*. Princeton, NJ: Princeton University Press.

APPENDIX

Table A1: Comparing Sample Sizes, Student Demographics, School Achievement, and School Location, Full NAEP Sample and HLM Sample

	Grade 4		*Grade 8*	
	Full Sample	*HLM Sample (% of full sample)*	*Full Sample*	*HLM Sample (% of full sample)*
Total Number of Students	190147	166736 (88%)	153189	131497 (86%)
Percent Female	49	49	50	50
Percent Black Students	18	18	17	17
Percent Hispanic Students	13	13	11	11
Pct American Indian	2	2	2	2
Pct Asian	4	4	4	4
Percent White	62	63	66	66
Pct LEP	7	7	4	4
Pct with IEPs	12	11	11	11
Pct free/reduced lunch	42	45	36	37
Home resources (out of 6 items)	3.7	3.7	4.3	4.3
Total Number of Schools	7485	6664 (89%)	6092	5377 (88%)
# Catholic Schools	216	180 (83%)	224	203 (91%)
# Lutheran Schools	88	73 (83%)	96	79 (82%)
# Cons. Chr. Schools	78	57 (73%)	90	70 (78%)
# Other Private Schools	157	124 (79%)	148	114 (77%)
# Charter Schools	149	126 (85%)	85	72 (85%)
# Public Schools	6797	6104 (90%)	5449	4839 (89%)
Mean School Achievement	235.3	235.7	279.3	279.7
Sch Lunch (scale 1-6)^	4.0	4.0	3.8	3.8
Mean Home Resources	3.7	3.7	4.4	4.4
Pct Minority	30	29	27	27
LEP (scale 1-6)^	1.6	1.6	1.4	1.4
Percent Large City School	17	18	15	15
Percent Rural Schools	41	41	43	43
Percent Schools in Northeast	18	18	17	17
Percent Schools in South	31	31	31	32
Percent Schools in West	26	24	24	23
Percent Schools in Midwest	25	27	28	28

^Scale was as follows: 1=0-5%, 2=6-10%, 3=11-25%, 4=26-50%, 5=51-75%, 6=76-100%

Note: The mean school achievement is the intercept of the base model in each case (i.e., weighted school means). The other percentages are the raw, unweighted descriptive statistics for each set of data.

Table A2: Raw, Unweighted Descriptive Statistics for Variables Used in the HLM Analysis

	Grade 4 *166,736 students, 6664 schools*				*Grade 8* *131,497 students, 5377 schools*			
	MIN	MAX	MEAN	SD	MIN	MAX	MEAN	SD
Student Level								
Plausible Value 1	97.53	332.82	233.62	28.31	114.17	415.14	276.68	35.70
Plausible Value 2	109.04	339.29	233.65	28.29	109.76	414.25	276.72	35.66
Plausible Value 3	99.31	337.54	233.64	28.25	97.43	400.27	276.74	35.73
Plausible Value 4	98.53	336.78	233.71	28.23	110.77	400.08	276.69	35.72
Plausible Value 5	106.96	341.66	233.67	28.23	107.41	411.10	276.66	35.67
Black	0.00	1.00	0.18	0.39	0.00	1.00	0.17	0.37
Hispanic	0.00	1.00	0.13	0.33	0.00	1.00	0.11	0.31
Indian	0.00	1.00	0.02	0.13	0.00	1.00	0.02	0.13
Asian	0.00	1.00	0.04	0.20	0.00	1.00	0.04	0.20
Female	0.00	1.00	0.49	0.50	0.00	1.00	0.50	0.50
LEP	0.00	1.00	0.07	0.25	0.00	1.00	0.04	0.19
IEP	0.00	1.00	0.11	0.32	0.00	1.00	0.11	0.31
Lunch	0.00	1.00	0.45	0.50	0.00	1.00	0.37	0.48
Home Resources	0.00	6.00	3.69	1.48	0.00	6.00	4.33	1.42
School Level								
Catholic	0.00	1.00	0.03	0.16	0.00	1.00	0.04	0.19
Lutheran	0.00	1.00	0.01	0.10	0.00	1.00	0.01	0.12
Cons. Christian	0.00	1.00	0.01	0.09	0.00	1.00	0.01	0.11
Other Private	0.00	1.00	0.02	0.14	0.00	1.00	0.02	0.14
Charter	0.00	1.00	0.02	0.14	0.00	1.00	0.01	0.12
Pct Lunch	1.00	6.00	4.02	1.55	1.00	6.00	3.76	1.53
Home Resources	0.00	6.00	3.72	0.66	0.00	6.00	4.35	0.64
Pct Minority	0.00	1.00	0.29	0.33	0.00	1.00	0.27	0.31
Pct LEP	1.00	6.00	1.61	1.20	1.00	6.00	1.39	0.92
Large City	0.00	1.00	0.18	0.38	0.00	1.00	0.15	0.36
Rural	0.00	1.00	0.41	0.49	0.00	1.00	0.43	0.50
Northeast	0.00	1.00	0.18	0.38	0.00	1.00	0.17	0.38
South	0.00	1.00	0.31	0.46	0.00	1.00	0.32	0.47
West	0.00	1.00	0.24	0.43	0.00	1.00	0.23	0.42