EFFECTS OF CATHOLIC SCHOOL ATTENDANCE ON STUDENT ACHIEVEMENT: A REVIEW AND EXTENSION OF RESEARCH

THOMAS B. HOFFER
National Opinion Research Center
at the University of Chicago

This paper begins with a review of basic descriptive data on achievement differences between public and Catholic school students and the main theories intended to explain Catholic sector effects on student achievement. The main theories are cast in terms of competitive markets, the communities in which the schools are embedded, and the historically institutionalized purposes of the schools. The analytical research is then reviewed and extended with some original analyses from recently collected national survey data on high school students. The main points from the review and extension of empirical research are as follows: (a) Catholic high schools have positive effects on verbal and mathematics achievement, but no discernable effects on science; (b) Catholic school effects are greater for students from disadvantaged backgrounds, especially with respect to family structure and functioning; and (c) the main schooling mechanism accounting for the Catholic school effects is the greater concentration of academic coursetaking among Catholic school students. The most glaring gap in the research record is the lack of data to assess effects of Catholic elementary school attendance. Finally, further work is needed to sort out the larger theoretical issues and practical implications of markets, charters, and communities.

Schools are intended and expected to have many effects on their students, ranging from the inculcation of relatively diffuse attitudes and norms to quite specific cognitive skills. How well schools achieve their various goals is currently a matter of great public concern, but the concern is perhaps greatest with respect to the cognitive goals of schooling. The symbolic
watershed event in the current era of reform is the report *A Nation at Risk* (National Commission on Excellence in Education, 1983). The report lamented a "rising tide of mediocrity" and sounded alarms about declines in national security and the standard of living to galvanize support for raising performance standards and outcomes.

The main themes of the report were based on evidence of college entrance examination scores that had been declining since the mid-1970s. Perhaps responding to that growing evidence, the U.S. government began to collect comparable data on the academic achievement of public and Catholic school students in the early 1980s. Researchers interested in finding evidence of effective alternatives to the practices—and, in some cases, to the very existence—of the public schools eagerly turned to these new databases for clues. Much of the large volume of research comparing the two sectors reflects an interest in identifying ways to improve American education generally, both public and private.

This report summarizes and critiques research comparing public and Catholic school student achievement scores in verbal skills, mathematics, science, and social studies. Findings from three national studies are reviewed: the National Assessment of Educational Progress (NAEP), High School and Beyond (HS&B), and the National Education Longitudinal Study of 1988 (NELS:88). The report is divided into three main sections: a review of theoretical arguments about the nature and size of sector effects, a review of the empirical results from the national surveys, and a conclusion discussing the strengths and weaknesses of the completed research and how future work might improve upon it.

**SECTOR DIFFERENCES AND THE THEORETICAL ARGUMENTS FOR SECTOR EFFECTS**

Theory and data are always closely bound, and the more exacting the connections, the better the theory one formulates and the better the data one collects. Where one steps into the circle of evidence and theory inevitably has at least an element of the arbitrary. For present purposes, I take as a point of departure a range of empirical evidence from national surveys showing that Catholic school students score significantly higher on standardized achievement tests than their public school counterparts. These data do not, however, necessarily indicate that Catholic schools are more effective than public schools, for that is an issue that must be resolved through analysis of the overall sector differences. The first theoretical issue is how to account for the overall achievement differences. Before considering the alternative explanations, it is useful to review the basic evidence.
OVERVIEW OF SECTOR DIFFERENCES IN ACHIEVEMENT

The National Assessment of Educational Progress (NAEP) data are the best source for national estimates of public and Catholic school student achievement differences and for historical trends in those differences. This is because NAEP uses large samples of schools and students and administers a much larger set of items in each subject area than other studies. Furthermore, NAEP includes national samples of students at three age and grade levels (ages 9, 13, and 17; grades 4, 8, and 12), whereas the national longitudinal studies start with eighth graders (NELS:88) or tenth graders (HS&B). The NAEP is designed especially for estimating average scores and charting historical trends in those averages for the three age-grade populations and major subpopulations within them; it is relatively useless for analyzing the scores in terms of factors predicting average differences. For that purpose, one must turn to the national longitudinal studies, HS&B and NELS:88.

Results are reported here for only mathematics, reading, and writing, despite the fact that NAEP also assesses achievement in science, history, and geography. Public-Catholic school breakdowns were not available at the time of writing for the latter subject areas, though the U.S. Department of Education indicates (personal communication) that those data will be available in late 1997.

The NAEP data presented in Table 1 show that, on average, Catholic school students in grades 4, 8, and 12 perform at levels significantly above their public school counterparts in the areas of mathematics, reading comprehension, and writing skills. Though the differences are slightly smaller in grade 4, they are, overall, very large in all cases. The NAEP scores are reported on a scale that ranges from 0 to 500. This scale is itself completely arbitrary, and one must define some standard against which to interpret any differences of interest. One minimal standard is statistical significance, that is, the likelihood that a given difference could occur by chance, through "the luck of the draw" (sampling error) in a sample survey. The differences in Table 1 are all sufficiently large that one can dismiss chance as an explanation: even the smallest of these differences is less likely than one-in-one-thousand to arise from sampling error.

The main standard that NAEP provides in its reports is a set of five "performance levels" that are defined for discrete 50-point ranges of the 0-500 scale scores. The performance levels are defined, first, by identifying the specific test items that students achieving at each level are on average likely to answer correctly and incorrectly. Subject-matter specialists then scrutinized these item sets to identify the kinds of skills and knowledge needed to answer the benchmark items correctly. NAEP reports these scores as percentages of students achieving at each level of performance. These scores are attractive
# Table 1.

**Average NAEP Test Scores for Public and Catholic School Students in the 1992 Assessments**

<table>
<thead>
<tr>
<th>Test, Grade Level, and Full Sample Standard Deviations</th>
<th>Public</th>
<th>Catholic</th>
<th>Difference</th>
<th>Effect Size (unadjusted)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Standard Deviation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Grade 4 (std. dev.=33.1)</td>
<td>217.3 (0.8)</td>
<td>226.6 (1.2)</td>
<td>9.3 (1.4)</td>
</tr>
<tr>
<td></td>
<td>Grade 8 (std. dev.=30.9)</td>
<td>266.1 (1.0)</td>
<td>277.2 (2.1)</td>
<td>11.1 (2.3)</td>
</tr>
<tr>
<td></td>
<td>Grade 12 (std. dev.=30.1)</td>
<td>296.6 (1.0)</td>
<td>310.4 (2.5)</td>
<td>13.8 (2.7)</td>
</tr>
<tr>
<td></td>
<td>Reading Comprehension</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Grade 4 (std. dev.=40.3)</td>
<td>215.9 (1.1)</td>
<td>230.2 (2.2)</td>
<td>14.3 (2.5)</td>
</tr>
<tr>
<td></td>
<td>Grade 8 (std. dev.=39.4)</td>
<td>258.1 (1.0)</td>
<td>275.4 (1.9)</td>
<td>17.3 (2.1)</td>
</tr>
<tr>
<td></td>
<td>Grade 12 (std. dev.=43.0)</td>
<td>288.7 (0.7)</td>
<td>306.3 (1.5)</td>
<td>17.6 (1.7)</td>
</tr>
<tr>
<td></td>
<td>Writing Skills</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Grade 4 (std. dev.=38.2)</td>
<td>220.1 (1.3)</td>
<td>233.5 (2.3)</td>
<td>13.5 (2.6)</td>
</tr>
<tr>
<td></td>
<td>Grade 8 (std. dev.=36.3)</td>
<td>259.9 (1.2)</td>
<td>274.4 (2.4)</td>
<td>14.5 (2.7)</td>
</tr>
<tr>
<td></td>
<td>Grade 12 (std. dev.=32.0)</td>
<td>283.3 (1.0)</td>
<td>304.8 (1.6)</td>
<td>21.5 (1.9)</td>
</tr>
</tbody>
</table>


\(^a\) Full sample standard deviations are taken from the Data Appendix of Campbell, Reese, O'Sullivan, and Dossey (1996).

\(^b\) Catholic-minus-public differences. Standard errors of the differences are calculated by taking the square root of the sum of squared standard errors for the separate public and Catholic means.

\(^c\) Effect sizes are calculated by dividing the Catholic-minus-public difference by the total sample standard deviation. These estimates are not adjusted for social background or other differences among public and Catholic school students.
since they indicate the extent to which students have attained relatively specific competencies in each subject area. Unfortunately, the NAEP reporting mechanisms do not include breakdowns by public and Catholic school attendance; these scores are thus unavailable to the research public.

An alternative standard uses a conventional "effect size" metric to express the sector differences. Often used to compare results across studies in meta-analyses, the effect size metric expresses group differences as fractions of the full sample standard deviation for the achievement score (Glass, McGaw, & Smith, 1981). Generally, effect sizes equal to or greater than 0.10 are considered substantial, but one must consider the time frame during which the effect was generated. The longer the time period in which students are exposed to a "treatment," the larger will be the effect estimate. The effect size estimates reported in the last column of Table 1 are much greater than the 0.10 minimum standard, and indeed indicate very large achievement differences among public and Catholic school students. But these differences do not control for social background differences among public and Catholic students, and they are differences that represent the accumulation of learning differences across not only the current grade levels listed, but also all prior grades.

How do the NAEP data compare with the national longitudinal studies? Direct comparisons are not possible, since the test scores are expressed in different metrics in each study. Nonetheless, it is possible to standardize the public-Catholic differences into the conventional "effect size" metric, and to compare effect sizes across studies. The standardization selected here is again to express the sector differences as proportions of the respective full-sample standard deviations. These are presented in Table 2.

These comparisons indicate that the overall achievement differences between public and Catholic school students are fairly consistent across studies and subject areas. Among twelfth graders, the differences in mathematics and reading comprehension estimated from NELS:88 are somewhat larger than those from NAEP and from the 1982 High School and Beyond study. The NELS:88 gaps are also greater than the HS&B sector differences in the tenth grade. The NELS:88 sector differences among eighth graders, though, are smaller than the 1992 NAEP data show. The sector differences are largest in reading and mathematics across the studies, and smallest in science.

The main points from Tables 1 and 2 are that (a) contemporary national studies with independent samples of schools and students which use similar broad-range achievement tests are consistent with respect to the presence and size of large achievement differences favoring Catholic school students; and (2) the limited historical record shows that the higher scores of Catholic school students found in the most recent data were also found in the early 1980s. Again, it is important to emphasize that the "effects" shown in Tables 1 and 2 are simply the average sector differences at each grade level. Some,
TABLE 2.
AVERAGE ACHIEVEMENT TEST SCORE DIFFERENCES BETWEEN PUBLIC AND CATHOLIC SCHOOL STUDENTS EXPRESSED AS PROPORTIONS OF THE TEST SCORE STANDARD DEVIATION, BY RESEARCH STUDY AND YEAR

<table>
<thead>
<tr>
<th>Study and Year, by Grade in School</th>
<th>Catholic School Effect Sizes (Unadjusted), by Subject Area</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mathematics</td>
</tr>
<tr>
<td>12th Graders</td>
<td></td>
</tr>
<tr>
<td>NAEP 1992</td>
<td>.46</td>
</tr>
<tr>
<td>NELS:88 1992</td>
<td>.53</td>
</tr>
<tr>
<td>HS&amp;B 1982</td>
<td>.48</td>
</tr>
<tr>
<td>10th Graders</td>
<td></td>
</tr>
<tr>
<td>NELS:88 1990</td>
<td>.44</td>
</tr>
<tr>
<td>HS&amp;B 1980</td>
<td>.40</td>
</tr>
<tr>
<td>8th Graders</td>
<td></td>
</tr>
<tr>
<td>NAEP 1992</td>
<td>.36</td>
</tr>
<tr>
<td>NELS:88 1988</td>
<td>.21</td>
</tr>
</tbody>
</table>

SOURCES: NAEP data are from NCES (1994). High School and Beyond data are from Coleman and Hoffer (1987), pp. 64-65. NELS:88 data are original tabulations for this report. See Ingels, et al. (1994) for complete information on the NELS:88 project and data files.

perhaps all, of the achievement differences may be due to social background differences between public and Catholic students. Even the apparent trend toward larger differences at higher grade levels may also be a result of selection effects, if, for example, Catholic schools tend to weed out weaker students through the high school years at a greater rate than public schools.

THEORETICAL EXPLANATIONS OF SECTOR DIFFERENCES

What accounts for the sector differences shown in Tables 1 and 2? This question takes us to the center of a lively debate that began in the early 1980s and has continued to the present. The main poles of the debate are “no effects of Catholic school attendance on achievement” versus “positive effects of
Catholic school attendance.” Explanations of the latter claim are captured in several models of how Catholic school effects are generated. These competing explanations, in turn, carry quite different practical implications, and we will turn to those after reviewing the main arguments.

**Individual Selection**
The null explanation is that sector differences simply reflect differences in the kinds of individual students attending public and Catholic schools. According to this theory, achievement in Catholic schools is higher simply because Catholic school students are higher achievers to begin with or come from families which promote achievement more effectively. If public and Catholic school students with similar backgrounds and initial levels of achievement were compared, this theory predicts that no differences in final levels of achievement would be found.

In contrast to the selection-based explanation, other theories predict greater effectiveness for Catholic schools. The following paragraphs identify four distinct hypotheses, or models, which have been articulated by different researchers over the past decade or so.

**Aggregate Student Composition**
The aggregate-student composition hypothesis has been best articulated in the public-Catholic school debate by McPartland and McDill (1982). This hypothesis asserts that Catholic schools have higher levels of achievement as a direct consequence of having a more selective student body. The key idea here is that a student will learn more if he or she has higher-achieving peers.

This contextual effect, over and above the individual effects of background, has been developed in two different directions in the sociology literature. The first conceives context as essentially a social-psychological, normative factor that affects the orientations and efforts of students and school staff. Aggregations of students along socially salient lines of stratification lead to certain kinds of “collective representations” or shared self-images that in turn shape norms of behavior and individuals’ attitudes and actions. The indicators of composition most widely used are average student socioeconomic status and the proportions of students who are minorities (typically African-American, but also Hispanic in some studies). McPartland and McDill (1982) emphasize this normative conception of school composition effects.

A second variant of the composition model can be derived from the sociological literature on ability grouping and tracking. Barr and Dreeben (1983) note that classes with initially higher average achievement tend to move at a more rapid pace than classes with lower average achievement. Students with the same level of initial achievement learn more in the faster classes than in the slower ones. While the composition of classes is the key variable in this
conception, class composition is likely to be shaped in important ways by school-level student enrollment characteristics. As Barr and Dreeben argue, there is a whole "technology" of transforming school-level student "inputs" into instructional units of classrooms and groups within classes. By implication, we can thus note that the consequences of school inputs for instructional group composition are by no means automatic. Two schools with the same inputs can divide and allocate the students to instructional contexts in very different ways, reflecting differences in other resources (e.g., the number and qualifications of the school's teachers and the number and size of classrooms), goals (e.g., all graduating students must be prepared to succeed in college, versus half of the graduates must be prepared for college), and the theories linking resources to goals.

The first social-psychological variant of the aggregate-student composition model leads to the same conclusion as the individual-selection model, that Catholic schools do not do anything better than public schools. The higher achievement in Catholic schools is instead a simple result of higher and perhaps more homogeneous student inputs. The second variant, in contrast, points to a range of specific actions that school administrators and teachers take to transform school-level student enrollments into classes and, within classes, instructional groupings. One should thus not control for class-level inputs when trying to determine the effects of schools, or more properly, school sectors. The appropriate variable on which to control for both variants is school composition, instead of class-level compositional variables.

**Competitive Market**

The *market hypothesis* is given its most elaborate articulation in Chubb and Moe's *Politics, Markets, and America's Schools* (1989). The main argument advanced by the authors is that the competitive position of private schools (Catholic and other private schools included) essentially forces them to be more responsive and accountable to their constituencies: and this in turn leads to higher student achievement. Public schools are largely shielded from market forces, due to the barriers to parental choice that the state has constructed. These barriers include the tax code, which forces parents with children in private schools to pay taxes to support public schools and to pay the private school tuition. It also includes the system of democratic control over the public school system, which leads to bureaucratic standardization and a lack of responsiveness to parental demands.

**Institutional Charter**

The *institutional charter argument* is most acutely developed in Bryk, Lee, and Holland's *Catholic Schools and the Common Good* (1993). Their argument is that Catholic schools were originally developed, or "chartered," to teach a common academic curriculum to all students, and that this mission
has been largely maintained. The academic tradition is rooted in and is still invigorated to a significant degree by Catholic religious and social ideals. Though perhaps most pronounced in schools sponsored by the Jesuit order, the ideas of literacy as a means of recovering the truths contained in sacred texts and commentaries, and the well-honed intellect as a tool to defend the faith are still part of the Catholic school ideology. Another key element of Catholic ideology that affects schooling is the belief in the fundamental equality of all peoples under God. The linking of this idea of equality to an active faith in God gives the notion of equality more immediacy than the secular idea of equality rooted in the abstract idea of “citizen” that defines the public school ideal. Perhaps as a result, Catholic schools tend to expect all students to complete an academic curriculum, whereas public schools are only beginning to consider ways to bring more students into the academic program fold.

Though not developed by Bryk, Lee, and Holland, another dimension of the historical charter of Catholic schools, one discussed by Greeley in his 1982 book *Catholic High Schools and Minority Students*, may also continue to have an impact. This dimension is more of a reflection of American Catholics’ struggle for social equality than of religious belief per se. Specifically, part of the motivation to maintain a common academic curriculum may be rooted in the upward-mobility orientation of the Catholic ethnic groups that many of the Catholic schools were built to serve. Discrimination from the Protestant establishment that controlled most local public school districts may well have had the effect of relegating the sons and daughters of Catholic immigrants to vocational or dead end programs, had they stayed in the public system. Outperforming their public school counterparts may then have been a goal of some Catholic educators and constituencies, for the prejudiced and unresponsive public schools would have limited the achievement and future life chances of Catholic youth. In the present, this orientation would be concretely manifest in a competitive spirit among Catholic educators vis-a-vis their public education counterparts, and this competitive spirit would be independent of the market position of the Catholic schools.

**Functional Community and Social Capital**

The functional community hypothesis is articulated by Coleman and Hoffer in *Public and Private High Schools: The Impact of Communities* (1987). Coleman and Hoffer see Catholic school advantages as accruing in essentially accidental ways. Rather than emphasizing the force of institutionalized practice, they point instead to the immediate social structures in which the schools are embedded. Particularly important, they argue, is the greater likelihood of Catholic school parents to know one another through church participation and to thus be able to exchange information and establish norms about the conduct of their children’s schooling. The social ties that are built
on the basis of participation in church (or other institutions, such as work) constitute a "functional community," which contrasts with a "value community" that is defined by shared values rather than concrete relations with face-to-face contacts. The ties among individuals in the functional community can be viewed as "social capital" that parents can draw upon to help steer their children in productive directions.

Like the aggregate-student composition model, this theory points to explanatory mechanisms that do not entail any greater effort, talent, or better organization on the part of Catholic school educators. The schools' greater effectiveness is rather due to the more fortunate circumstances of Catholic school students, particularly the greater social capital available to them from their parents' social networks. But it is important to emphasize that this social capital is constructed by parents through participation in a community; such capital contrasts with the aggregated human capital or emergent class consciousness of the composition model, which does not necessarily entail social relations among parents. It is thus the mutually reinforcing fit of school and community that Coleman and Hoffer emphasize.

THEORIES OF CATHOLIC SCHOOL EFFECTS, INDIVIDUAL DECISIONS, AND SOCIAL POLICY

While there can be no argument over whether Catholic school students score higher on achievement tests than their public school counterparts, the different possible explanations carry significantly different connotations for both individual families and public policy. For families, the individual selection effects model implies that Catholic schools produce no added value to their students' levels of achievement. If greater achievement is the parents' goal in enrolling their child in Catholic school, evidence supporting this theory would suggest that parents are wasting their money, at least in terms of enhancing their child's measured achievement. All of the other models, in contrast, do imply a value-added dimension to the positive effect of Catholic schools and support the notion that their children are learning more in the Catholic schools than in their public alternatives.

Similarly, the different theories hold different implications for public policy. Policy arguments which favor reducing the financial constraints on parents enrolling their children in Catholic schools are supported by images of more effective teachers, sounder curricula, and more astute administrators. In contrast, few would argue in favor of expanding access on the basis of individual or aggregate selection effects. These simply point to the power of segregating students rather than to superior effort or school organization.

This is obviously true for the individual selection model, but one could argue that aggregate selection effects are still a positive effect of the Catholic schools. That is, by concentrating students from more advantaged back-
grounds in a single Catholic school instead of a range of more heterogenous public schools, Catholic schools are producing higher levels of achievement among those students than they would otherwise have realized. But what is missing from that argument is a balancing of the gains with the negative effect of concentration on students remaining in the public schools. The positive aggregation effects thus imply negative segregation effects, but the exact balance of the gains and losses is not known. The gains may exceed the losses, or gains and losses may balance in a "zero-sum" way, or the losses may exceed the gains. In any case, to the degree that aggregate selection effects are present, an expansion of the Catholic sector is not optimal in the classical Pareto formulation of at least some gains coupled with no losses.

The functional community hypothesis, in contrast, points to a Pareto-optimal condition. According to this hypothesis, Catholic schools achieve better results because of the match of the school with the community of families it serves. The ties among families are built through membership in the church and participation in church-related activities. A school linked to that community will benefit from the strengths of the community. If the school a child attends is not linked to the community in which his or her family participates, then the benefit of the community dissipates and does not go elsewhere.

**REVIEW OF EMPIRICAL RESEARCH**

The point of departure for a review of national survey-based research is the comparison of public and Catholic schools presented in Coleman, Hoffer, and Kilgore's *High School Achievement: Public, Catholic, and Other Private Schools Compared* (1982). Though the data they analyzed were cross-sectional (from the 1980 base year of the High School and Beyond survey), the definition of the research problem and the methods they used established standards that virtually all subsequent research has adopted. This is true of the several analyses of the High School and Beyond longitudinal data, as well as the analyses of the NELS:88 data that have appeared in the past few years. The key aspects of that approach are as follows:

(a) A "Catholic school effect" on achievement (and other outcomes) consists of the unexplained difference between public and Catholic school students remaining after statistically controlling for differences in the students' backgrounds.

(b) The presence and size of the Catholic school effect may depend on the background of the students. Coleman et al. (1982) claimed that Catholic schools were more effective for students from relatively disadvantaged social backgrounds, leading them to characterize Catholic schools as more closely approximating the "common school" ideal of American education than their public school counterparts.
A complete explanation of the Catholic school effect would entail two levels of explanatory effort. The first consists of identifying the specific mechanisms or ways in which public and Catholic schools differ, such that once those factors are statistically controlled, the Catholic school effect disappears. Beyond that statistical accounting, an adequate theory of sector effects would explain why public and Catholic schools differ on the mechanisms of interest. In other words, one might find that the Catholic school effect disappears when one compares the achievement of students taking rigorous college-preparatory programs of courses in high school. In one sense, the different academic demands of public and Catholic schools may be regarded as an explanation of the achievement differences. But why do public and Catholic schools differ in the academic demands they make of students? An adequate explanation of sector effects would answer both of these questions.

The review section assesses previous studies on these three points. Specifically, the questions posed of each study are: (a) For each achievement outcome analyzed, what is the size of the estimated Catholic school effect? How was the analytic sample defined, which statistical methods were used, and which background variables were controlled and omitted in estimating the effects? (b) Were effects of Catholic school attendance estimated separately for different subpopulations, or were the effects assumed to be equal for all groups? (c) Did the authors attempt to account for Catholic school effects in terms of specific schooling variables, and if so, which were most important in explaining the Catholic school effects? How did the authors interpret their findings with respect to the larger theoretical models set forth in the first section?

SECTOR EFFECTS ON AVERAGE ACHIEVEMENT SCORES

Findings from the 1980 Base Year of High School and Beyond

Almost all contemporary literature on Catholic school effects on academic achievement is tied, either directly or indirectly, to the Coleman, Hoffer, and Kilgore (1982) analysis of the 1980 High School and Beyond data. Those data were drawn only from the base year of HS&B, and were thus cross-sectional. Coleman et al. (1982) used multiple regression to control for a large set of social background variables which were likely to be correlated both with being in a Catholic school versus public school, and with the achievement outcomes. The control variables included a set that is clearly not affected by sector and a set that may be affected to some extent by sector. Variables that are clearly not influenced by the student’s enrollment in a Catholic school included:
Family income
Mother's and father's education
Race and Hispanic ethnicity
Number of siblings
Both parents present in the home
Number of rooms in the home
Mother's working before child was in elementary school
Mother's working when child was in elementary school

Variables that may be affected by enrollment in a Catholic school, but which Coleman et al. also treated as controls included:

- Encyclopedia or other reference books in the home
- More than 50 books in the home
- Typewriter in the home
- Pocket calculator in the home
- How often student talks to parent about personal experiences
- Whether mother thinks student should go to college after high school
- Whether father thinks student should go to college after high school

When these variables were controlled for in the regression analyses, Coleman et al. (1982) found significant effects of Catholic school attendance on achievement among both the 1980 sophomores and 1980 seniors. The estimated Catholic school effects are about one half to one third the size of the unadjusted effects shown in Table 2. In the effect size metric of Tables 1 and 2, Coleman et al. estimated effects of Catholic school attendance on achievement scores of the 1980 sophomores to be 0.16 in reading, 0.20 in vocabulary, and 0.15 in mathematics. The effects for the 1980 seniors were similar: 0.12 in reading, 0.29 in vocabulary, and 0.14 in mathematics.

These results were quickly criticized in several journal articles (special editions of the *Harvard Education Review* in 1981 and *Sociology of Education* in 1982 and 1983). The main criticism was that Coleman et al. did not adequately control for background differences among public and Catholic school students, especially the ability or initial level of high school achievement among students. Without randomized assignments of students to schools, one can never fully rule out the possibility of selection effects. Controlling for the effects of background variables associated with the type of school attended and the achievement outcomes is one strategy for approximating the experimental situation; using longitudinal data to compare changes on the outcome variables is another. Since the High School and Beyond project administered the same achievement tests to the 1980 sophomores two years later, when most were then seniors, this method also became available to the research community.

When the HS&B first follow-up data were released in late 1984, several teams of researchers began to assess the effects of Catholic schools on
achievement gains over the two-year period. The results largely confirmed the findings of the cross-sectional analyses, indicating significant positive effects of Catholic school attendance on test scores. The methodology was essentially the same as before, but with the added control for sophomore level of achievement. Adding the sophomore score to the model meant that the Catholic school effects would be confined to those occurring during the last two years of high school. The cross-sectional analysis, in contrast, does not yield effect estimates that specifically relate to the high school years. The cross-sectional estimates are instead cumulative effects that have built up over the full span of schooling. A serious limitation of the Coleman et al. (1982) analyses was that it was not possible to determine when the Catholic schools had their effects. The effects could have been generated (a) in equal increments over the entire Catholic school (elementary and secondary) career of the student, or (b) only during some subset of that career. While it was assumed in the debate that the high school years were the critical ones, this assumption had no supporting data. Interestingly, the HS&B 1982 data include retrospective self-reported indicators of whether the student was enrolled in a public, Catholic, or other private school at each elementary and secondary grade level, and those data would allow one to see whether the size of the sector effects depends on the duration and on specific grades in which a student is enrolled in Catholic school. This is addressed again shortly.

Sector Effects from the HS&B 1982 First Follow-up
In any case, the availability of the HS&B 1982 first follow-up data allowed researchers to compare growth rates of public and Catholic students from the end of the sophomore year to the end of the senior year. The first round of analyses was featured in a dedicated issue of Sociology of Education in spring of 1985; Coleman and Hoffer published a monograph titled Public and Private High Schools which extended those analyses and analyzed other outcomes in 1987. The models for estimating the Catholic school effects on achievement included essentially the same set of background variables as the base year analyses, plus controls for the sophomore achievement scores. Analyses by Hoffer, Greeley, and Coleman (1985), Willms (1985), and Alexander and Pallas (1985) converged on the finding that Catholic schools contribute from 0.03 to 0.04 additional standard deviation units during the junior and senior years of high school (Jencks, 1985). This statistic is an average across all six achievement tests in the HS&B battery: Reading, Vocabulary, Mathematics, Writing, Science, and Civics. The Catholic school effects were largest in Writing, Vocabulary, Mathematics, and Reading, but close to zero in Science and Civics.

These effects are small in absolute terms, but become quite large if cumulated over several years. The question is, "From how far back do they cumulate?" Jencks (1985) proposed one estimate, based on the students'
reports of how many years they spent in Catholic schooling. From those data, it appears the average Catholic high school senior has spent about 9.5 years total in Catholic schooling. As an estimate of the cumulative effect of Catholic schooling on achievement by the end of high school, Jencks uses the predicted senior score difference (the average across the six tests) between public and Catholic students with average social backgrounds (restricted to SES and race-ethnicity). Those estimated cumulative sector effects range from only about 0.11 to 0.22 standard deviation units. When these are divided by the 9.5 average number of years, the average annual effect becomes only about 0.01 to 0.02 standard deviations per year.

These estimated effects are much smaller than those for the last two years of high school. Jencks (1985) notes three possible explanations for the difference: (1) the estimates may differ simply by chance, since the differences at the level of one-year estimates are probably not statistically significant; (2) public and Catholic elementary schools may be equally effective, so that the only Catholic school effects on achievement are during the high school years; and (3) sector effects may decay over time, so that advantages of Catholic elementary school attendance may have been realized but then lost in later years. This last point is not clear, for Jenck's analogy of the decaying effects of compensatory education program participation does not correspond to the sector comparisons. Decaying effects are usually conceptualized with respect to what happens after one leaves a treatment, but this was not the typical case for the Catholic high school students for whom the cumulative effects were estimated.

Tests of the Aggregate-Student-Composition Hypothesis

This hypothesis asserts that the effects of student background must be assessed at both the individual and school aggregate levels. None of the published analyses of the HS&B base year data included measures of school-average or other aggregated student characteristics, but the hypothesis was addressed in two separate analyses of the HS&B first follow-up survey. Neither study found much support for the hypothesis. Willms (1985) included school-mean student socio-economic status (SES), the proportion black, and the proportion Hispanic in his models. Including these measures did not, however, substantially reduce the estimated Catholic sector effects for any of the achievement outcomes, and Willms thus dropped them from his final models.

Coleman and Hoffer (1987) augmented Willms's aggregate measures with school-average sophomore verbal and mathematics achievement scores, and the proportion of students in the school who reported (retrospectively, when they were sophomores) that they planned to go to college when they were in the ninth grade. This set of variables thus provided a much more rigorous test of the hypotheses than Willms's set. The results showed that these
added controls explained only about one tenth of the Catholic school effect on verbal and mathematics achievement growth of blacks and Hispanics, none of the Catholic effect on white students’ verbal score growth, and about one third of the Catholic school effect on math achievement growth of white students (Coleman and Hoffer, 1987).

NELS:88 Analyses of Sector Effects
The next major data collection effort that would permit analyses of the effects of Catholic schooling on achievement scores was the National Education Longitudinal Study of 1988 (NELS:88). In light of the numerous articles and books published on the subject with the High School and Beyond data, it is surprising that no comparable studies have yet appeared using the NELS:88 data. This database presently has test scores available for when students were in the 8th, 10th, and 12th grades. It thus allows researchers both to attempt cross-validation with HS&B, and, if effects are found, to gain a better picture of the timing of the effects.

The only published analysis of Catholic school effects on achievement growth using the NELS:88 data is an article by Gamoran (1996) analyzing the relative effectiveness of different kinds of high schools located in urban areas. His analysis is confined to the 1988-1990 period, from when the NELS:88 students were in the spring of their eighth grade to the spring of their tenth grade (unless they dropped out or repeated a grade). Controlling for initial achievement and social background variables, Gamoran found no evidence of significant Catholic school effects on achievement growth in reading, science, or social studies compared to students in comprehensive public schools. He did, however, find evidence of a positive effect of Catholic schools on mathematics achievement of the same order of magnitude as the HS&B estimates: about .06 in the effect size metric for the two year span, or about .03 per year.

Schiller (1994) analyzed the NELS:88 mathematics scores and found significant Catholic sector effects over the 10th to 12th grade span. Controls used here included the previous cycle math score plus an extensive set of social background variables that compared closely with the set used by Coleman et al. and their critics in the HS&B analyses. These probably overestimate the Catholic sector effects, since the HS&B analyses showed that additional controls for the other achievement tests reduced the Catholic effect estimates (Jencks, 1985; Willms, 1985). Schiller does not provide information on the metric sizes of the Catholic sector effects or the test score standard deviations, and it is thus not possible to produce effect size estimates to compare with the HS&B estimates.

As an initial step toward filling out the picture of Catholic sector effects in the NELS:88 data, ordinary least squares regressions of the 1992 test scores were run on dichotomous indicators of whether the student was
enrolled in a Catholic or other private high school in 1992, the 1988 (eighth-grade) test scores, plus controls for SES, race, Hispanic ethnicity, and gender. A summary of the sector effect estimates is shown in Table 3.

The results in Table 3 are rough estimates in the sense that they do not allow for interactions of sector with the control variables, and do not include controls for school-aggregate student composition or for individual student family structure and functioning. The lack of sector interactions may diminish the Catholic sector effect estimates (a consequence of such an omission in the HS&B data), and the lack of controls for school-aggregate student composition, and individual family structure and functioning differences may increase the Catholic sector effect estimates. Past research suggests that the latter bias (overestimating the sector effects) is not likely to be large, since each regression controls for all four eighth-grade achievement tests.

At first glance, the Catholic school effects on reading, mathematics, and social studies appear to be much larger than the HS&B estimates, for which the comparable effect sizes ranged from 0.06 to 0.08. But it must be remembered that the results in Table 3 are for four years instead of two. When the NELS:88 results are divided by four, the estimated annual effects for verbal skills and mathematics match the annual estimates from HS&B fairly closely (Jencks, 1985, Table 2, shows an annual effect size estimate of about .03).

SECTOR EFFECTS ON ACHIEVEMENT SCORES OF DIFFERENT SUBPOPULATIONS

The results reviewed thus far pertain to an "average student," defined in terms of the background variables included as controls in the models of achievement. But what holds true on average may be very far from the mark for students from different backgrounds. Coleman, Hoffer, and Kilgore (1982) found that the overall impact of social background on achievement, as indexed by the $R^2$ statistic for the achievement-on-background regression equations, was weaker among Catholic school students than among public school students. Specifically, they found that (a) parental education and minority status had smaller effects on achievement among both the HS&B sophomores and seniors, and (b) the effects of parent education and race-ethnicity were smaller among Catholic school seniors than among Catholic school sophomores, whereas the opposite pattern was found among public school seniors and sophomores. Thus, social background effects were smaller and appeared to diminish as the students progressed in the Catholic schools. This led Coleman et al. to argue that the Catholic schools were closer to the "common school" ideal that animated the development of the public school system than were the public schools themselves. Andrew Greeley (1982) independently obtained similar results and drew similar conclusions from his analysis of minority student academic outcomes, also using the HS&B base year data.
TABLE 3.
ESTIMATED EFFECTS OF CATHOLIC HIGH SCHOOL ATTENDANCE ON 8TH GRADE TO 12TH GRADE ACHIEVEMENT GAINS, CONTROLLING FOR 8TH GRADE ACHIEVEMENT AND SOCIAL BACKGROUND: NELS:88 1992 SECOND FOLLOW-UP AND 1988 BASE YEAR DATA (N=12,105)

<table>
<thead>
<tr>
<th>Achievement Test</th>
<th>b</th>
<th>t-value</th>
<th>1992 Std. Dev.</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Comprehension</td>
<td>1.32</td>
<td>3.53</td>
<td>10.36</td>
<td>.128</td>
</tr>
<tr>
<td>Mathematics</td>
<td>1.81</td>
<td>4.06</td>
<td>14.65</td>
<td>.124</td>
</tr>
<tr>
<td>Science</td>
<td>0.23</td>
<td>1.02 (ns)</td>
<td>6.31</td>
<td>.037</td>
</tr>
<tr>
<td>History/Civics</td>
<td>0.84</td>
<td>4.08</td>
<td>5.49</td>
<td>.153</td>
</tr>
</tbody>
</table>

NOTES: Estimates obtained from OLS regressions of 1992 test scores (expressed in what NELS:88 refers to as the "true score number of items correct" metric) on dichotomous indicator of whether the student was enrolled in a Catholic high school in 1992, whether the student was enrolled in another private high school in 1992, all four 1988 tests, plus family SES, 0-1 indicators of whether the student is black, Hispanic, or Asian; and 0-1 indicator of whether student is female. The regressions were weighted with the second follow-up design weight. The t-values reported here reflect a downward adjustment for an assumed design effect of 1.5. The standard deviations of the 1992 test scores are for the full sample of students.

Although these findings were not pursued in relation to a prior theory, they attracted considerable interest and criticism. The main point from the HS&B base year debates and reanalyses was that the observed pattern was consistent with either of two alternatives: (a) Catholic schools do a better job of promoting learning among traditionally less-advantaged youth; or (b) the nominally "less-advantaged" youth that Catholic schools enroll are an especially select group, with unmeasured background characteristics which lead to the higher levels of achievement remaining after controlling for measured background (Goldberger & Cain, 1982). While there was no evidence for the validity of (b), it could not be refuted, since the critical factor was not measured.

The common school hypothesis about Catholic schooling was also re-examined with the HS&B 1982 follow-up data. Probably the best operationalization one could make of the Goldberger-Cain unmeasured selection hypothesis would be to control for the sophomore achievement scores in regressions of senior achievement on the background variables of interest. Regressing senior achievement on sophomore achievement plus SES, race, and Hispanic ethnicity, Hoffer, Greeley, and Coleman (1985) found that the effects of sophomore achievement and the three social background variables were consistently smaller than the effects in the public schools. Across the six HS&B achievement tests, the effects of the four variables were closer to zero in the Catholic schools in 22 of 24 comparisons. However, only a few of the differences in coefficients proved to be statistically significant, and the
results are probably best regarded as suggestive, rather than conclusive (Jencks, 1985).

In later work with the HS&B follow-up data, Coleman and Hoffer (1987) examined relative effects of additional background factors on achievement in the two school sectors. They distinguished between measures of "traditional disadvantage," which include social class, gender, race, and ethnicity and "deficiency," which include household structural and functional problems. Measures of deficiency they used in their analysis included whether there was a single parent head of household, whether the mother worked outside the home before the child started school, the extent to which the parent and child talked about personal issues, and whether the parents expect the child to go to college after high school. These are not particularly good measures of the constructs, but they were all the HS&B data had to offer, and all showed significant effects on achievement in the full sample. The results for the measures of deficiency indicated that they were much less related to verbal achievement in the Catholic schools than in the public, but only slightly less related in the Catholic schools to mathematics achievement.

WHAT DO CATHOLIC SCHOOL STUDENTS DO DIFFERENTLY?

The preceding two sections have described the basic findings on Catholic school effects on achievement outcomes. The results indicate that Catholic high school attendance has positive effects on high school students' achievement in verbal skills and mathematics, but not in science or social studies. Moreover, Catholic school effects are greater for lower-SES, minority students, and for students from families with one or more structural or functional obstacles. The question now addressed is how Catholic schools produce these effects. The main explanations that have been advanced are couched at two levels: the within-school mechanisms that are affected by school sector and which affect student achievement, and the larger differences between schools that shape these internal schooling factors.

As Bidwell and Kasarda (1980) argue, differences in average achievement between schools are ultimately explained by differences in the specific kinds of schooling experiences students have within schools. In other words, if one wants to understand why school-level outcomes differ, one must look for differences in the outcome-producing processes internal to the schools. This points to a certain logical priority of the within-school factors, and the research has in fact developed accordingly. The first efforts to give an account of the Catholic school effects were undertaken by Coleman et al. (1982). The framework they used generally distinguished two sets of factors: disciplinary standards and academic standards.

The HS&B data contained several measures of individual students' con-
compliance to school rules, including absenteeism, tardiness, and class cutting. Students also reported their perceptions of the disciplinary climate in the school. The measures of academic demands included hours of homework and coursework in academic areas.

When these were factored into the regression analysis of achievement outcomes, Coleman et al. (1982) found that the discipline problems were much lower in the Catholic schools, and had significant effects on achievement outcomes. The better discipline in Catholic schools accounted for most of the Catholic sophomores' higher achievement in reading and mathematics, compared to public school sophomores. In contrast, the measures of academic demands accounted for relatively little of the Catholic sophomore advantages. Catholic students did more homework than public students, and that accounted for a small but significant part of the achievement differences, but the coursework measures did not explain much at all. For the 1980 seniors, though, the Catholic advantage in mathematics was explained by the greater number of math courses completed by Catholic students. The different effect of the math coursework variable on the sophomore and senior achievement scores reflects the fact that the measures of math coursework were much better for the seniors.

The explanatory picture changed when the HS&B follow-up data were analyzed. The Coleman et al. (1982) base year analyses of the schooling variables, like their analysis of achievement scores more generally, were criticized because of the lack of controls for prior achievement (Goldberger & Cain, 1982). A number of critics also made the point that Catholic school students looked better on the discipline and academic variables because they are primarily academic-track, college-bound students. If Catholic school students were compared to academic-track public school students, it was suggested, the differences would disappear.

Coleman and Hoffer (1987) responded to the first of these criticisms by incorporating controls for sophomore achievement differences in their regressions for estimating Catholic school effects on the schooling variables. The argument about Catholic school students being best compared with academic-track public school students is based on a faulty premise, however. The argument assumes that all of the Catholic students would be placed in an academic track if they were in the public sector. But that is not supported by the data. Catholic school students are indeed much more likely to report being in an academic or college-preparatory program of studies than public school students with comparable social backgrounds. Overall, about 70% of Catholic school seniors reported being enrolled in an academic program, versus about 40% of public school students. But when the effects of sophomore achievement scores, SES, parent expectations, and the other social background variables on program are controlled for, the 30% sector difference reduces to about 18%, which is far from having disappeared (Coleman &
Hoffer, 1987). In other words, there is a substantial independent effect of Catholic school attendance on the likelihood of academic program enrollment.

Coleman and Hoffer (1987) present an analysis that parallels the Coleman, Hoffer, and Kilgore (1982) base year explanatory effort, but which extends the original analysis to include the track variable and better measures of coursework. They also carried out the analysis separately for minority (black and Hispanic students combined) and non-Hispanic white students. The discipline variables were essentially the same as in the 1980 analysis, and included indices both of individuals' own behavior and reports of school-wide disciplinary problems. The results showed that Catholic school students completed more math, science, and foreign language courses; completed more homework; had better attendance; and observed fewer school-wide discipline problems than public school students with similar sophomore achievement and social backgrounds (Coleman & Hoffer, 1987).

When these schooling variables were entered into the achievement regressions, Coleman and Hoffer (1987) found that the academic variables (track placement, homework, and coursework) explained most of the Catholic school effect on verbal skills, and all of the effects on mathematics scores. These findings have been largely corroborated by other researchers using the HS&B data (Lee & Bryk, 1988), and, to a limited extent, with the NELS:88 data (Gamoran, 1996).

**CONCLUSIONS: WHY ARE CATHOLIC SCHOOLS MORE EFFECTIVE?**

The main points from the review and extension of empirical research are:

- Catholic high schools have positive effects on verbal and mathematics achievement, but no discernible effects on science achievement. The main evidence here is still from HS&B, but preliminary analyses of the NELS:88 data confirm the pattern at least for the last two years of high school.
- Catholic school effects are greater for students from disadvantaged backgrounds, especially with respect to family structure and functioning. The evidence for "common school effects" is consistent but statistically weak for minority and low-SES students, and more work is needed on that issue. The evidence is still entirely from HS&B; work with NELS:88 is needed.
- The main "schooling" mechanism accounting for the Catholic school effects is the greater concentration of academic coursetaking among Catholic school students. Discipline and climate variables explain a small amount of the Catholic school effects on achievement. Again, the evidence is still largely from HS&B, and NELS:88 replications and extensions are needed.
- The main "macro" theories of the Catholic sector effects are: (a) aggregate student composition, (b) market competition, (c) institutional charter, and (d)
functional community. Counter evidence from HS&B weakens the case for (a), but the other hypotheses have not been rigorously addressed with available survey or archival data.

One important shortcoming in the research to date is the lack of specificity about the presence and size of Catholic school effects at different grade levels. The most glaring gap in the research record is the lack of data to assess effects of Catholic elementary school attendance. The NAEP fourth-grade aggregated data can be used in a very limited way to compare sector differences at grades 4, 8, and 12, controlling for parent education only. But new studies are needed to collect data that would allow the kinds of comparisons that have been done with the High School and Beyond and NELS:88 data. This need will be met with the Early Childhood Longitudinal Study scheduled to start in fall 1998 with a national sample of 23,500 kindergartners in over 1,000 schools. These children will be surveyed and assessed each year through the fifth grade.

The question of when Catholic schools have their effects is also present at the high school level. The NELS:88 data suggest that the Catholic school effects are small if present at all in the ninth and tenth grades, but resemble the effects estimated with the HS&B data in the eleventh and twelfth grades. Why this inconsistency should occur is not anticipated in any of the theoretical frameworks presented in this review, and should be investigated more carefully.

Several other limitations of the research on high school achievement can also be noted. One concerns the achievement criteria upon which public and Catholic schools are compared. While the tests administered by NAEP, HS&B, and NELS:88 are valid indicators of achievement, the achievement they measure is mainly recall of facts and concepts, and literal reading comprehension, rather than problem solving and ability to "go beyond the text." These tests may measure necessary but not sufficient conditions for the effective applications of intelligence. Further research might examine whether public and Catholic school students differ in their higher-order thinking skills as well.

A second limitation of the research to date is that it has not shed much light on the inconsistency of Catholic school effects across subject areas. In particular, it is not clear why Catholic schools do not have positive effects on science achievement. Detailed comparisons of coursework should be done, using the NELS:88 transcript data.

Third, the "internal" explanatory efforts are most successful for math and less successful for reading. This may reflect the quality of coursework measures in verbal skill areas. Future data collection projects should look hard to develop stronger "opportunity to learn" indicators for language arts (and social studies).
Where does the research stand with respect to the "big theory" explanations that have been advanced? As we have seen, the individual selection accounts for at least half of the overall public-Catholic differences in verbal skills and mathematics in high school, and all of the sector difference in science. But neither the individual selection nor the aggregate-student composition explanations hold up for the positive effects of Catholic schooling on growth in verbal skills and mathematics over the last two years of high school.

What about the competing explanations of the positive effects of Catholic schools? The inconsistency of the Catholic advantages across subjects and grade levels is difficult to reconcile with the competitive market model, which predicts that Catholic schools should be more effective for all of their students than the public alternative. The institutional charter and functional community hypotheses, in contrast, are not as exacting in their predictions about achievement effects. Catholic schools may have strong charters with respect to academic emphasis and internal community, but those structures may not carry over to consistently higher performances at all grades and in all subjects. The strongest test of the charter hypothesis is to assess the extent to which the structures and beliefs that Bryk, Lee, and Holland (1993) posit are present and operative in Catholic schools. The specific conditions under which those structures and beliefs translate into academic excellence, however, remain to be identified.

Similarly, the functional community hypothesis is most directly tested, not by achievement comparisons, but by assessing the extent to which the social networks that Coleman and Hoffer (1987) predict are found to exist. Data are available in the NELS:88 survey for assessing the hypothesis with respect to parent contacts with other parents, and parent involvement in the schools, but the analyses remain to be done. Again, the specific conditions under which these community relations translate into higher achievement need to be identified and assessed.

In conclusion, the issue of the relative effectiveness of public and Catholic schools in promoting academic achievement remains unresolved in several respects. One aspect on which there is no debate is whether the Catholic schools do worse than their public counterparts in any area, after taking into account student background. No study has produced evidence of lower effectiveness. Further, where positive effects of Catholic schooling have been found, researchers have also been largely successful in linking the advantages to specific mechanisms, particularly greater academic demands. While the larger theoretical issues of markets, charters, and communities need to be further developed and tested, the model of higher academic standards that Catholic schools have come to symbolize, if not always embody, has already had a positive effect on contemporary American education, public and private.
REFERENCES


Thomas B. Hoffer, Ph.D., is a senior research scientist at the National Opinion Research Center at the University of Chicago. He works primarily on the design, execution, and analysis of large-scale education surveys. His main areas of research interest are in educational policy and social stratification. Correspondence concerning this article should be addressed to Thomas Hoffer, Senior Research Scientist, NORC, 1155 E. 60th Street, Chicago, IL 60637.

This paper was developed with support of the Lilly Endowment as part of the project, "Legacy at the Crossroads: The Future of Catholic Schools," sponsored by the Life Cycle Institute, The Catholic University of America.