Schools and Student Achievement: More Evidence from the Milwaukee Parental Choice Program

Cecilia Elena Rouse

Many states are considering programs that would provide vouchers for (low-income) children to attend private schools because policymakers believe that traditional reforms—such as reducing class sizes—will not fix an educational system that is "broken." Advocates of vouchers argue that teachers' unions and bloated bureaucracies impede such reforms from reaching the classroom and increasing student achievement. Furthermore, because children are required to attend their neighborhood school, the system has no incentive to change. Wealthier parents can voice dissatisfaction with their residential school by moving to another neighborhood or enrolling their children in a private school; however, poorer—particularly inner-city—parents cannot. Vouchers would, at a minimum, provide disadvantaged children with more educational options. If the students also received a better education in the private schools, the program might offer a cost-effective way to improve student achievement, at least for those students who use the vouchers.

In 1990, Wisconsin became the first state in the nation to implement a publicly funded school voucher program. The Milwaukee Parental Choice Program provides a voucher, worth approximately $4,373 in 1996-97, to low-income students to attend nonsectarian private schools. The program began with seven private schools, although by 1996 the number had risen to twenty.[1] At this time, religious schools are not permitted to participate in the program.[2] The participating private schools offer a variety of educational approaches, including Montessori and Waldorf, as well as bilingual and African-American cultural emphases. Although the tuition charged by many of the "choice" schools is quite low (ranging from less than $200 to about $4,000), actual expenditures per pupil are generally higher (on the order of $4,000 to $5,000 per pupil in 1996-97).[3] The balance of the revenues comes from grants, donations, and fund-raising by parents. In

Cecilia Elena Rouse is an assistant professor of economics and public affairs at Princeton University and a faculty research fellow of the National Bureau of Economic Research.

addition, because the schools are nonsectarian, many also receive Title I funding from the federal government.

Because the parental choice program is targeted to the most disadvantaged public school students, only students whose family income is at or below 1.75 times the national poverty line are eligible. In principle, the student in a family of three with a family income of approximately $21,000 is eligible to apply; in practice, the mean family income of applicants is approximately $12,300. Choice applicants are considerably more disadvantaged than the average student in the Milwaukee public schools (whose family income is $24,000); they are also more likely to be minority and have lower preapplication math and reading test scores. However, the parental education of choice applicants is comparable to that of nonapplicants.

Some argue that an unrestricted voucher program would improve the schooling of all children. In the most unrestricted program, all (or a substantial percentage) of the students in the public schools would be eligible to attend a private school. Since state funding would be tied to student enrollments, public schools would have to compete for students, as in the marketplace, which would give the schools an incentive to improve. If such an unrestricted voucher program were successful, the academic outcomes of students in public and private schools would equalize over the long run. While such effects are theoretically possible, the Milwaukee Parental Choice Program is too small to provide insight into the potential student achievement benefits of an unrestricted voucher program.[4] It cannot show whether providing vouchers would also improve the schooling of students who remain in the public schools. An analysis of the Milwaukee Parental Choice Program can, however, indicate whether the private schools participating in the program (the choice schools) are "better" than the public schools in Milwaukee.

In this paper, I review the three existing studies of the effects of the choice schools on student achievement. Two of the studies report significant gains in math for the choice students and two report no significant effects in reading. I also extend the analysis to compare the achievement of students in the choice schools with that of students in three different types of public schools: regular attendance area schools, citywide (or magnet) schools, and attendance area schools with small class sizes and supplemental funding from the state of Wisconsin ("P-5" schools). The results suggest that students in P-5 schools have math test score gains similar to those in the choice schools, and that students in the P-5 schools outperform students in the choice schools in reading. In contrast, students in the citywide schools score no differently than students in the regular attendance area schools in both math and reading. Given that the pupil-teacher ratios in the P-5 and choice schools are significantly smaller than those in the other public schools, one *potential* explanation for these results is that students perform well in schools with smaller class sizes.

Existing Studies of the Achievement Effects of the Choice Program

Three studies to date have evaluated the achievement effects of the Milwaukee Parental Choice Program. The first, conducted by Witte, Sterr, and Thorn (1995), concludes that there were no relative achievement gains among the choice students (see also Witte [1997]). The second, by Greene, Peterson, and Du (1997), finds that the choice students made statistically significant test score gains in both reading and math by their third and fourth years in the program. The third study, by Rouse (forthcoming), reports that the students selected to attend a choice school experienced significantly faster gains in math scores, but showed no differential gains in reading. To understand why these three studies generated conflicting results, one must consider two aspects of the evaluations: the selection of the control, or comparison, group and the method of controlling for family background and student ability.

Selection of the Control, or Comparison, Group

Ideally, to establish whether choice schools are better than the Milwaukee public schools, one must ascertain whether students who attended the choice schools had higher achievement gains than they *would have had* if they had attended a Milwaukee public school. Because this counterfactual is impossible to obtain, one must instead identify a group of students who did not attend a choice school; their test scores provide the yardstick against which to measure the effect of the program. This group is called a control, or comparison, group.[5] The best control group is constructed using a randomized experiment. In this social experiment, children are randomly assigned to attend a choice school (the treatment group), while others are assigned to attend public schools (the control group). After some period of time, one would compare outcomes—such as test scores, high school graduation rates, or labor market success—for the treatment and control groups. Since, on average, the only difference between the groups would be their initial assignment—which was randomly determined—any differences in outcomes could be attributed to the type of school attended.

Such an experiment, however, was not implemented in Milwaukee (nor anywhere else), forcing researchers to devise statistical methods that attempt to mimic a randomized experiment. One cannot simply compare the achievement of students in choice schools with that of a comparison group of students in the Milwaukee public schools. In Milwaukee, this simple comparison would likely show that students enrolled in choice schools fare no better than students in the Milwaukee public schools.[6] One might be tempted to conclude that the choice schools are no different than the public schools. However, such an interpretation might be misleading. Students who qualify for the parental choice program come from disadvantaged families. As a result, they generally score lower on standardized tests than wealthier, more advantaged students and would likely have continued to do so even if they had remained in the public schools. One would attribute the test score results to the schools when the results may, in fact, be due to the characteristics of the students. To estimate the true effect of the choice schools, one must control for family background (such as family income and parental education) and student ability. The goal is to control for all individual characteristics that are correlated with attending the choice school and to explain the higher test scores in such a way that the only difference between the two groups of students is enrollment in a choice school. In general, the more similar the two groups are to begin with, the more credible the evaluation of the program will be.

The choice of a control, or comparison, group is one area in which the existing analyses of the Milwaukee voucher program differ. Greene, Peterson, and Du (1997) compare the test scores of choice students with those of the group of students who applied to the program but were not accepted (the "unsuccessful applicants"); Witte, Sterr, and Thorn (1995) compare choice students with a random sample of students from the Milwaukee public schools; and Rouse (forthcoming) compares selected choice students with both the unsuccessful applicants and the students in the Milwaukee public schools. There are advantages and disadvantages to both control/comparison groups.

The unsuccessful applicants are an appealing control group because all of these students were interested in attending a choice school. Therefore, the unsuccessful applicants likely have parents who are similarly motivated to the parents of the successful applicants. In addition, the parents of all applicants must expect that their children will be well served in the program, which may not be true for the children who did not apply. There are problems with using the unsuccessful applicants as a control group, however. The first is that since the parents of all applicants were interested in enrolling their children in a private school, the parents of the unsuccessful applicants may have been more likely to enroll their children in a private school outside of the choice program, rather than re-enrolling them in a Milwaukee public school. This decision was made easier by a parallel, privately funded program—Partners for Advancing Values in Education (PAVE)—that provided scholarships to students interested in attending (primarily) Catholic schools. If post-application data on these students were available, this would not be a problem. However, the data do not track students who enrolled in either a public school outside of the Milwaukee public school system or a nonchoice private school. The second problem is that the sample sizes are extremely small. By the fourth year of the program, there were fewer than forty unsuccessful applicants to use in evaluating the program, which makes estimated effects of the program sensitive to unusually high or low test scores (Witte 1997).

One can also compare the achievement of students in the choice schools with that of a random sample of students from the Milwaukee public schools. This comparison group yields a much larger sample and is, perhaps, less subject to nonrandom attrition (after all, these students were ostensibly not interested in leaving the Milwaukee public schools). At the same time, the random sample of students from the Milwaukee public schools may have refrained from applying to the parental choice program because they thought it would not serve them well, or because their parents are less motivated or involved, which would lead to an overstatement of the achievement effects of the program. As a result, using this comparison group requires a statistical strategy that adequately controls for student characteristics.

Method of Controlling for Student Characteristics and Family Background

The second area in which the existing analyses of the Milwaukee program differ is the method of controlling for family background and student ability. Greene, Peterson, and Du (1997) control for "application lotteries"; Witte, Sterr, and Thorn (1995) control for the student's prior test scores; and Rouse (forthcoming) controls for "individual fixed effects." Again, each methodology has advantages and disadvantages.

Consider first the strategy employed by Greene, Peterson, and Du. The choice schools are not allowed to discriminate in admitting students, which is interpreted to mean that if more students apply for the school than there are seats available, the students are randomly selected from among the applicants. If a choice school is not oversubscribed, it is required to take all who apply, with only a few exclusions. Therefore, in each school in which students are randomly selected (through an application lottery), a mini-randomized experiment is conducted. If the schools truly select the students at random, then, on average, the only difference between the successful and unsuccessful applicants is whether they have been randomly selected. As a result, in theory, one could simply compare the outcomes of successful applicants with the outcomes of unsuccessful applicants and attribute the difference to whether the students were selected to attend a choice school. Moreover, because selection was random (that is, not related to student ability or parental background), one need not control for individual characteristics.[7] One can also combine all of these mini-experiments and control for variables indicating the application lottery in which each student participated.[8] (Naturally, this strategy requires using the unsuccessful applicants as a control, or comparison, group.) The primary advantage of using the unsuccessful applicants as a control group and controlling for application lotteries is that, if selection is truly random, this strategy should uncover the true effect of the parental choice program on student test scores using a method that closely resembles a randomized experiment, at least in theory.

In practice, this strategy has some disadvantages. First, the data do not contain information on the actual school(s) to which a student applied. As a result, one cannot recover the actual application lotteries. Greene, Peterson, and Du have devised a creative solution to this problem, but it is not clear how close their imputation comes to the actual lotteries.[9] A second disadvantage is that even if the lotteries are truly random and the imputation reasonably mimics them, it appears that the motivated unsuccessful applicants were more likely to attend another private school—one outside of the choice program (Rouse forthcoming; Witte 1997). As a result, by not controlling for family background, one may overstate the effectiveness of the program.

There are also advantages and disadvantages to controlling for prior test scores—the methodology implemented by Witte, Sterr, and Thorn. On the one hand, controlling for these scores has the advantage of accounting for student ability that changes over time, rather than controlling for characteristics at a fixed point in time. In addition, this methodology allows one to develop a dynamic model of test score growth in which a child's test score this year is a direct function of his or her test score last year. On the other hand, test scores may not be a good measure of ability (even ability at a fixed point in time). Moreover, the strategy may not be appropriate when applied to data on students who have been enrolled in a choice school for several years (Rouse forthcoming). Finally, one can only include students who have prior test scores in the analysis, which is a potential problem in Milwaukee, where the majority of students are not tested each year.

The strategy implemented by Rouse (forthcoming) controls for all student characteristics (both observed and unobserved) that do not change over time (that is, they are fixed, or time-invariant). These characteristics include more motivated parents, parental education, and innate student ability. The methodology is referred to as controlling for individual fixed effects. This strategy requires fewer assumptions than one using application lotteries and allows for larger samples than one controlling for prior test scores. Its primary disadvantage, however, is that it does not control for time-varying student characteristics.

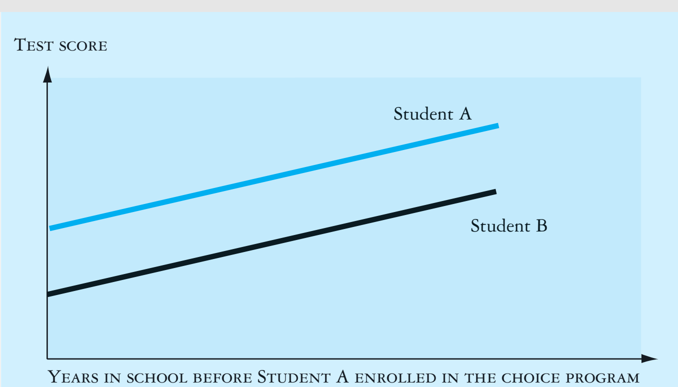
To understand this strategy (which I employ in the rest of the paper), consider two students: Student A, who enrolled in a choice school, and Student B, who did not. The diagram depicts two possible test score trajectories for the two students *before Student A enrolled in the choice program* (see box).[10] Suppose that the prechoice test scores of Student A and Student B evolve as shown in the left portion of the diagram. Here, Student A scored higher than Student B each year before Student A enrolled in the program. This may reflect the fact that Student A was more "able" than Student B, and one would not want to attribute the test score difference to the choice schools since it existed even before Student A enrolled in the choice program and it would likely have continued to exist even if Student A had remained in a Milwaukee public school. Fortunately, in this case the fixed-effects analysis will uncover the true (unbiased, in statistical terms) effect of the choice schools on student achievement.

The fixed-effects analysis will, however, lead to an overstatement of the program's effects if Student A had faster test score gains than Student B before Student A enrolled in the choice program. In this case (shown in the right portion of the diagram), the fixed-effects analysis will attribute the faster achievement growth to the choice program when, in reality, students in the choice program would have had faster test score growth even if they had remained in the Milwaukee public schools. To assess whether this potential problem likely explains the entire estimated program effect, I analyzed the preapplication test score trajectories of students in the choice program and those in the Milwaukee public schools. This exercise indicated that the results obtained using individual fixed-effects estimates are probably not overstated.

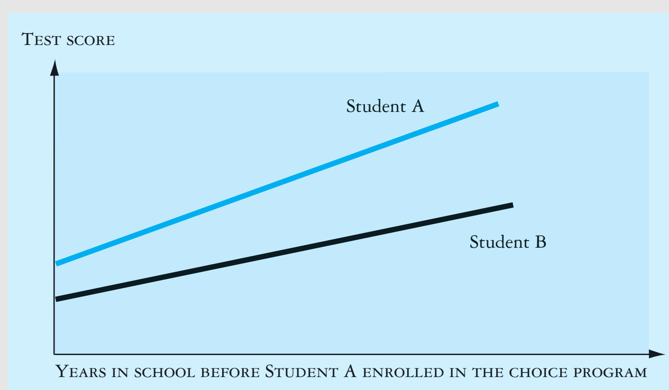
Understanding Individual Fixed-Effects Estimates

Student A is enrolled in a choice school and Student B is enrolled in a Milwaukee public school. Consider their test scores before Student A enrolled in the choice program:

Individual fixed-effects estimates will generate the true effect if both Student A and Student B had the same growth in test scores before Student A enrolled in the choice program.



Individual fixed-effects estimates will overstate the effect if Student A had faster test score gains even before he or she enrolled in the choice program.



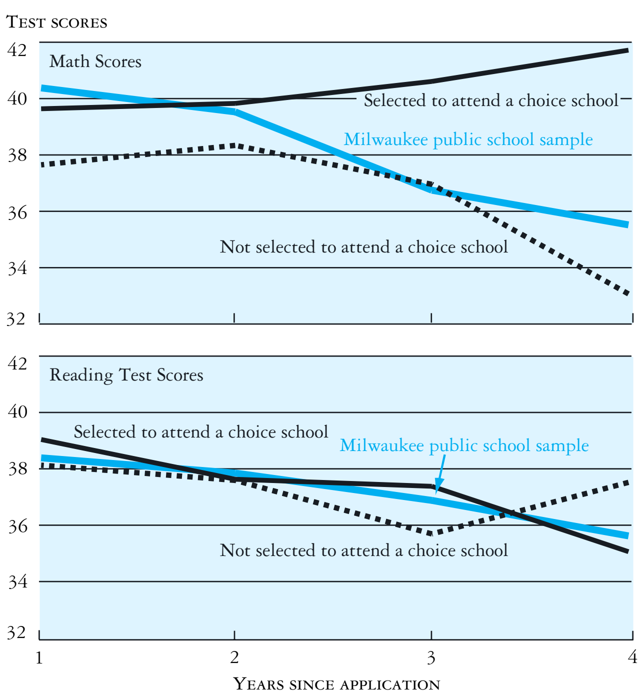
Are the Private Schools in the Choice Program "Better" Than the Milwaukee Public Schools?

Comparing Choice Schools with All Milwaukee Public Schools

Chart 1 compares the test scores of students selected to attend a choice school with those of both the unsuccessful applicants and the random sample of students from the Milwaukee public schools, controlling for individual fixed effects.[11] Note that I use the test scores of those *selected* to attend a choice school, whether or not the student ever enrolled in a choice school or eventually returned to the Milwaukee public schools. I show these results for two reasons. First, making vouchers available is the only policy instrument open to policymakers. If the state of Wisconsin decides to provide educational vouchers to all low-income students, not all will take advantage of the program and not all who enroll will remain in it. In the extreme case in which no students actually use the vouchers, even if the choice schools are much better at educating children than the public schools are, there will be no achievement gains from the program. Thus, comparing the test scores of students who are selected (whether or not they actually are enrolled in a choice school) reflects the overall potential gains from offering the vouchers. Second, students who leave the choice schools may do so because they are not flourishing there. In this case, an analysis that compares the test scores of students who remain enrolled in a choice school may overstate the true effect of the program.[12]

Chart 1

Estimates of Math and Reading Test Scores for Students Selected to Attend a Choice School, Applicants Not Selected, and Students in the Milwaukee Public Schools



Source: Rouse (forthcoming).

Note: The estimates control for individual fixed effects (for example, they are corrected for ability and family background).

The top panel of Chart 1 shows that students selected for the choice program made yearly gains in math achievement, particularly beginning in the second year after application. It also reveals that both the unsuccessful applicants and the students in the Milwaukee public school sample experienced large declines in their math test scores in the third and fourth years. The bottom panel shows the trends for reading scores. It is clear that there are no discernible differences in the reading test scores between the three groups.

Given that the trends for the unsuccessful applicants and the students in the random Milwaukee public schools sample are similar, Chart 1 shows that any differences between the three existing analyses do not hinge on the selection of a control, or comparison, group (provided that family background is adequately controlled for). In addition, the math results in the chart are consistent with those reported by Greene, Peterson, and Du (1997), and the reading results accord with those reported by Witte, Sterr, and Thorn (1995). The fact that the math results agree with those reported by Greene, Peterson, and Du indicates that *in these data*, if one adequately controls for student characteristics, it does not make a large difference whether one defines choice students as those who are *selected* to attend a choice school or as those who are *enrolled* in a choice school. In contrast, the reading results conflict with those reported by Greene, Peterson, and Du, largely because the authors' results disappear when one includes individual fixed effects. The math results conflict with those reported by

Witte, Sterr, and Thorn because of differences in our specifications and samples.[13]

It is also worth noting that these data are far from ideal for an evaluation of the choice program. The fact that students who were not enrolled in either a choice school or a Milwaukee public school were not included in the data leads to concerns about nonrandom sample attrition. In addition, because of changes in the tests administered in the public schools, some data are imputed.[14] I continue to estimate results similar to those presented in Chart 1 when I attempt to control for both sample attrition and data imputations. Nevertheless, statistical techniques cannot substitute for better data, so these data deficiencies should be kept in mind when interpreting the results.

Comparing Choice Schools with Different Types of Public Schools

Other studies have also found that private schools perform better than public schools (see, for example, Coleman, Hoffer, and Kilgore [1982a, 1982b], Evans and Schwab [1995], Neal [1997], and Sander [1996]). Many attribute the observed superiority of private schools to the fact that these schools compete for students.[15] However, few have attempted to look within the "black box" of private school success to understand why the schools may be successful. Those who *have* looked point to differences in homework, curriculum, decentralized governance, and social integration (Bryk, Lee, and Holland 1993; Coleman, Hoffer, and Kilgore 1982a; Coleman and Hoffer 1987). I attempt to look more closely at the apparent Milwaukee private-public school differences in achievement by focusing more intensely on the public schools.[16]

The Milwaukee public school district consists of approximately 145 schools. The district operates a controlled choice program in which first-time students in Milwaukee's public schools, students who reach the top grade of their school, and students desiring to transfer from their attendance area school are required to select three schools in which they would like to enroll. If a school is oversubscribed, selection is based on a random lottery with preference given to children attending the feeder schools, those with siblings already enrolled in the school, and those living in the attendance area or nearby (Milwaukee Public Schools 1997).

Within the district there are approximately thirty citywide (or magnet) schools, which were created in the 1970s to facilitate desegregation. Many of these schools are specialized, offering foreign language immersion, gifted and talented and performing arts instruction, and Montessori, Waldorf, and Global Learning educational approaches. Approximately 22 percent of the total Milwaukee public school enrollment is in citywide schools.[17] Many researchers (for example, Archbald [1995]) hypothesize that citywide schools should be better than regular attendance area schools because citywide schools compete for students (at least within the district). In Milwaukee, this competitive effect may be muted, however, because although the citywide schools are designed to accommodate students from all over Milwaukee, many of them allocate over half of their available seats to children who live close to the school (Milwaukee Public Schools).

Finally, a group of fourteen schools (known as "Project Rise Schools") whose students are predominately minority and extremely disadvantaged were exempted from desegregation. Instead, they were provided with extra funding from the state. Today, these fourteen schools, along with about seven others, participate in the Preschool to Grade 5 Grant Program, and are known as P-5 schools;[18] they enroll about 15 percent of the total public school students and 25 percent of the elementary school students. This program provides supplemental state grants to schools with high proportions of economically disadvantaged and low-achieving students. In theory, eligible schools are required to maintain pupil-teacher ratios of under twenty-five to one, institute annual testing in basic skills, identify students needing remedial education, increase parental involvement, provide in-service training, and conduct staff evaluations (Clancy, Toulmin, and Bukolt 1995). In practice, the schools primarily comply with the small class size requirement. In 1993-94, Wisconsin allocated $6.7 million to the P-5 schools, which amounted to grants of approximately $500 per child.[19]

To assess whether student achievement varies among the different types of public schools, I estimate the effect of the total number of years in which the student has continuously been enrolled, or had ever been enrolled, in the particular type of school.[20] Thus, I estimate the gap in test scores between students in "regular" Milwaukee public schools and those enrolled in choice, citywide, and P-5 schools. I control for family background and student ability by including individual fixed effects, as described above.

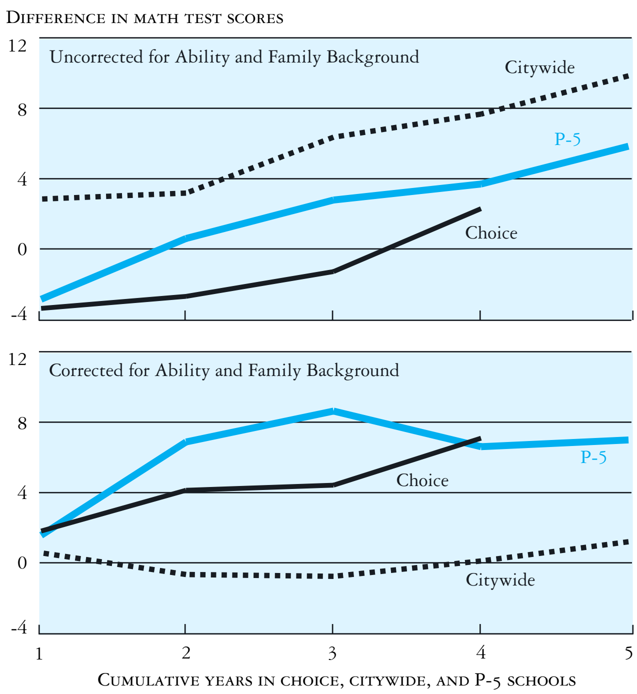
Chart 2 shows the results for math scores.[21] The differences in the top panel do not adjust for student ability and family background; those in the bottom panel do. Consider, first, the results that do not adjust for family background. These figures suggest that students in the citywide schools consistently score higher than students in the regular public schools, and the gap increases with the cumulative number of years the students have been enrolled in the citywide schools. This finding is consistent with much of the existing evidence on magnet schools (for example, Blank [1990], Crain, Heebner, and Si [1992], and Gamoran [1996]). In addition, the results indicate that although the students in the P-5 and choice schools have lower scores (than students in the regular public schools) in the first year, the rate of increase is (roughly) similar to that for students in the magnet schools.[22]

Notice, however, the effect of controlling for student ability and family background using individual fixed effects, as shown in the bottom panel of Chart 2. Once student characteristics have been accounted for, the gap in math scores between the citywide students and regular public school students disappears. At the same time, the gap between those in the P-5 and choice schools becomes large and statistically meaningful. Significantly, there is no difference in the math achievement gains of students in the P-5 and choice schools.

Chart 3 presents the reading score results. Again, before controlling for student ability and background (with individual fixed effects), I find that students in citywide schools score substantially higher than students in the regular public schools and in the choice schools (top panel). Students in the P-5 schools make incremental yearly gains, although these gains are not statistically distinguishable from zero. The bottom panel again shows that once one adjusts for individual

Chart 2

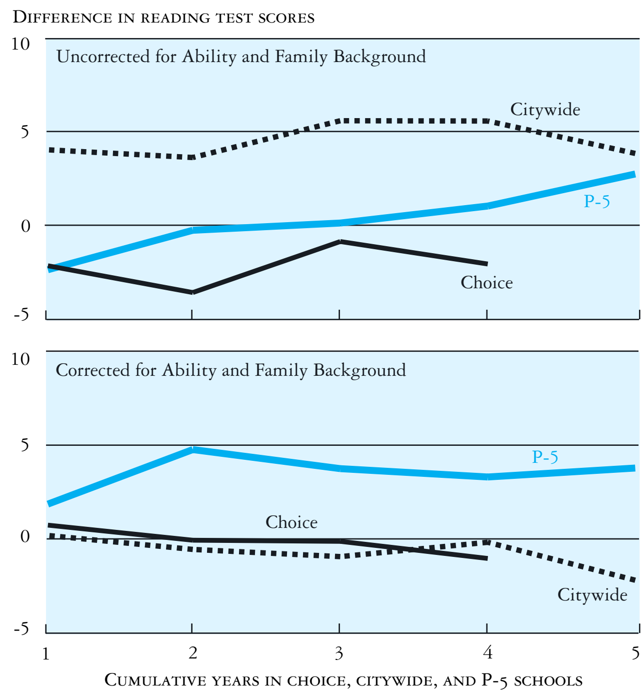
The Difference in Math Test Scores between Choice, Citywide, and P-5 Schools, and "Regular" Milwaukee Public Schools



Notes: The top panel does not control for individual fixed effects; the bottom panel does. A P-5 school participates in Milwaukee's Preschool to Grade 5 Grant Program.

Chart 3

The Difference in Reading Test Scores between Choice, Citywide, and P-5 Schools, and "Regular" Milwaukee Public Schools



Notes: The top panel does not control for individual fixed effects; the bottom panel does. A P-5 school participates in Milwaukee's Preschool to Grade 5 Grant Program.

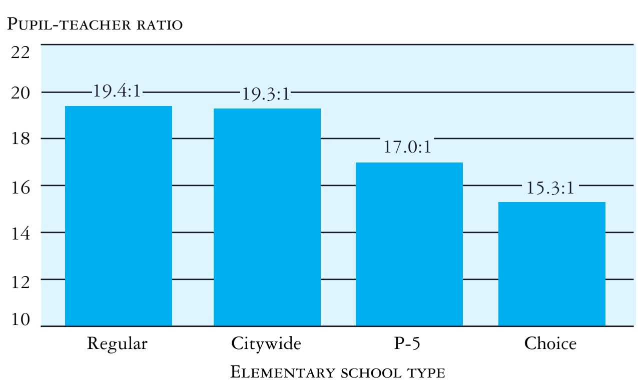
fixed effects, students in the citywide and choice schools are found not to have faster reading test score gains than students in the regular public schools.[23] In contrast, students in the P-5 schools have substantially faster gains in reading than those in the other public schools and choice schools.

Overall, these results suggest that the observed superiority of the citywide schools in Milwaukee can be attributed to the fact that they enroll higher achieving students.[24] The results also suggest that students in the P-5 schools have math score gains equal to those of students in the choice schools and reading score gains that are greater. After four years, the P-5 and choice test score advantage is about 0.37 of a standard deviation for math; the P-5 advantage in reading is about 0.16 of a standard deviation.[25] These gains are relatively large for education productions, and are comparable to the effects from the Tennessee class size experiment (Finn and Achilles 1990; Krueger 1997).

What might explain the fact that the P-5 and choice schools generally outperform the other public schools? While there are undoubtedly many factors that might explain this result, one relatively easily observed characteristic that they have in common is a small pupil-teacher ratio, which is often used as a proxy for class size.[26] Chart 4 shows the average pupil-teacher ratio by school type.[27] The average pupil-teacher ratio in the P-5 schools is 17.0 students per teacher; the average ratio in the choice schools is 15.3.

Chart 4

AVERAGE PUPIL-TEACHER RATIO BY ELEMENTARY SCHOOL TYPE



Notes: Ratios are enrollment-weighted. A P-5 school participates in Milwaukee's Preschool to Grade 5 Grant Program.

Both are significantly smaller than the pupil-teacher ratios in the regular and citywide public schools.

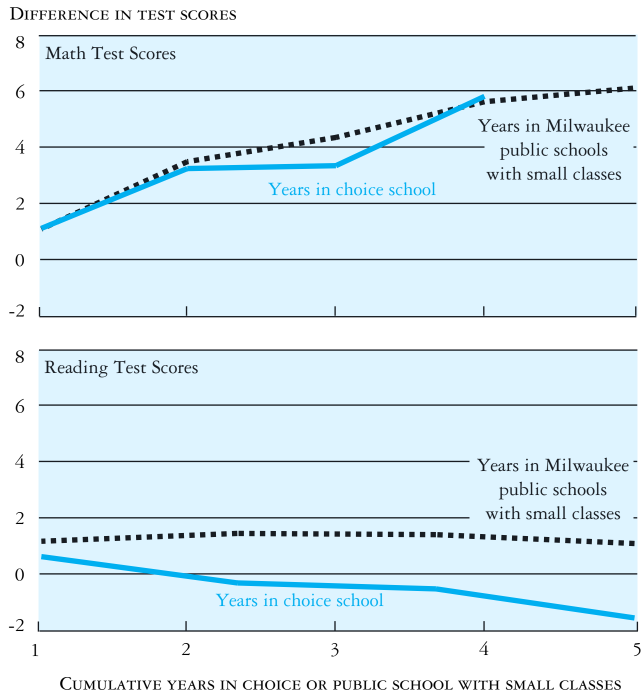
To gauge the extent to which small pupil-teacher ratios might explain the achievement effects of the choice program, I first estimate the effect of the choice schools on test scores relative to all Milwaukee public schools. Next, I estimate the achievement gains that accrue to students enrolled in public schools with low pupil-teacher ratios relative to those in public schools with higher pupil-teacher ratios.[28] This latter analysis uses only students enrolled in the Milwaukee public schools. I then compare the two sets of achievement gains. A finding that the gains among the public schools with low pupil-teacher ratios largely correspond to the gains in the choice schools provides *indirect* evidence that low pupil-teacher ratios (and perhaps small class sizes) may explain part of the observed private-public school achievement differentials (Chart 5).[29]

The solid line in the top panel of Chart 5 shows the math test score growth of students in the choice schools relative to students in all Milwaukee public schools. These results essentially replicate those presented in the top panel of Chart 1. The dotted line shows the math test score progression of students in public schools with small pupil-teacher ratios relative to students in public schools with larger pupil-teacher ratios. The two lines almost entirely coincide. The results for reading are in the bottom panel. In this case, the two lines do not overlap to the same degree as those for math; however, none of the gaps is significantly different from zero, either.

These results indicate that lower pupil-teacher ratios (or class sizes) *may* explain the differential math gains by students in the choice schools (as well as the lack of gains in reading). They do not, however, explain why the P-5 schools appear to perform so well in reading. It is important to understand that this exercise does not *prove* that low pupil-teacher ratios explain either the public-private school or the P-5–regular school achievement difference. Rather, this exercise highlights the need for a much better understanding of why the choice schools in Milwaukee may, on average, be better (at least in teaching mathematics) than the average public school, and why the P-5 schools appear as strong as the choice schools and stronger than the

Chart 5

Can Small Classes Explain the Choice School Achievement Effects?



Notes: The chart depicts the difference in test scores between choice schools and all Milwaukee public schools, and between public schools with small pupil-teacher ratios and public schools with larger pupil-teacher ratios (controlling for ability and family background with individual fixed effects). Milwaukee public schools with small classes are those with a pupil-teacher ratio less than or equal to seventeen students per teacher.

other public schools. I have looked at pupil-teacher ratios because they are a readily available measure that partially defines the P-5 schools and because representatives from the choice schools I contacted emphasized their small class sizes. However, there are likely to be other equally compelling school-specific factors that may explain the differences. Moreover, it is critical that we understand these factors better if we are to improve education for America's urban youth.

Conclusion

The results in this paper suggest that there are significant differences between the public schools in Milwaukee. In particular, students who attend a subset of schools distinguished by, among other characteristics, their small class sizes and additional state funding have test score gains in math that keep pace with those in the private schools that participate in the Milwaukee Parental Choice Program. In addition, this subset of schools has significantly faster reading score gains than either the choice or the other public schools.

In order to evaluate these results fully, one must consider not only student achievement, but costs as well. Evidence that students performed just as well (or better) in the choice schools, but at lower cost, would indicate that private schools are more efficient. Unfortunately, I know of no definitive accounting of the cost differences between the two sectors for providing the same mix of services. Some researchers argue that private schools cost 50 to 60 percent less than public schools (for example, Hoxby [1998]). Coleman and Hoffer (1987) report that, among high schools, overall private school expenditures are 91 percent of public school expenditures. However, both "other, non-Catholic" and "high-performance" private schools spend more than public schools.[30] In addition, Levin (forthcoming) presents an extremely rough estimate of the costs in the Milwaukee public and choice schools. He concludes that the choice schools may have only slightly lower costs (for the same services). Therefore, particularly compared with the P-5 schools, the choice schools may not have an unambiguous efficiency advantage. Clearly, a careful comparison of the educational costs in public and private schools would make an invaluable contribution to the literature and the public policy discussion.

This analysis provides direct evidence that not all public schools are created equal. In addition, not all private schools are created equal. For example, while the overall results suggest that students in the choice schools have no faster gains in reading than do students in the (average) Milwaukee public school, Hispanic students in the choice program—90 percent of whom attend one private school—do make significant gains in reading.[31] If we really want to "fix" our educational system, we need a better understanding of what makes a school successful, and we should not simply assume that market forces explain sectoral differences and are therefore the magic solution for public education.

APPENDIX

Table A1
ELEMENTARY AND MIDDLE SCHOOLS CLASSIFIED AS P-5 AND CITYWIDE

P-5	Citywide
Auer*	Brown
Clarke*	Craig
Franklin*	Elm
Green Bay	Fratney
Holmes	Garfield Avenue
Hopkins*	Grant Avenue
Kagel*	Greenfield
Keefe*	Hawley Road
Kilbourne	Lincoln Center for the Arts
LaFollette*	Lloyd
Lee*	MacDowell
Martin Luther King, Jr.*	Meir Elementary School
Palmer	Milwaukee Education Center
Phillis Wheatley*	Milwaukee French Immersion
Pierce*	Milwaukee German Immersion
Riley	Milwaukee Spanish Immersion
Siefert*	Morgandale
Thirty-first Street (Westside)*	Morse
Thirty-seventh Street	Robinson
Twenty-seventh Street*	Roosevelt
Vieau	Sara Scott
	Starms Discovery
	Thirty-eighth Street
	Thurston Woods
	Tippecanoe
	Townsend Street
	Twenty-first Street
	Urban Waldorf

Note: A P-5 school participates in the Preschool to Grade 5 Grant Program.

*Denotes an original Project Rise School.

Table A2
ORDINARY LEAST SQUARES AND INDIVIDUAL FIXED-EFFECTS ESTIMATES OF THE EFFECT OF CITYWIDE, P-5, AND CHOICE SCHOOLS ON MATH AND READING TEST SCORES

	Dependent Variable			
	Math Scores		Reading Scores	
	Ordinary Least Squares	Fixed-Effects	Ordinary Least Squares	Fixed-Effects
Currently enrolled in citywide school	4.240	0.233	2.565	-1.033
	(1.105)	(1.054)	(1.000)	(0.986)
Enrolled one year	-1.375	0.347	1.471	1.245
	(1.094)	(1.308)	(0.996)	(1.226)
Enrolled two years	-1.037	-0.882	1.062	0.504
	(1.212)	(1.411)	(1.091)	(1.336)
Enrolled three years	2.139	-0.986	3.018	0.121
	(1.401)	(1.533)	(1.264)	(1.438)
Enrolled four years	3.448	-0.114	3.003	0.883
	(1.975)	(1.878)	(1.765)	(1.755)
Enrolled five years	5.620	0.978	1.279	-1.134
	(3.142)	(2.683)	(2.679)	(2.392)
Currently enrolled in P-5 school	2.446	1.810	3.529	0.439
	(0.850)	(0.688)	(0.762)	(0.649)
Enrolled one year	-5.271	-0.234	-5.871	1.441
	(0.741)	(1.743)	(0.670)	(1.623)
Enrolled two years	-1.821	5.067	-3.773	4.342
	(0.885)	(1.777)	(0.798)	(1.653)
Enrolled three years	0.363	6.820	-3.388	3.328
	(1.071)	(1.893)	(0.953)	(1.759)
Enrolled four years	1.271	4.799	-2.483	2.885
	(1.367)	(2.054)	(1.159)	(1.914)
Enrolled five years	3.417	5.180	-0.779	3.361
	(2.014)	(2.372)	(1.746)	(2.222)
Currently enrolled in choice school	0.338	-2.631	0.297	-1.558
	(1.739)	(1.391)	(1.547)	(1.331)
Enrolled one year	-3.683	4.450	-2.428	2.321
	(1.656)	(1.762)	(1.484)	(1.673)
Enrolled two years	-2.999	6.766	-3.853	1.519
	(1.844)	(1.839)	(1.651)	(1.743)
Enrolled three years	-1.592	7.054	-1.139	1.458
	(2.193)	(2.045)	(1.980)	(1.943)
Enrolled four years	1.980	9.721	-2.336	0.549
	(3.113)	(2.560)	(2.797)	(2.421)
Memo:				
Control for individual fixed effects?	No	Yes	No	Yes
R^2	0.057	0.819	0.039	0.795
Number of observations	10,186	10,186	10,224	10,224

Notes: Standard errors are in parentheses. All specifications include a constant and dummy variables indicating the grade level of the student when he or she took the test. The math score regressions include a dummy variable indicating if the test score was imputed. "Enrolled" is the total number of years the student has continuously been enrolled, or had ever been enrolled, in the particular type of school. A P-5 school participates in the Preschool to Grade 5 Grant Program.

Table A3
INDIVIDUAL FIXED-EFFECTS ESTIMATES OF THE EFFECT OF CHOICE SCHOOLS AND PUBLIC SCHOOLS WITH SMALL PUPIL-TEACHER RATIOS ON MATH AND READING TEST SCORES

	Sample of Choice and Public Schools		Sample of Only Public Schools	
	Dependent Variable			
	Math Scores	Reading Scores	Math Scores	Reading Scores
	(1)	(2)	(3)	(4)
Currently enrolled in school with small pupil-teacher ratio			-0.232	-2.383
			(1.111)	(1.022)
Enrolled one year			1.333	3.542
			(1.360)	(1.254)
Enrolled two years			3.725	3.821
			(1.518)	(1.401)
Enrolled three years			4.585	3.778
			(1.727)	(1.593)
Enrolled four years			5.851	3.463
			(1.985)	(1.829)
Enrolled five years			6.332	4.247
			(2.294)	(2.101)
Currently enrolled in choice school	-3.459	-2.312		
	(1.365)	(1.297)		
Enrolled one year	4.584	2.926		
	(1.734)	(1.651)		
Enrolled two years	6.707	1.992		
	(1.813)	(1.713)		
Enrolled three years	6.810	1.781		
	(2.012)	(1.903)		
Enrolled four years	9.269	0.749		
	(2.526)	(2.376)		
Memo:				
R^2	0.816	0.795	0.819	0.803
Number of observations	10,186	10,224	7,171	7,241

Notes: Standard errors are in parentheses. All specifications also include a constant and dummy variables indicating the grade level of the student when he or she took the test, and individual fixed effects. The math score regressions include a dummy variable indicating if the test score was imputed. The regressions in columns (1) and (2) compare the choice schools with all Milwaukee public schools; those in columns (3) and (4) include only the Milwaukee public schools. "Enrolled" is the total number of years the student has continuously been enrolled, or had ever been enrolled, in the particular type of school.

ENDNOTES

The author thanks Alan Krueger and Michele McLaughlin for useful conversations and Howard Fuller for helping her to classify (and understand) the Milwaukee public schools. Michele McLaughlin also provided expert research assistance. Any errors are the author's.

1. For excellent descriptions of the program, see Witte, Thorn, Pritchard, and Claibourn (1994) and Witte, Sterr, and Thorn (1995).

2. As a result, the schools participating in the voucher program are not representative of the typical private school, since only 21 percent of private schools are nonsectarian (U.S. Department of Education 1996). However, until the constitutionality of whether religious schools can participate in voucher programs has been decided, the experience in Milwaukee will be relevant for other cities considering such reforms.

3. I obtained this information by calling the five schools enrolling the largest proportion of choice students. Combined, these schools enroll over 95 percent of the choice students.

4. Originally, the private schools in the choice program were only allowed to admit up to 49 percent of their students as part of the program; this level was raised to 65 percent in 1994. In addition, the number of students who could participate in the choice program was originally limited to 1 percent of the Milwaukee public school enrollment in the first four years but was increased to 1.5 percent in 1994. Given the total enrollment in the Milwaukee public schools, there could be a maximum of only about 1,000 students in the program at any one time.

5. The term control group is generally reserved for randomized experiments, while comparison groups are developed from survey or administrative data.

6. In most other settings, the comparison would show that students in private schools outperform those in public schools.

7. In principle, if one had measures of all the characteristics in which students in the choice schools and students in the public schools differed, one could simply control for these and generate the true effect of the program. The problem, however, is that one is never sure that every characteristic has been controlled for, and indeed we rarely have measures of all (relevant) aspects of the students and their parents. With application lotteries, one does not need these measures.

8. One must control for the application lotteries because applicants to some schools were more likely to be selected than applicants to other schools.

9. There are several places where there could be slippage between the actual lotteries and the imputations. For example, children with siblings who are already enrolled in a choice school are exempted from the lottery, children can apply to more than one school at a time, and the Greene, Peterson, and Du (1997) imputation assumes that a child's race completely determines the school to which he or she applies. In addition, Witte (1997) expresses concern that the choice schools may have abused the permitted exclusions in order to have more control over which students they enrolled.

10. Another way to think about this diagram is that it represents the test score trajectories for both students in the absence of the choice program.

11. The test scores used in this paper are the normal curve equivalent scores of the Iowa Tests of Basic Skills. See Rouse (forthcoming) for more information about the sample.

12. See Rouse (1997) or Rouse (forthcoming) for an elaboration of these points.

13. The fact that the individual fixed-effects strategy can accommodate students missing prior test scores appears to explain a significant portion of the difference in our results.

14. Beginning in 1993, there was no "total math score" (from the Iowa Tests of Basic Skills) for a substantial percentage of students in the Milwaukee public schools. Therefore, I predict (or impute) the total score from the subset of students in the Milwaukee public schools who took the entire battery of math tests (see Rouse [forthcoming] for more details).

15. Others have argued that the observed private school effect is due to the selection process that leads higher achieving students to attend private schools. That is, they argue that the researchers have not controlled for all of the differences between the students in the private schools and the comparison group of students in the public schools. (See, for example, Cain and Goldberger [1983], Cookson [1993], Goldberger and Cain [1982], Murnane [1984], and Witte [1992]).

16. Ideally, I would also disaggregate the achievement gains by the individual choice schools. However, the state of Wisconsin has asked that such an analysis not be undertaken in order to preserve the confidentiality of the choice students.

17. This is my calculation, based on the Common Core of Data for 1991-92.

18. See Table A1 in the appendix for a list of the schools categorized as P-5 and citywide.

19. This is my calculation, based on the Common Core of Data for 1991-92.

20. I estimate the effect of being enrolled in—rather than being selected to attend—the different types of schools because the results in Rouse (forthcoming) suggest that the analyses yield similar results. In addition, estimating the effect of being selected to attend the different types of public schools (and estimating the effect of "years since application") does not make as much sense.

21. The underlying coefficient estimates and standard errors for Charts 2 and 3 are in Table A2 in the appendix.

22. The gap between students in the P-5 and regular public schools becomes statistically significant in the third year. The gap for the choice schools is not statistically significant.

23. These results differ from those reported by Archbald (1995), who found that students in the Milwaukee magnet schools scored higher on math and reading tests than those enrolled in the attendance area schools.

24. Not all citywide schools perform equally. In particular, when these schools are divided into "gifted," "language immersion," "special program" (such as Waldorf, Montessori, or Global Learning), and "other"—and individual fixed effects are included—the students in the language immersion schools have substantially faster gains in reading than students in all other types of schools, and students in the gifted schools have significantly slower gains in mathematics than students in the regular schools. Because in some years the number of students in some of these school categories is small, these results should be regarded as tentative.

25. I used the within-sample standard deviation of 19 for this calculation. Nationally, the standard deviation for normal curve equivalent scores is 21.

26. Although highly correlated, the pupil-teacher ratio does not always directly correspond to the average class size. Rather, the two measures diverge as intraschool variation in class size increases due, for example, to special and compensatory education (Boozer and Rouse 1997). To illustrate, the average pupil-teacher ratio in the choice schools is 15.3 students per teacher; however, the schools' average class size is 23.6 students. Unfortunately, data on average class size for the Milwaukee public schools were not readily available.

27. The estimates of the pupil-teacher ratios for the choice schools are based on the schools I contacted. I estimated the pupil-teacher ratios for the public schools using the Common Core of Data for 1991-92.

28. Schools with low pupil-teacher ratios have ratios less than or equal to seventeen to one. I chose seventeen because it is the maximum pupil-teacher ratio in the choice schools I contacted. According to this criterion, 43 percent of all Milwaukee public schools and 52 percent of the P-5 schools are considered to have low pupil-teacher ratios.

29. See Table A3 in the appendix for the estimated coefficients and standard errors.

30. Other, non-Catholic private schools spend 38 percent more than public schools, while high-performance private schools spend 131 percent more than public schools. Expenditures in Catholic private schools, however, are lower than those in public schools. The fact that Catholic school costs differ from those in other types of private schools may reflect lower teacher salaries and greater in-kind subsidies (including facilities) from the Catholic church.

31. These results are not reported here but are available from the author on request.

References

Archbald, Doug. 1995. "A Longitudinal Cohort Analysis of Achievement Among Elementary-Magnet Students, Neighborhood-School Students, and Transfer Students." JOURNAL OF RESEARCH AND DEVELOPMENT IN EDUCATION 28, no. 3 (spring): 161-8.

Blank, Rolf K. 1990. "Analyzing Educational Effects of Magnet Schools Using Local District Data." SOCIOLOGICAL PRACTICE REVIEW 1, no. 1: 40-51.

Boozer, Michael A., and Cecilia Elena Rouse. 1997. "Intra-School Variation in Class Size: Patterns and Implications." Revision of NBER Working Paper no. 5144 (June 1995).

Bryk, Anthony S., Valerie E. Lee, and Peter B. Holland. 1993. CATHOLIC SCHOOLS AND THE COMMON GOOD. Cambridge: Harvard University Press.

Cain, Glen G., and Arthur S. Goldberger. 1983. "Public and Private Schools Revisited." SOCIOLOGY OF EDUCATION 56 (October): 208-18.

Clancy, Dan, Charlie Toulmin, and Merry Bukolt. 1995. "Wisconsin Public School Finance Programs, 1993-94." In Steven D. Gold, David M. Smith, and Stephen B. Lawton, eds., PUBLIC SCHOOL FINANCE PROGRAMS OF THE UNITED STATES AND CANADA 1993-94. New York: American Education Finance Association.

Coleman, James S., and Thomas Hoffer. 1987. PUBLIC AND PRIVATE HIGH SCHOOLS: THE IMPACT OF COMMUNITIES. New York: Basic Books.

Coleman, James S., Thomas Hoffer, and Sally Kilgore. 1982a. HIGH SCHOOL ACHIEVEMENT: PUBLIC, CATHOLIC, AND PRIVATE SCHOOLS COMPARED. New York: Basic Books.

————. 1982b. "Cognitive Outcomes in Public and Private Schools." SOCIOLOGY OF EDUCATION 55, no. 2-3 (April-July): 65-76.

Cookson, Peter W. 1993. "Assessing Private School Effects: Implications for School Choice." In Edith Rasell and Richard Rothstein, eds., SCHOOL CHOICE: EXAMINING THE EVIDENCE. Washington, D.C.: Economic Policy Institute.

Crain, Robert L., Amy L. Heebner, and Yiu-Pong Si. 1992. "The Effectiveness of New York City's Career Magnet Schools: An Evaluation of Ninth Grade Performance Using an Experimental Design." National Center for Research in Vocational Education working paper.

Evans, William N., and Robert M. Schwab. 1995. "Finishing High School and Starting College: Do Catholic Schools Make a Difference?" QUARTERLY JOURNAL OF ECONOMICS 110 (November): 941-74.

Finn, Jeremy D., and Charles M. Achilles. 1990. "Answers and Questions About Class Size: A Statewide Experiment." AMERICAN EDUCATIONAL RESEARCH JOURNAL 27, no. 3 (fall): 557-77.

Gamoran, Adam. 1996. "Student Achievement in Public Magnet, Public Comprehensive, and Private City High Schools." EDUCATIONAL EVALUATION AND POLICY ANALYSIS 18, no. 1 (spring): 1-18.

Goldberger, Arthur S., and Glen G. Cain. 1982. "The Causal Analysis of Cognitive Outcomes in the Coleman, Hoffer and Kilgore Report." SOCIOLOGY OF EDUCATION 55, no. 2-3 (April-July): 103-22.

Greene, Jay P., Paul E. Peterson, and Jiangtao Du. 1997. "The Effectiveness of School Choice: The Milwaukee Experiment." Harvard University Education Policy and Governance Occasional Paper no. 97-1, March.

Hoxby, Caroline M. 1998. "What Do America's 'Traditional' Forms of School Choice Teach Us about School Choice Reforms?" Federal Reserve Bank of New York ECONOMIC POLICY REVIEW 4, no. 1.

Krueger, Alan B. 1997. "Experimental Estimates of Education Production Functions." Princeton University Industrial Relations Section Working Paper no. 379, May.

Levin, Henry M. Forthcoming. "Educational Vouchers: Effectiveness, Choice, and Costs." JOURNAL OF POLICY ANALYSIS AND MANAGEMENT.

Milwaukee Public Schools. 1997. DIRECTIONS: YOUR SCHOOL SELECTION GUIDE FOR THE 1997-98 SCHOOL YEAR.

Molnar, Alex. 1996. GIVING KIDS THE BUSINESS: THE COMMERCIALIZATION OF AMERICA'S SCHOOLS. Boulder, Colo.: Westview Press.

Murnane, Richard J. 1984. "A Review Essay—Comparisons of Public and Private Schools: Lessons from the Uproar." JOURNAL OF HUMAN RESOURCES 19: 263-77.

Neal, Derek. 1997. "The Effect of Catholic Secondary Schooling on Educational Achievement." JOURNAL OF LABOR ECONOMICS 15, no. 1, part 1 (January): 98-123.

Rouse, Cecilia Elena. 1997. "Lessons from the Milwaukee Parental Choice Program." POLICY OPTIONS 18, no. 6 (July-August): 43-6.

————. Forthcoming. "Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program." QUARTERLY JOURNAL OF ECONOMICS.

Sander, William. 1996. "Catholic Grade Schools and Academic Achievement." JOURNAL OF HUMAN RESOURCES 31 (summer): 540-8.

U.S. Department of Education. National Center for Education Statistics. 1996. DIGEST OF EDUCATION STATISTICS, 1996. NCES 96-133, by Thomas D. Snyder, Charlene M. Hoffman, and Claire M. Geddes. Washington, D.C.

Witte, John F. 1992. "Private School vs. Public School Achievement: Are There Findings That Should Affect the Educational Choice Debate?" ECONOMICS OF EDUCATION REVIEW 11 (December): 371-94.

———. 1997. "Achievement Effects of the Milwaukee Voucher Program." Unpublished paper, University of Wisconsin, January.

Witte, John F., Troy D. Sterr, and Christopher A. Thorn. 1995. "Fifth-Year Report: Milwaukee Parental Choice Program." Unpublished paper, University of Wisconsin, December.

Witte, John F., Christopher A. Thorn, Kim M. Pritchard, and Michele Claibourn. 1994. "Fourth-Year Report: Milwaukee Parental Choice Program." Unpublished paper, University of Wisconsin, December.