NBER WORKING PAPER SERIES

PRIVATE SCHOOL VOUCHERS AND STUDENT ACHIEVEMENT: AN EVALUATION OF THE MILWAUKEE PARENTAL CHOICE PROGRAM

Cecilia Elena Rouse

Working Paper 5964

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 1997

I thank Orley Ashenfelter, David Card, Hank Farber, and Alan Krueger for insightful suggestions and conversations, and participants at the National Bureau of Economic Research's Program on Children Conference and the Princeton Labor Lunch for helpful comments. I particularly thank Lisa Boeger, Jay Greene, and John Witte for insights about the program and lots of help with the data. Jeff Wilder provided expert research assistance. All errors are mine. This paper is part of NBER's research programs on Children, Labor Studies and Public Economics. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

Private School Vouchers and Student Achievement:
An Evaluation of the Milwaukee Parental Choice Program
Cecilia Elena Rouse
NBER Working Paper No. 5964
March 1997
JEL No. I20
Children, Labor Studies, Public Economics

ABSTRACT

In 1990, Wisconsin became the first state in the country to provide vouchers to low income students to attend non-sectarian private schools. In this paper, I use a variety of estimation strategies and samples to estimate the effect of the program on math and reading scores. First, since schools selected students randomly from among their applicants if the school was oversubscribed, I compare the academic achievement of students who were selected to those who were not selected. Second, I present instrumental variables estimates of the effectiveness of private schools (relative to public schools) using the initial selection as an instrumental variable for attendance at a private school. Finally, I used a fixed-effects strategy to compare students enrolled in the private schools to a sample of students from the Milwaukee public schools. I find that the Milwaukee Parental Choice Program appears to have had a positive effect on the math achievement of those who attended a private school; but had no benefits for reading scores. I have found the results to be fairly robust to data imputations and sample attrition, however these limitations should be kept in mind when interpreting the results.

Cecilia Elena Rouse
Industrial Relations Section
Firestone Library
Princeton University
Princeton, NJ 08544
and NBER
rouse@princeton.edu

At the cornerstone of many school reform proposals lies the premise that private schools are more efficient than public schools. This premise is particularly prominent in the current "school choice" debate. Proponents of school choice argue that governments should offer tuition vouchers to families who wish to send their children to private, rather than public schools. If private schools are indeed more effective than public schools, a voucher program may offer a cost-effective way to improve the quality of education. The key evidence of more effective private schooling is a study by Coleman, Hoffer, and Kilgore (1982a,b) who used the first year of the High School and Beyond data to show that students in private schools have higher achievement levels than those in public schools. Critics of school choice programs have argued that private schools would not necessarily do a better job educating students who are currently attending public schools. Rather, they argue that the observed superiority of private school students arises from the selection process that leads higher-achieving students to attend private schools (Goldberger and Cain (1982))[1].

Ideally, the issue of the relative effectiveness of private versus public schooling could be addressed by a social experiment in which children in a well-defined universe were randomly assigned to a private school (the "treatment group"), while others were assigned to attend public schools (the "control group"). After some period of time, one could compare outcomes, such as test scores, high school graduation rates, or labor market success between the treatment and control groups. Since, on average, the only differences between the groups would be their initial assignment -- which was randomly determined -- any differences in outcomes could be attributable to the type of school attended.[2] While such an experiment has never been implemented, the legislative

[1] See, as well, the collection of articles in the Sociology of Education (1982) and the Harvard Education Review (1981), as well as the review pieces by Cookson (1993), Murnane (1984), and Witte (1992).

[2] Short of an experiment, one could also look for an exogenous event that is correlated with the likelihood that an individual attends a private school but is uncorrelated with the error term in the outcome equation. Evans and Schwab (1995) conduct such an analysis on the relative value of Catholic high schools using whether the individual is Catholic as the exogenous event. They conclude that students who attend Catholic schools are more likely to finish high school and to enter a four-year college than students who attend public schools.

requirements of a recently enacted school voucher program in Milwaukee, Wisconsin allow one to come close to such an idealized experiment.

In 1990, Wisconsin became the first state in the country to implement a school choice program that would provide vouchers to low income students to attend non-sectarian private schools.[3] The number of students in any year was limited to 1% of the Milwaukee public school membership in the first four years, but was expanded to 1.5% in 1994. Only students whose family income was at or below 1.75 times the national poverty line were eligible to apply. In principle, a family of three with an income of approximately $21,000 would be eligible to apply to the program; in practice, as shown in Table 1, the mean family income of applicants was approximately $12,300. The choice students were considerably more disadvantaged than the average student in the Milwaukee public schools (whose average family income was $24,000), and the average non-choice private school student in Milwaukee (whose average family income was about $43,000; Witte, Thorn, and Pritchard (1995)). The means in Table 1 also show that choice applicants were more likely to be minority and had lower math and reading test scores than the average student in the Milwaukee public schools. These test scores are nationally normed, suggesting that the students who applied to choice were scoring considerably lower than the national average, as well. On the other hand, the parental education for choice applicants, at least for those responding to the surveys, was comparable (or even a little higher) than non-applicants from the Milwaukee public schools.

As the program limited participation to independent, secular private schools, the participating schools were not representative of all private schools in Milwaukee, as the majority of private school enrollments were likely in religious schools (Witte, Thorn, and Pritchard (1995)). That said, until the constitutionality of whether religious schools can participate in voucher programs has been decided,

[3] There was an attempt to include religious schools in the program, however the Wisconsin supreme court ruled that it would violate the separation between church and state. I only briefly describe the program here. For more details, see Witte, Thorn, Pritchard, and Claibourn (1994) and Witte, Sterr, and Thorn (1995). I use the terms "private schools" and "choice schools" interchangeably.

Table 1

Mean Characteristics of Applicants to the Choice Program and Students in the Milwaukee Public Schools

	Applicants		Milwaukee Public School Sample
	Selected	Not Selected	
African-American	0.721 (0.011)	0.780 (0.012)	0.550 (0.007)
Hispanic	0.217 (0.011)	0.154 (0.011)	0.097 (0.004)
"First" Math Score (National Percentile Ranking)	34.560 (1.196)	35.073 (1.552)	43.353 (0.494)
"First" Math Test Score, Missing	0.704 (0.012)	0.779 (0.012)	0.382 (0.007)
"First" Reading Score (National Percentile Ranking)	33.869 (1.050)	34.341 (1.369)	40.227 (0.462)
"First" Reading Score, Missing	0.699 (0.012)	0.781 (0.012)	0.382 (0.007)
Family Income (÷ 1000)	$12.123 (0.297)	$12.715 (0.509)	$23.897 (0.534)
Family Income, Missing	0.491 (0.013)	0.760 (0.012)	0.727 (0.006)
Mother's Education	12.559 (0.067)	12.399 (0.109)	12.108 (0.055)
Mother's Education, Missing	0.555 (0.013)	0.763 (0.012)	0.724 (0.006)
Father's Education	12.048 (0.092)	11.521 (0.152)	12.250 (0.068)
Father's Education, Missing	0.697 (0.012)	0.846 (0.010)	0.796 (0.005)
Grade of Application	2.997 (0.079)	3.875 (0.119)	3.477 (0.058)
Has Test Score Two Years After Application	0.536 (0.015)	0.284 (0.019)	0.535 (0.009)
Has Test Score Three Years After Application	0.467 (0.017)	0.236 (0.024)	0.405 (0.009)
Has Test Score Four Years After Application	0.407 (0.025)	0.183 (0.027)	0.364 (0.011)
(Maximum) Number of Observations	1544	1219	5318

Notes: Standard errors are in parentheses. For applicants to the choice schools, the "first" test scores are the "pre-application test scores" (the test scores from the year of application) or earlier ones if the pre-application test scores are missing; for students in the Milwaukee public schools, the "first" test scores are the first non-missing (post-1989) test scores. See text or Data Appendix for how I construct a "year of application" for the Milwaukee public school sample. The data in this table do not represent my analysis sample.

the experience in Milwaukee will be relevant for other cities considering such reforms. In the first year, 7 private schools participated; by 1995, this number had risen to 12. And, the schools represented a variety of educational approaches including Montessori and Waldorf schools, and schools with bilingual and African-American cultural emphases. These schools were not elite private schools. For example, the voucher was worth approximately $3,200 in 1994-1995 which contrasts with a range of tuition and fees for schools participating in the choice program of $1,080-$4,000 (in 1993-1994). In fact, the voucher program helped to improve the financial status of several of the participating schools. While one could believe, a priori, that a program providing vouchers to elite private schools, such as Exeter and Andover, would result in higher achievement gains, it is not so clear that a program providing vouchers to local non-sectarian private schools (that are willing to participate) would have such effects.[4]

Finally, choice schools were only allowed to admit up to 49% of their students as part of the choice program, and were not allowed to discriminate in which students they took. This was interpreted to mean that if the school was oversubscribed, the students would be randomly selected from among the applicants. If a choice school was not oversubscribed, it was required to take all who applied, with only a few exclusions.

It is this feature of the Milwaukee Parental Choice Program that I primarily analyze in this paper. By using applicants who are not selected as a control group, I take advantage of this quasi-natural experiment to estimate whether private school vouchers could improve school achievement of participating students in the short-run, and whether non-sectarian private schools are, indeed,

---

[4] On the other hand, in Table 1, the sample from the Milwaukee public schools scored, on average, at the 43-45 percentile which suggests that the Milwaukee public schools are performing below the national average. Coleman, Hoffer, and Kilgore (1982a,b) find that other-private schools perform at least as well as "Catholic schools".

better than public schools.[5] I also use a fixed-effects estimator to compare the achievement effects of students in the private schools to a random sample of students in the Milwaukee public schools.

I am not the first to analyze the Milwaukee Parental Choice Program. The original evaluation of the fourth year of the program, conducted by Witte, Sterr, and Thorn (1995), compared the test scores of students in the choice schools to a random sample of all Milwaukee public school students and to low-income students (comparison groups) and concluded there were no significant relative achievement gains among the choice students. A subsequent analysis by Greene, Peterson, Du, Boeger, and Frazier (1996) criticized the Witte, Sterr, and Thorn study by noting that students in their comparison group were from substantially more advantaged family backgrounds than students in the choice program. Although Witte, Sterr, and Thorn argue that their comparison group is, in fact, comparable to choice students, conditional on observable covariates, it is nevertheless possible that unobserved factors remain which would tend to obscure any relative achievement gains among the choice student population. As an alternative to the use of a general comparison group, Greene, et. al. propose the use of the unsuccessful applicants as a "quasi-experimental" control group. Relative to the unsuccessful applicants, Greene, et. al. conclude that the choice students made statistically significant test score gains by their third and fourth years in the program. However, the analysis by Greene, et. al. makes a potentially important assumption by excluding from the choice group students who were successfully admitted to the choice schools, but did not attend them, or only attended for a short period of time.

I begin my analysis by returning to the "true" source of exogenous variation in the Milwaukee

---

[5] Because the choice program was so small, it cannot provide an analysis of the potential benefits of "school choice". In the most unrestricted school choice program, all (or a substantial fraction) of the students in the public schools would be eligible to attend a private school. Since state funding would be tied to student enrollments, the public schools would have an incentive to improve, leading to no differences in the outcomes of students in "public" and "private" schools. Analysis of the Milwaukee Parental Choice Program can, however, indicate whether parents would prefer to send their children to a private school (see Witte, Sterr, and Thorn (1995)), and whether, in the short run, the academic achievement of those children that attend the private schools would likely increase.

Choice program -- that of selection into the pool of students eligible to attend a choice school. I argue that a complete evaluation of the program's effects on student achievement should consider not only whether private schools are better than public schools (as do Greene, et. al (1996) and Witte, Sterr, and Thorn (1995)), but also whether students who are selected to attend a choice school will actually enroll and remain there. I do so by estimating the effect of being selected to attend a private school on student achievement. I next present instrumental variables estimates of the effectiveness of private schools using the initial selection as an instrumental variable for attendance at a private school; I also present fixed-effects estimates of the private school effect using both the unsuccessful applicants and the sample of students in the Milwaukee public schools as comparison groups. I find that students selected for the choice program scored approximately 1-2 extra percentage points per year in math compared to the control group, and those who were enrolled in a private school scored approximately 1.5-2 percentage points more per year those who were in the Milwaukee public schools. I do not observe differences in the reading scores. I have found the results to be fairly robust to data imputations and sample attrition, however these limitations should be kept in mind when interpreting the results.

In the next section I lay out the empirical framework, in section three I describe my sample, I present the results in section four, and conclude in section five.

II. Empirical Strategy

A. Evaluating the Choice Program vs. Establishing Whether Private Schools are Better

In an "ideal" experiment the random assignment process completely determines the status of individuals in the treatment and control groups. In many experimental settings involving human subjects, however, there is slippage between the random assignment status of experimental subjects and whether or not they actually receive the treatment. For example, in randomized trials of medicines, some individuals in the treatment group may fail to take the prescribed treatment (see, for example, Coronary Drug Project (1980) and Efron and Feldman (1991)). In some social experiments, the slippage can occur in both directions. For example, in a job training experiment, some people assigned to the treatment group may fail to show up, whereas others in the control group may receive training from other providers (Hausman and Wise (1985) and Heckman and Smith (1994)).[6] In the specific context of the Milwaukee Choice program, the slippage between random assignment status -- whether the individual was chosen to attend a choice school -- and "treatment status" -- whether an individual actually attended a choice school -- is significant because the treatment lasted over several years. In fact, only about 50 percent of students who were randomly selected to attend choice schools in the first year of the program were still attending the school two years later. In order to understand the nature of alternative estimates of the relative effect of private schooling in this setting, it is necessary to develop a model that takes account of the link between the random selection process and the decision to attend a choice school.

Consider the following linear probability model that models child i's actual attendance at a choice school in year t-1, P_{it-1} ($P_{it-1} = 1$ for students who attend a choice school; $P_{it-1} = 0$ for those who do not), as a function of i's initial selection status, S_i ($S_i = 1$ for students who were selected, $S_i = 0$ for those who were not):

$$P_{it-1} = a + \rho S_i + X_i g + Z_i g' + v_{it-1}\,. \qquad (1)$$

Here, X_i and Z_i are vectors of individual characteristics (to be distinguished later). Note that if individuals were forced to attend the school to which they were assigned, then $\rho=1$ and $a=g=g'=0$. More generally, however, students' attendance decisions depend on a variety of factors, such as their family's residential location, whether the school is attended by friends or siblings, and so forth. In

[6] Another example is that of the Vietnam draft lottery. Angrist (1991) explains racial differences in the effect of the lottery on enlistment by the fact that the effect of the draft lottery on enlistment was greater for Whites than for non-Whites.

this case, the parameter ρ may be smaller than 1.

Consider next an outcome equation for the test score of student i in year t. To fix ideas, suppose that test scores depend on the characteristics of i and on her attendance at a choice school in the previous year. (I discuss the specification of the outcome equation more fully in the next section.) Specifically, assume that the test score of student i in year t (T_{it}) is determined by:

$$T_{it} = \alpha + \beta P_{it-1} + X_i\gamma + Z_i\Gamma + \varepsilon_{it} \,. \tag{2}$$

In equation (2), the parameter β reflects the relative effectiveness of choice schools over the public schools attended by students in the sample.

Finally, it is useful to combine equations (1) and (2) into a reduced-form equation:

$$T_{it} = \pi_0 + \pi_1 S_i + X_i\Pi_2 + Z_i\Pi_3 + \nu_{it} \,. \tag{3}$$

Note that $\pi_1=\rho\beta$. Thus, the reduced-form effect of selection into the group eligible to attend choice schools is a combination of two effects: the effect of selection on the relative likelihood of attending a choice school (ρ) -- what might be called a "takeup effect", and the "true" effect of choice schools on student achievement (β). In the treatment literature (e.g. Rubin (1974), Efron and Feldman (1991)), the parameter π_1 is referred to as the "intent to treat" effect.

There are at least two reasons why we might be interested in π_1 rather than the constituent parameters ρ and β. First, it is the only policy instrument available to policy makers. 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. For example, see Table 2. Of the approximately 300 students who applied in 1990 and were selected to attend a choice school, 98% ever enrolled in a choice school, but only 76% remained through the first spring semester.[7] By the

[7] I can only observe if a student was still enrolled in the choice program in the spring if he took an achievement test. If a student was absent the day of the test, this measure has the potential to inflate the attrition from the program. As a result, for this table, if a student is observed to be enrolled the following year, I also assume he is enrolled in the current year (thus, for example, if a student is observed to be enrolled in the spring of 1992 and he applied in 1990, I assume he was

Table 2

Numbers of Applicants, Selections, and Enrollments

	Year of "First" Application			
	1990	1991	1992	1993
Number of Applicants	408	375	314	232
Number Selected	301	310	182	161
# Ever Enrolled in Private School in Fall	294	281	169	136
# Ever Enrolled in Private School in Spring	228	255	164	110
Enrolled in Spring of:				
1991	228			
1992	175	255		
1993	125	180	162	
1994	86	125	128	110

Notes: The year of "first" application is the first year that a student ever applied if they were never selected, and the first year that a student applied and was selected if they applied more than once. "Enrolled" means that a student was enrolled in a private school. Enrolled in the fall means that they were enrolled as of October of the year; enrolled in the spring means they took the Iowa Tests of Basic Skills in a private school indicating that they were still enrolled in the private school in the spring. The numbers in this table are based on African-American and Hispanic students who applied in grades K-8 between 1990 and 1993 (i.e., my analysis sample).

second spring, approximately 58% remained. There are similar patterns in later cohorts. Clearly, ρ can be substantially less than one. In the extreme case in which $\rho=0$, even if private schools are much better at educating children than the public schools, there will be no achievement gains from the program. Thus, the reduced form estimates reflect the overall potential gains from offering the vouchers. Secondly, as in many experimental settings, the randomization only occurred in the intent to treat and as such, the reduced-form estimate is the only unambiguously unbiased estimate that one can obtain from an ordinary least squares (OLS) regression, assuming the initial selection was truly random.

However, the reduced-form estimate is rather unsatisfying for those parents and policy makers who desire an estimate of the benefits that are likely to accrue to those who actually enroll and remain in a private school for a specified period of time (the effect of "treatment on the treated"). That is, it does not establish whether private schools are better than public schools precisely because not all who are assigned to attend a private school do so. Unfortunately, it is also more difficult to estimate whether private schools are better than public schools using the Milwaukee data. If, conditional on X_i and Z_i, take-up and attrition from the program is orthogonal to the error term in equation (2), the estimate of β in equation (2) by OLS will generate an unbiased estimate of the effectiveness of private schools relative to public schools. Greene, et. al. (1996) make this assumption when they exclude from their analysis children who were selected but did not enroll in a private school, and those who left the choice program and re-enrolled in a Milwaukee public school. However, in practice it is unlikely that students who do not take up the program or who leave the program are a random sample of students who applied to the program. In particular, students who are not being well served by the private schools may be more likely to leave, and those who do not

enrolled in the spring of 1991 as well.)

take up the program may not expect to be well served.[8]

I can, however, use the initial selection as an instrumental variable that is correlated with P_{it-1} and uncorrelated with the error term in the education production function. The instrumental variables estimate of β will provide a consistent estimate of a "causal" effect of attending private school for a period of time on test scores (Angrist, Imbens, and Rubin (forthcoming)). However, not all students who start at a private school remain there, rather some receive partial treatment; and I do not have a second instrumental variable that is correlated with students leaving the choice program (and not correlated with the error in equation (2)).[9] However, I can define the treatment in many ways. For example, I can define the treatment as having attended private school for 1, 2, 3, or 4 years and compare the effect to those who have not attended for as long. For, no matter how one defines P_{it-1} (the treatment), the initial selection is still random.

However, with these definitions of treatment, one must interpret both the OLS and IV estimates of equation (2) with caution. Suppose that the true education production function were,

$$T_{it} = \delta_0 + \beta_0 P_{it-1} + \psi D_{it-1} + \omega_{it} \tag{4}$$

where, as in equation (2), T_{it} represents the test score of student i in year t, P_{it-1} is whether the student was enrolled in a private school, D_{it-1} indicates if a student had previously been enrolled in a private school in the choice program but is now enrolled in a Milwaukee public school (i.e. the student received partial treatment), β_0 represents the achievement gains of students who attended private school relative to students who never attended a private school, ψ is the partial treatment effect of

[8] Take up was particularly a problem in the first year of the Milwaukee Voucher Program as it was not clear that the program would be declared constitutional. For evidence that there may have been negative selection in the take up in the first year, see Appendix Table 1a.

[9] I have tried using the miles to the nearest choice school as a second instrumental variable that is correlated with leaving the choice program (and, hopefully, uncorrelated with the error term in the education production function), however the estimated standard errors were so large as to render the exercise uninformative.

private school for those who left the program, and ω_{it} is an error term. (I have excluded the vectors, X_i and Z_i for convenience.) In this case, the OLS estimate of β in equation (2) is likely to be an underestimate of the effect of private school relative to students who have <u>never</u> attended a private school. The reason is that the base group will include students who received partial treatment. That is,

$$\hat{\beta}_{OLS} = \beta_0 + \phi - \Psi\rho' \tag{5}$$

where φ is the selection bias in the OLS estimate of β, and ρ' is the proportion of students in the Milwaukee public schools who attended a choice school for some period of time. Except in the cases where there are no students who receive partial treatment or there is no effect of receiving partial treatment (ρ'=0 or ψ=0), or the selection effect is particularly large, the OLS estimate will be too small, i.e., $\hat{\beta}_{OLS}<\beta_0$.

The opposite problem occurs with the IV estimate of β, as I do not have a second instrumental variable to identify a separate effect for those students who leave the choice program. To see this, consider the point made above that the reduced-form estimate of the "intent to treat" is a function of take-up and the causal effect of private schools on achievement scores, i.e., $\pi_1=\rho\beta$, or $\beta=(\pi_1/\rho)=(\delta T_{it}/\delta S_{it-1})/(\delta P_{it-1}/\delta S_{it-1})$. This estimate is known as the Wald estimator when the instrumental variable is binary and is equivalent to the IV estimator.[10] To consider the potential bias in the IV estimate[11], re-write the true education production function (equation (4)) as,

$$T_{it} = \delta_0 + \beta_0\rho S_{it-1} + \psi(1-\rho)S_{it-1} + \omega_{it} \tag{4'}$$

where S_i is whether the individual had been randomly selected to participate in the choice program,

[10] See Wald (1940) and Durbin (1954).

[11] See Gruber (1996) for a similar derivation.

and ρ is the proportion of students selected who are actually enrolled in a private school. Further, consider a modified version of equation (1),

$$P_{it-1} = a + \rho S_{it-1} + v_{it-1} \tag{1'}$$

where, again, the vectors X_i and Z_i have been suppressed. From these equations,

$$\hat{\beta}_{IV} = \frac{\dfrac{\delta T_{it}}{\delta S_{it-1}}}{\dfrac{\delta P_{it-1}}{\delta S_{it-1}}} = \beta_0 + \psi\Theta \tag{8}$$

assuming that $\delta\rho/\delta S_i=0$ (i.e. that the marginal take-up is zero) and letting $\Theta=(1-\rho)/\rho$.[12] If $\psi = 0$ (those who receive partial treatment receive no additional benefits), then $\hat{\beta}_{IV} = \beta_0$. In general, however, as long as those who receive partial treatment receive some benefit (and $\rho<1$), the IV estimate will be upward biased, i.e., $\hat{\beta}_{IV} > \beta_0$. Combining the biases in the OLS and IV estimators, we can generally expect that $\hat{\beta}_{OLS} < \beta_0 < \hat{\beta}_{IV}$ if private schools generate achievement gains for those who attend for some period of time.

Finally, evaluations using experimental data are not panaceas. One problem with the model outlined above is that I must assume a constant treatment effect across the population in order for the IV estimate to provide a consistent estimate of the common treatment effect. If, however, the effectiveness of private schools varies across the population (there are heterogeneous treatment effects), then the IV estimate may not provide a consistent estimate of even the average treatment effect. Unfortunately, to relax the assumption of a constant treatment effect would require larger samples than available with these data (Heckman (1995a,b), Manski (1992,1995)). A second problem is that randomization must occur at some point in the experiment and because in this program

[12] I estimate a similar range of estimates of the IV bias when I let $\delta\rho/\delta S_i\neq0$. In this case, $\Theta=[(1-\rho)-S_i(\delta\rho/\delta S_i)]/[\rho+S_i(\delta\rho/\delta S_i)]$.

randomization was based on an applicant pool, both the causal and reduced-form estimates reflect the effect of the program or the effect of private schools relative to public schools among students interested in attending a private school. It is not the treatment effect that one would estimate for the general population (Angrist, Imbens, and Rubin (forthcoming), Heckman and Smith (1993)). To assess how sensitive the results are to using the not-selected applicants and to estimate an effect of private schools that would come closer to a treatment effect for the entire at-risk population, I conclude by estimating OLS and fixed-effects estimates using the sample of students from the Milwaukee public schools.

B. *Modelling Educational Achievement*

A second issue to consider is the proper form of the education production function. In the most basic equation, students' outcomes are modeled as a function of schooling inputs, such as whether the student attended a private school, as in equation (2). By only including one dummy variable indicating attendance at a private school, such a model implies that private schools generate a one time increase in student achievement but that the rate of learning is similar between public and private schools. A more sophisticated model estimates a change in student achievement as a function of schooling inputs (a value-added specification) or includes individual fixed-effects to compensate for a lack of previous schooling inputs and to control for any unobserved differences between students in private and public schools (Hanushek (1979)). This specification suggests that private schools may change the slope of the equation, but because most data have only one or two years of schooling information, one cannot separately identify whether the private schools have changed the slope of the education function, the intercept, or both.

However, the longitudinal structure of the Milwaukee data allows me to estimate richer and less restrictive models. In particular, I can estimate the effects of having been in a private school for some period of time as it is quite possible that a child's achievement does not improve immediately as she may need some time to adjust to her new environment.[13] In the basic model, I simply include dummy variables indicating how long the student has attended a private school, that is,

$$T_{it} = \alpha + \sum_{k=1}^{4} \beta_k P_{it-k} + X_i\gamma + Z_t\Gamma + \varepsilon_{it} \tag{9}$$

where P_{it-k} are dummy variables indicating the number of years at private school[14], and the vector β_k reflects the impact of previous attendance at private school on current test scores. For example, a pattern in which $\beta_k > \beta_{k-1}$ would suggest that the impact of private schools rises over time, i.e., private schools increase the rate at which children learn. I also constrain the year-to-year gains to be equal, and estimate the linear effect of years in the program on student achievement. Finally, I estimate a spline function that allows me to test for an intercept shift, a change in the slope, and discontinuities in the slope,

$$\begin{aligned} T_{it} = {} & \alpha' + \alpha_p' P_{it-1} + \delta' Y_{it-1} + \beta' Y_{it-1} P_{it-1} \\ & + \delta_2' d_2 (Y_{it-1} - 2) + \beta_2' d_2 (Y_{it-1} - 2) P_{it-1} \\ & + \delta_3' d_3 (Y_{it-1} - 3) + \beta_3' d_3 (Y_{it-1} - 3) P_{it-1} \\ & + X_i \gamma' + Z_t \Gamma' + \varepsilon_{it}' \end{aligned} \tag{10}$$

[13] This is the argument advanced by Greene, et. al. (1996). Also, according to John Witte, the private schools in the choice program reported they were more concerned about getting the children interested in learning in the first couple of years than in increasing their test scores. Others have interacted school resources with schooling to estimate the effect of school quality on the slope of the production function. See, for example, Akin and Garfinkel (1980), and Behrman and Birdsall (1983).

[14] I use the notation of P_{it-1} suggesting enrollment in a private school, but I also estimate these specifications for the reduced-form equations.

where Y_{it-1} is the number of years since the student applied to the choice program[15], $d_2 = 1$ if the student has been in the choice program for at least two years, and $d_3 = 1$ if a student has been in the choice program at least three years.

III. Sample

The samples I analyze in this paper are drawn from the Milwaukee Parental Choice Program public release data files (Witte and Thorn (1995)). These files contain administrative data on information such as test scores, race, sex, grade level, whether the student applied and was accepted to the choice program, etc., as well as some descriptive information about the Milwaukee public schools the students are currently attending, or attended either pre- or post-enrolling in a choice school. In addition, there is family background information based on respondents to a survey administered each fall and spring to all first time choice applicants from the previous year, and to families from the Milwaukee public school sample in 1991. In total, there are data on approximately 2300 applicants to the choice program, and on a sample of approximately 5300 students from the Milwaukee public schools.

The most difficult feature of the Milwaukee Parental Choice Program public use data set for the modelling framework suggested by equations (1)-(3) is the fact that the data do not indicate the choice school to which a student applied. This is a problem because S_i, the probability of selection, is only random <u>conditional on knowing to which school and grade a student applied</u>. This is because applicants to some schools in a certain grade were more likely to be selected than applicants to other schools and/or grades. Therefore, if one does not control for the "application lotteries" (indicators for the school and grade to which an individual applied), any estimated effect of private schools could be spurious. Formally, this implies that the vector of control characteristics included in equations (1)-(3) must include indicators for the specific "application lotteries". In the notation of

[15] Note that Y_{it-1} will be negative for the pre-application test score observations.

equations (1)-(3), these are represented by the vector, Z_4. Although one cannot observe the precise school to which an individual applied, Greene, et. al. (1996) note that one can infer the lottery in which an individual participated by knowing an individual's race, grade to which she applied, and year of application. They do so by highlighting that over 80 percent of the choice students were enrolled in one of three schools. Almost all students who applied to one of these schools were Hispanic, and almost all students who applied to the two others were African-American. Thus, for African-American and Hispanic students, selection can be assumed random, conditional on a set of dummy variables that represent interactions between a child's race, the grade to which he applied and the year of application.[16]

For purposes of most of the analysis in this paper, I use a subsample of the Milwaukee data: those who applied as a voucher student to attend one of the choice schools. I outline, in detail, how I constructed the sample in the Data Appendix, although I highlight some features here. First, I consider a student as having been selected for the choice program if they applied and were selected. I consider a student to have been in the choice program if they took an achievement test while in a private school.[17] For students who either are never selected or are selected the first time they applied, I consider the year of application to be the first year they applied. For those who applied more than once, were initially not selected, but were selected later, I consider their year of application the year in which they were selected. I determine the grade to which a student applied from the grade level of his achievement test the following spring.[18] My basic analysis sample consists of

[16] There are 72 application lotteries (2 races*9 grades*4 years) when all four application cohorts are used. The median number of students in each lottery is about 28.

[17] This means that I do not classify as choice students (in the structural analysis) some students who enrolled in a private school in the fall, but had left the school by the spring.

[18] These are largely the same definitions used by Greene, et. al. (1996). Generally speaking the results are robust to alternative definitions of the year of application and the grade to which a student applied. The one exception is that the third year results (as in Appendix Tables 1a and 1b) are sensitive to the grade at application; that is, using alternative ways of determining the grade to which the student applied, the coefficient drops in magnitude and significance.

African-American and Hispanic students who applied to the choice program between 1990 and 1993 for grades K-8.[19] Sample descriptive statistics are in Appendix Table 2.

I attempt to assess whether the selection was random in Table 3. In the upper panel, I regress the characteristic in the top row on whether the student was selected and (approximately 72) dummy variables indicating the student's applicant pool. In each case, I only include those individuals for whom there is non-missing information on the family background characteristic. It is clear that in four of the five cases, selection appears insignificantly related to the characteristic at the top. In the bottom panel, I "free up" the relationship between being selected and the individual's applicant pool by interacting whether the individual was selected with the applicant pool dummy variables. I present the p-value of the F-test of the joint significance of the interactions. In this case I treat each applicant pool as a separate lottery and test whether, jointly, selection appeared to be random. Whether the individual is female and the student's pre-application test score appear uncorrelated with selection. On the other hand, family income, mother's education, and father's education are not. These last three measures, however, are all drawn from a survey of choice applicants and the fact that the interactions are not jointly zero may reflect the survey response more than the randomization of the experiment. As a result, I consider selection broadly randomly, conditional on the approximations to the applicant pools.

That said, there are notable places where slippage between the actual application lotteries and these constructed lotteries could theoretically occur. First, the grade at the school may not have been oversubscribed. Second, race may not completely determine to which school an individual applied (note, in particular, that African-Americans are assumed to apply to one of two schools). Third, students could apply to more than one school, although I cannot adjust for that with these data.

---

[19] As the Iowa Tests of Basic Skills does not appear to have been administered to students in grades 9-12, my analysis excludes the private high schools that participated in the program (so do Greene, et. al. (1996) and Witte, et. al (1995). The high schools were predominately "alternative" high schools designed for at-risk students.

Table 3

Was Selection Conditionally Random?
Conditional Differences in Characteristics Between those Selected to Attend Private School and Those Not Selected

	Female	Pre-Application Test Score	Family Income	Mother's Education	Father's Education
Constrained Estimate:					
Selected to Attend Private School	0.055 (0.034) [0.104]	1.779 (1.779) [0.318]	342.91 (852.63) [0.402]	0.290 (0.177) [0.102]	0.478 (0.157) [0.002]
Unconstrained Estimate:					
p-value of F-test on the joint significance of interactions between being selected and one's applicant pool	0.713	0.498	0.038	0.017	0.018
Number of Observations	1321	553	709	621	651

Notes: The first row is the regression coefficient on being selected from a regression of the characteristic in the top row on whether selected and applicant pool fixed-effects. Standard errors are in parentheses and p-values are in brackets. The p-values in the second row are based on a regression of the characteristic in the top row on dummies indicating the student's applicant pool and interactions between the applicant pool and whether the student was selected. The pre-application test score is the math test score.

Fourth, the imputation of the grade of application assumes that students continue to progress at grade level. Finally, if the sibling of a student is already enrolled in the choice program, then the student is admitted without having to go through the randomization.

Because I am not completely confident that the approximation perfectly matches the actual application lotteries, I control for whether the student is female, the family income of the student, and in the cross-sectional analyses by year, the student's pre-application test score (these constitute the elements in the vector X_i in equations (1)-(3)).[20] As noted above, the family income has a high rate of missing values as only 60% of those selected and about 43% of those not selected have non-missing values. One may be concerned that including this measure biases against finding an effect of the program due to the differential response rates. A typical concern might be that only the wealthier families among those not selected may have responded to the survey. However, as shown in Table 3, those who were selected had marginally higher levels of family income and parental education than those who were not selected. Thus, while controlling for family income may come at some cost, I believe that it results in a more credible estimate of the effectiveness of the program. The pre-application test score is from the year the student applied to the program. Approximately one-half of the students are missing these test scores (over one-half of those missing scores applied to kindergarten or first grade) and so I also include a dummy variable indicating that it is missing. In the analyses in which I stack the data (and create a panel), any pre-application test scores are included as additional (pre-application) observations. I also use the panel data to estimate models with individual fixed-effects which completely avoids the problem of missing observable covariates.

The test scores are the normal curve equivalent reading and math scores from the Iowa Tests of Basic Skills. These tests were administered to the choice students every spring. They were also administered to Chapter I students every year. Other students in the Milwaukee public schools were

---

[20] I have also tried controlling for mother's and father's education as well, with similar results.

tested less frequently (reportedly in grades 2, 5, and 7).[21] Beginning in 1993, schools were no longer required to administer the entire battery of math sub-tests in order to receive Chapter I funds, rather they were only required to administer the problem solving component.[22] As a result, I do not observe a "total math score" for a substantial fraction of students in the Milwaukee public schools.[23] At the same time, the public release version of the data do not include the problem solving component of the math tests for the students in the choice program. Fortunately, a subset of students in the Milwaukee public schools were administered the entire battery of math tests. Therefore, I impute the total math score by regressing the total score on the problem solving component using the random sample of students in the Milwaukee public schools who did not apply to choice and who took the entire test (see the Data Appendix for the equations). I use the predicted total score for those with only the problem solving component, and include a dummy variable indicating whether the score was imputed. I explore the likely effect of this imputation by comparing it with alternative ways of handling the missing data in Appendix Table 3. I believe that the results are fairly robust to alternative ways of handling this missing data problem.

IV. Results

A. *Evaluating the Choice Program: Reduced-Form Estimates*

I begin by estimating the mean effect of selection to the choice program on math and reading test scores for each calendar year, i.e., equation (2). The results are in Tables 4a and 4b. In the upper

[21] There are very few students in the Milwaukee public schools with test scores from every spring, and the pattern of missing grades does not obviously suggest that all students were tested in grades 2, 5, and 7.

[22] And in 1995, the Milwaukee public schools began administering state mandated tests rather than the Iowa Tests of Basic Skills (Witte, Sterr, and Thorn (1995)); as such I cannot evaluate the program using the 1995 test scores.

[23] In 1993 approximately 40% of the control group is missing the total score, and the percentage rises to 68% in 1994.

Table 4a

The Mean Effect of Selection to the Choice Program on Math Scores: By Year of the Test

	Without Other Covariates			
	1991	1992	1993	1994
	(1)	(2)	(3)	(4)
Selected to Attend Private School	3.143 (2.895)	0.219 (2.035)	3.493 (1.895)	2.793 (2.167)
Female	-0.607 (2.288)	-1.029 (1.587)	1.089 (1.394)	1.518 (1.417)
Imputed Test Score			1.475 (2.097)	1.505 (1.951)
R^2	0.128	0.059	0.125	0.165
Number of Observations	280	585	668	740
	With Other Covariates			
	1991	1992	1993	1994
	(5)	(6)	(7)	(8)
Selected to Attend Private School	1.851 (2.612)	-0.078 (1.966)	2.775 (1.790)	3.498 (2.021)
Female	-0.675 (2.052)	-0.690 (1.513)	1.194 (1.317)	1.384 (1.316)
Imputed Test Score			1.819 (2.009)	1.113 (1.811)
Family Income (÷1000)	0.335 (0.161)	0.355 (0.128)	0.342 (0.108)	0.362 (0.105)
Family Income Missing	-2.115 (2.090)	-0.432 (1.534)	-1.020 (1.331)	-0.772 (1.317)
Pre-Test Math Score	0.586 (0.075)	0.490 (0.068)	0.472 (0.057)	0.584 (0.057)
Pre-Test Score Missing	-1.023 (2.370)	-3.085 (1.879)	-0.692 (1.602)	-1.616 (1.682)
R^2	0.313	0.154	0.228	0.290
Number of Observations	280	585	668	740

Notes: Standard errors are in parentheses. All regressions include a constant and "applicant pool" fixed effects.

Table 4b

The Mean Effect of Selection to the Choice Program on Reading Scores: By Year of the Test

	Without Other Covariates			
	1991	1992	1993	1994
	(1)	(2)	(3)	(4)
Selected to Attend Private School	5.626 (2.599)	1.841 (1.788)	-0.278 (1.618)	1.536 (1.691)
Female	2.518 (2.058)	4.047 (1.426)	2.255 (1.270)	3.696 (1.272)
R^2	0.077	0.082	0.083	0.097
Number of Observations	270	573	675	731
	With Other Covariates			
	1991	1992	1993	1994
	(5)	(6)	(7)	(8)
Selected to Attend Private School	5.953 (2.357)	1.588 (1.745)	-0.775 (1.553)	1.463 (1.594)
Female	1.408 (1.861)	3.922 (1.367)	2.472 (1.211)	3.905 (1.186)
Family Income (÷1000)	0.246 (0.140)	0.246 (0.115)	0.411 (0.099)	0.288 (0.095)
Family Income Missing	-1.424 (1.880)	-0.470 (1.396)	0.142 (1.214)	-2.236 (1.177)
Pre-Test Reading Score	0.732 (0.093)	0.481 (0.070)	0.440 (0.060)	0.595 (0.060)
Pre-Test Score Missing	0.326 (2.114)	-3.514 (1.654)	0.033 (1.493)	-1.180 (1.493)
R^2	0.267	0.164	0.177	0.226
Number of Observations	270	573	675	731

Notes: Standard errors are in parentheses. All regressions include a constant and "applicant pool" fixed effects.

panels of both tables, I report estimates of the "intent to treat" without controlling for family income and pre-application test scores; in the bottom panel I report estimates that control for these covariates.[24] In all cases, I include a constant and fixed effects for the applicant pools described above. In the upper panel of Table 4a, I generally find that the students selected to attend private school scored approximately 3 (roughly) percentage points higher on the math test, although the treatment effect is only statistically significant (at the 10% level) for 1993 (column (3)). In the bottom panel when I control for other covariates, the estimated coefficient falls in 1991, 1992, and 1993 by about 20-40%. On the other hand, the estimate from 1994 increases by 25%. Again, the point estimates suggest that the program has a mean effect of raising math test scores 2-3 percentage points, the p-value of the effect in 1993 is 0.122 and that in 1994 is 0.084.

The results for reading scores are in Table 4b. In both panels, I estimate a large, 5.6-6 percentage point increase in reading scores in 1990 (that is statistically significant at the 5 percent level) for those selected to the choice program. However, the point estimates in later years suggest a smaller intent to treat effect of only about 1.5 percentage points, although the standard errors are generally so large as to make inference difficult. Note, again, that the effect of conditioning on family background leads to an increase in the coefficient in one year (1991) and a slight decrease in all others.

In Tables 5a and 5b, I estimate whether the choice program altered the slope of the education production function, rather than just the intercept, by estimating the effect of years since selection on test scores, both constraining the effect to be linear and allowing the effect to differ by year and cohort. In the top panel of Table 5a, I estimate the linear effect of years since selection on math test

---

[24] I do not control for the grade of the student at the time of the test. This appears to make little difference to the coefficient estimates on the effect of the choice program. I, however, note the instances in which it does affect the estimates for the control group. I do not include the grade in the results presented here since whether a child is at grade level could be considered an outcome itself while also being a policy instrument of the school. It appears that approximately 5% of the students are held back in any year (and the overall rate does not differ between the public and private schools). Controlling for whether the student was held back does not affect the results.

Table 5a

The Effect of Selection to the Choice Program on Math Scores: By Year and Years Since Selection

	Year of Test Score			
	1991	1992	1993	1994
	(1)	(2)	(3)	(4)
Number of years since selection	1.851 (2.612)	0.272 (1.169)	1.116 (0.798)	1.685 (0.689)
R^2	0.313	0.154	0.227	0.293
Number of Observations	280	585	668	740
	Year of Test Score			
	1991	1992	1993	1994
	(5)	(6)	(7)	(8)
1st year after selection (Year 1)	1.851 (2.612)	-2.311 (3.120)	3.785 (2.977)	-1.377 (4.404)
2nd year after selection (Year 2)		1.360 (2.509)	0.699 (3.128)	1.146 (3.282)
3rd year after selection (Year 3)			3.514 (2.796)	1.868 (4.182)
4th year after selection (Year 4)				9.328 (3.354)
p-value of F-test of jt. sign. of being selected		0.653	0.383	0.091
p-value of H_0: Year1=Year2		0.357	0.460	0.631
p-value of H_0: Year2=year3			0.494	0.888
p-value of H_0: Year4=Year3				0.154
p-value of H_0: Year3=Year1			0.946	0.582
p-value of H_0: Year4=Year2				0.072
p-value of H_0: Year4=Year1				0.049
R^2	0.313	0.155	0.229	0.295
Number of Observations	280	585	668	740

Notes: Standard errors are in parentheses. All regressions include a constant, "applicant pool" fixed effects, a dummy indicating if female, family income, and pre-application math test score, and indicators if income and pre-test score are missing. Columns (3), (4), (7), and (8) also include a dummy indicating if the test score was imputed.

Table 5b

The Effect of Selection to the Choice Program on Reading Scores:
By Year and Years Since Selection

	Year of Test Score			
	1991	1992	1993	1994
	(1)	(2)	(3)	(4)
Number of years since selection	5.593 (2.357)	1.102 (1.035)	0.121 (0.709)	0.576 (0.564)
R^2	0.267	0.165	0.176	0.226
Number of Observations	270	573	675	731
	Year of Test Score			
	1991	1992	1993	1994
	(5)	(6)	(7)	(8)
1st year after selection (Year 1)	5.593 (2.357)	0.155 (2.779)	-3.619 (2.634)	3.138 (3.938)
2nd year after selection (Year 2)		2.498 (2.221)	-1.360 (2.840)	-0.812 (2.698)
3rd year after selection (Year 3)			2.379 (2.562)	1.366 (3.488)
4th year after selection (Year 4)				3.306 (2.961)
p-value of F-test of jt. sign. of being selected		0.531	0.396	0.712
p-value of H_0: Year1=Year2		0.508	0.558	0.407
p-value of H_0: Year2=year3			0.327	0.620
p-value of H_0: Year4=Year3				0.670
p-value of H_0: Year3=Year1			0.103	0.735
p-value of H_0: Year4=Year2				0.304
p-value of H_0: Year4=Year1				0.973
R^2	0.267	0.165	0.180	0.227
Number of Observations	270	573	675	731

Notes: Standard errors are in parentheses. All regressions include a constant, "applicant pool" fixed effects, a dummy indicating if female, family income and pre-application reading test score, and indicators if income and pre-test score are missing.

scores in each year. In 1993 (column (3)), each additional year since being selected led to a 1.1 percentage point increase in math scores (p-value of 0.16). In 1994, the effect was even larger, 1.7 percentage points, and the yearly increase is significantly different from zero.

Note that in 1991 the only cohort included in the analysis is the group of students who applied in 1990; in 1992, the cohorts who applied in 1990 and 1991 are included, and so on. Thus, in the bottom panel one can follow a cohort by comparing the coefficients on the diagonals. For example, the lowest diagonal represents the yearly progress of the 1990 cohort; the diagonal just above it represents the progress of the 1991 cohort. In the first year, the 1990 cohort had an increase in test scores of 1.9 percentage points; in 1991, the same cohort had a difference of 1.4 percentage points. However, by 1993, the gap between those selected and those not selected had grown to 3.5 percentage points and by 1994 the gap was 9.3 percentage points. (The fourth year is the only one that is statistically significantly different from earlier years, and it is only different from the first and second years.) The 1991 cohort had a similar growth in their test scores, although the coefficient estimates are not significantly different from zero. The point estimates in this table suggest that there may have been faster math achievement gains among the private school students than among the public school students.

The magnitudes of the point estimates for reading, in Table 5b, are not so clear. In the top panel, I again find that the yearly increase, though positive, is far from statistically significant in 1992-1994. In the bottom panel, the yearly progress of the 1990 applicant cohort was not monotonic as they experienced a huge increase in test scores in the first year that was unmatched in subsequent years. On the other hand, the coefficient estimates for both the 1991 and 1992 cohorts suggest yearly gains, although the standard errors are quite large as well. In general, the results from Tables 5a and 5b suggest that being selected to a choice program may have altered the path of educational achievement, and not just the mean, although the confidence intervals are generally large.

In the previous tables, I estimated separate effects for each cohort in each year. There may be some efficiency gain by fitting a pooled model that assumes the yearly effects are similar across cohorts and years. To implement this pooled model, I created a panel data set from my analysis sample, treating each year in which students took an achievement test (including pre-application tests) as a separate observation. I then estimated spline functions (see equation (10)) which allow me to measure the achievement path of the control group and test for discontinuities in the slope of the achievement gains. I present OLS estimates with Huber standard errors that allow for individual correlations and that are also robust to heteroskedasticity. I also present models that include individual fixed-effects.[25] In Tables 6a-6c I only present the effects on math test scores as I found no significant discontinuities or significant differences between the control and treatment groups for reading scores.[26]

The basic results for the pooled model are in Table 6a. In the first specification (columns (1) and (2)), I estimate a basic spline function in which I allow only the intercept and slope to differ between the control and treatment groups. The estimates in column (1) are by OLS and those in column (2) include individual fixed-effects (FE). Both columns indicate that the intercept for those selected to the choice program is slightly lower than that for the control group, although the difference is not statistically different from zero. On the other hand, it appears that the two groups have significantly different slope estimates. The math test score for those who were not selected for the program decreased, on average, approximately one percentage point per year, however this may be due to students being in higher grades.[27] The test scores of those selected for the program increased an additional (statistically significant) 1.5 percentage points per year.

[25] I have also estimated generalized least squares random effects models with similar results.

[26] The results for reading scores are available from the author upon request.

[27] When I control for the grade at the time of the test, the control group actually had a positive (and significant) 4 percentage point increase in test scores per year. The magnitude and statistical significance of the coefficient on the interaction between years since application and whether selected remains basically unchanged.

Table 6a

The Effect of Selection to the Choice Program on Math Scores, 1991-1994: OLS and Fixed-Effects (FE) Spline Functions

	OLS	FE	OLS	FE	OLS	FE	OLS	FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Selected to attend private school	-0.841	-1.539	-0.062	-1.125	-0.197	-1.430	-0.271	-1.133
(α'_p)	(1.061)	(0.962)	(1.347)	(1.287)	(1.108)	(1.032)	(1.355)	(1.297)
Number of years since application	-1.056	-0.952	-0.698	-0.760	-0.809	-0.830	-0.744	-0.773
(δ')	(0.384)	(0.350)	(0.440)	(0.380)	(0.403)	(0.362)	(0.443)	(0.381)
Number of years since application*Selected	1.509	1.487	0.459	0.916	0.779	1.252	0.744	0.944
(β')	(0.510)	(0.468)	(0.868)	(0.815)	(0.580)	(0.543)	(0.879)	(0.830)
$\geq$ 2 Years (δ'_2)			-2.412	-1.575			-0.841	-0.887
			(1.445)	(1.245)			(2.113)	(1.917)
($\geq$ 2 years)*Selected (β'_3)			3.516	2.199			0.757	1.436
			(1.845)	(1.647)			(2.524)	(2.335)
$\geq$ 3 Years (δ'_3)					-5.465	-3.149	-4.128	-1.799
					(2.815)	(2.476)	(4.071)	(3.813)
($\geq$ 3 Years)*Selected (β'_3)					8.028	3.735	6.752	1.967
					(3.242)	(2.847)	(4.545)	(4.194)
p-value of H_0: $\delta'=0$, $\delta'_2=0$, $\delta'_3=0$			0.006	0.012	0.003	0.011	0.010	0.027
p-value of H_0: $\beta'=0$, $\beta'_2=0$, $\beta'_3=0$			0.007	0.011	0.002	0.004	0.007	0.027
p-value of H_0: $\delta'_2=0$, $\delta'_3=0$							0.151	0.402
p-value of H_0: $\beta'_2=0$, $\beta'_3=0$							0.066	0.366
p-value of H_0: $\delta'_2=\delta'_3$							0.571	0.866
p-value of H_0: $\beta'_2=\beta'_3$							0.359	0.930
R^2	0.098	0.008	0.099	0.009	0.099	0.009	0.010	0.009

Notes: There are 3283 observations. Standard errors are in parentheses. The OLS columns report Huber standard errors that allow for correlations "within" an individual (since these data are stacked). The OLS regressions include a constant, "applicant pool" fixed effects, a dummy indicating if female, family income, and an indicator if income is missing. All regressions include a dummy indicating if the test score was imputed.

Table 6b

The Effect of Selection to the Choice Program on Math Scores, 1991-1994: Robustness of Spline Functions (OLS Estimates)

	Basic Regression*			Spline Function with Discontinuities*		
	(1)	(2)	(3)	(1)	(2)	(3)
Selected to attend private school (α'_p)	-0.841	-2.335	-0.079	-0.062	-3.395	-0.167
	(1.061)	(1.392)	(1.112)	(1.347)	(1.693)	(1.355)
Number of years since application (δ')	-1.056	-0.559	-0.871	-0.698	-0.566	-0.798
	(0.384)	(0.484)	(0.411)	(0.440)	(0.498)	(0.447)
Number of years since application*Selected (β')	1.509	1.454	0.755	0.459	2.305	0.720
	(0.510)	(0.773)	(0.581)	(0.868)	(1.093)	(0.879)
$\geq$ 3 Years (δ'_2)				-2.412	0.391	-0.960
				(1.445)	(3.702)	(2.133)
($\geq$ 3 years)*Selected (β'_2)				3.516	-2.372	0.855
				(1.845)	(4.158)	(2.537)
p-value of H_0: $\delta'=0$, $\delta'_2=0$				0.006	0.518	0.093
p-value of H_0: $\beta'=0$, $\beta'_2=0$				0.007	0.104	0.449
R^2	0.098	0.114	0.100	0.099	0.114	0.100
Number of Observations	3283	2102	3096	3283	2102	3096

Notes: Huber standard errors (that allow for correlations within an individual) are in parentheses. All regressions include a constant, "applicant pool" fixed effects, a dummy indicating if female, family income, an indicator if income is missing, and a dummy indicating if the test score was imputed.

* Column Descriptions:
- Column 1: Uses Full Sample
- Column 2: Only Uses 1991-1993 Cohorts
- Column 3: Excludes 1994 Test Scores for 1990 Cohort

Table 6c

The Effect of Selection to the Choice Program on Math Scores, 1991-1994: Robustness of Spline Functions (Individual Fixed-Effects Estimates)

	Basic Regression*		
	(1)	(2)	(3)
Selected to attend private school (α'_p)	-1.539	-3.024	-1.494
	(0.962)	(1.297)	(1.040)
Number of years since application (δ')	-0.952	-0.691	-0.756
	(0.350)	(0.402)	(0.369)
Number of years since application*Selected (β')	1.487	1.955	1.156
	(0.468)	(0.686)	(0.547)
R^2	0.008	0.010	0.005
Number of Observations	3283	2102	3096

Notes: Standard errors are in parentheses. All regressions include individual fixed-effects and a dummy indicating if the test score was imputed.

* Column Descriptions:
  - Column 1: Uses Full Sample
  - Column 2: Only Uses 1991-1993 Cohorts
  - Column 3: Excludes 1994 Test Scores for 1990 Cohort

In the remaining columns, I attempt to establish if there were discontinuities in the achievement gains, as Greene, et. al. (1996) argue that the choice program led to significant jumps in achievement in the third year and then again in the fourth year. In columns (3) and (4) I allow for a discontinuity three years after selection. Now, the linear effect of being selected for the program (β') suggests a positive, but insignificant yearly increase in test scores of about 0.5-1 percentage point. On the other hand, the OLS estimate suggests there was a break in the slope at the third year for the treatment group that is statistically significant. In columns (5) and (6) I allow for a break in the slope in the fourth-year, and not the third, and a find similar result for the OLS estimate. The fixed-effects estimate, in column (6), indicates the linear effect of being selected was positive (and significant at the 5% level), but there was an insignificant increase in test scores in the fourth year. Finally in the last two columns I allow for discontinuities in both the third and the fourth years. Allowing for both reveals that the discontinuity in the third year estimated in column (3) was likely driven by the increase in the fourth year; further, the jump in test scores in the fourth year is no longer significant for either the OLS or fixed-effects estimates. Also note that the control group had a significant downward jump in test scores in the fourth year, as well, increasing the intent to treat effect.

As highlighted in Table 5a, the mean difference in test scores between those selected and those not selected was much higher in the fourth year for the 1990 applicant cohort than for any other year or cohort. To assess the extent to which the results might be driven by this cohort, I tried using alternative samples. First, I used the full sample. Second, I estimated the spline functions excluding the 1990 cohort. Third, I excluded the 1994 test scores for 1990 cohort, but included their 1991-1993 test scores. The results are in Tables 6b and 6c. The first three columns of Table 6b are the basic spline functions that do not allow for discontinuities in the yearly gains; the second three columns allow for a discontinuity in the third year.[28] The results for the basic spline function in the

[28] Given that I did not estimate significant discontinuities using the fixed-effects estimator, I do not present those results in Table 6c. I continue to find no significant effects using the alternative samples; these results are available from the author upon request.

first three columns are quite similar across the first two samples. The test scores of those selected for the program increase by about 1.5 percentage points each year above those who were not selected. Excluding the 1994 test scores for the 1990 cohort, however, the coefficient on the yearly increase in test scores drops by about one-half, and the estimate is no longer statistically significant. This differs from the fixed-effects estimates in Table 6c in which I estimate a (statistically significant) yearly increase of 1-2 percentage points per year using all three samples.

In the last three columns of Table 6b I allow for a discontinuity in the third year. In this specification, the three samples all suggest positive (in columns (1) and (3)), although insignificant, yearly gains in the treatment group relative to the control group. However, the discontinuity in the third year only obtains in the first column which includes the 1994 test scores from the 1990 cohort. Thus, while the yearly gains appear fairly robust, it is much less clear that there are discontinuities in the trend.[29]

Before I turn to "causal" estimates of the relative effectiveness of private schools, I assess the potentially important effect of sample attrition on the parameter estimates. For example, only 46% of students who applied in 1990 had non-missing 1994 test scores (49% among those selected for the program and 35% among those not selected). Test scores may be missing for four reasons. First, the student may have still been enrolled in a choice or Milwaukee public school, but was absent the day of the test. Test score data missing for this reason can likely be considered random. Second, as mentioned earlier, students in the Milwaukee public schools were not necessarily tested every year. For those who were not tested because the school did not test their grade level, their scores could also be considered randomly missing. However, Chapter I students were allegedly tested every year which would suggest that those missing data from the Milwaukee public schools come from families with higher incomes than those who were tested more frequently (or that a student whose family

[29] I have also tried excluding those students who leave the choice school because they reached the terminal grade for the school. The results remain essentially unchanged.

income increases would be tested frequently in the earlier years, but tested less frequently in later years). Third, it appears as though students in grades 9-12 were not tested; therefore, older applicants would be missing test scores in later years. Finally, any student who enrolled in a private school not in the choice program or who transferred to a public school outside of Milwaukee is missing test scores.[30] Again, it is not clear that one should assume this group is randomly missing from the sample.

Figures 1a-1d are kernel densities of the distributions of (conditional) pre-application test scores by whether the student is missing a test score in 1994 or not, separately for the treatment and control groups.[31] The distributions of the pre-application math scores are in Figures 1a and 1b. For both those selected for the program and those not selected, the pre-application math test score distributions of those missing the 1994 math test score and those with test scores largely overlap. Among the controls, however, it appears that those without test scores are of slightly lower "ability". On the other hand, for the treatment group, it appears that those without 1994 test scores are of slightly higher "ability". Together, these figures suggest that the previous estimates do not overstate the effectiveness of the choice program on math scores, and may well understate its effectiveness. Similar distributions for reading scores are in Figures 1c and 1d. There is even more overlap in the distributions for the reading scores, although in both figures, those with 1994 reading test scores appear to be marginally positively selected.

[30] Witte (1996) argues that a large fraction of students who were not selected in the lotteries chose to attend another private school (not participating in the choice program). This was made easier by a parallel, privately funded program (Partners for Advancing Values in Education (PAVE)) that provided scholarships for Catholic schools. It is for this reason that Witte, Sterr, and Thorn (1995) prefer to use the random sample of students from the Milwaukee public schools as a comparison group.

[31] These test scores are adjusted for whether the student is female, family income and applicant pool fixed effects, and for the math scores, whether the test was imputed because the student only took the problem solving component of the math test.

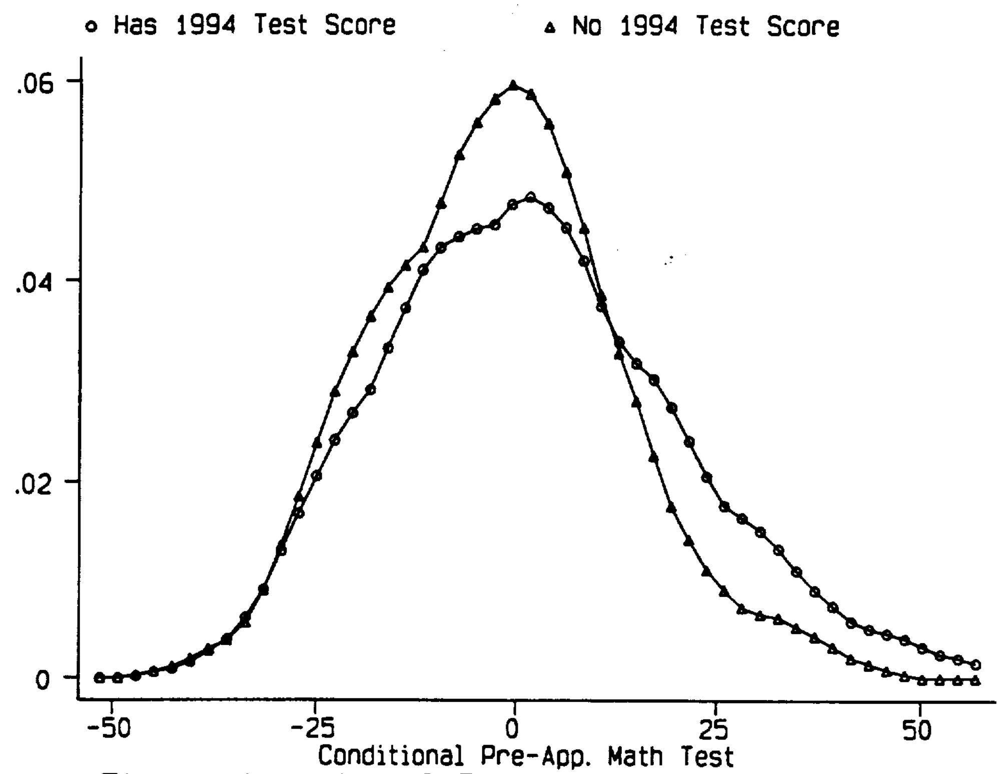

Figure 1a. Kernel Densities for Control

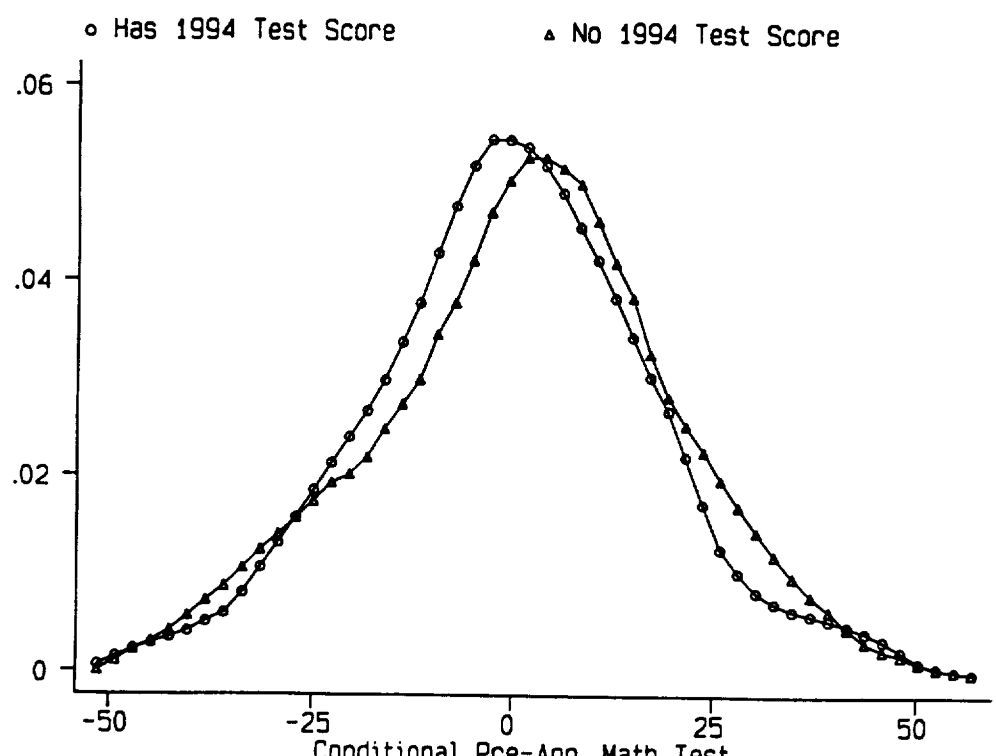

Figure 1b. Kernel Densities for Treatment

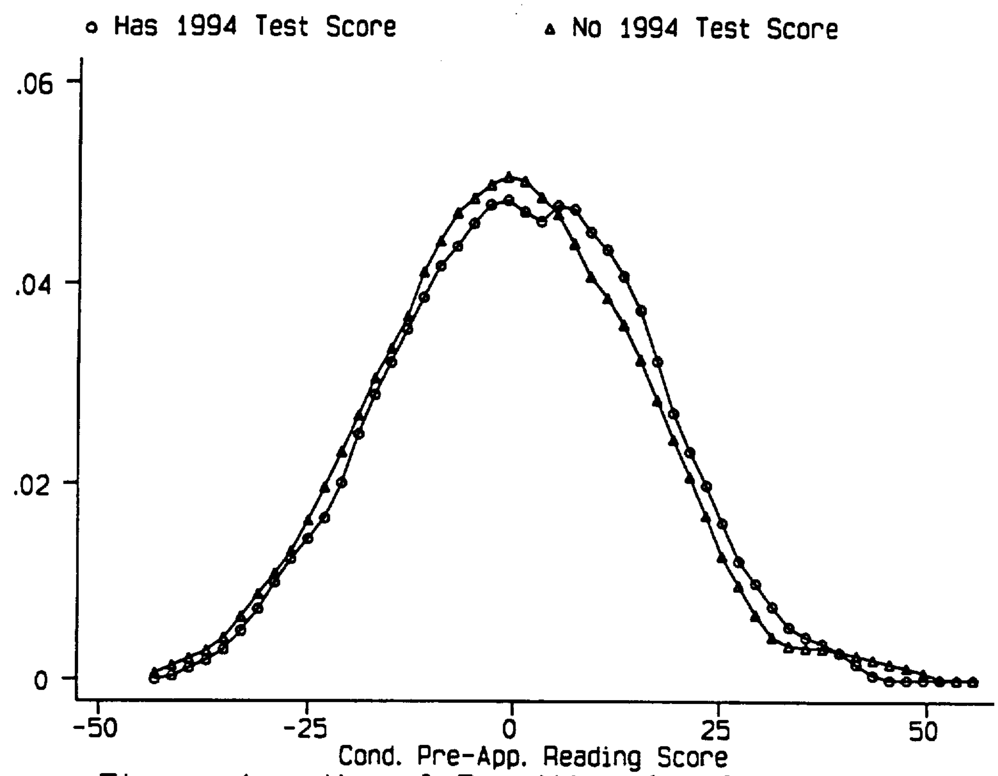

Figure 1c. Kernel Densities for Control

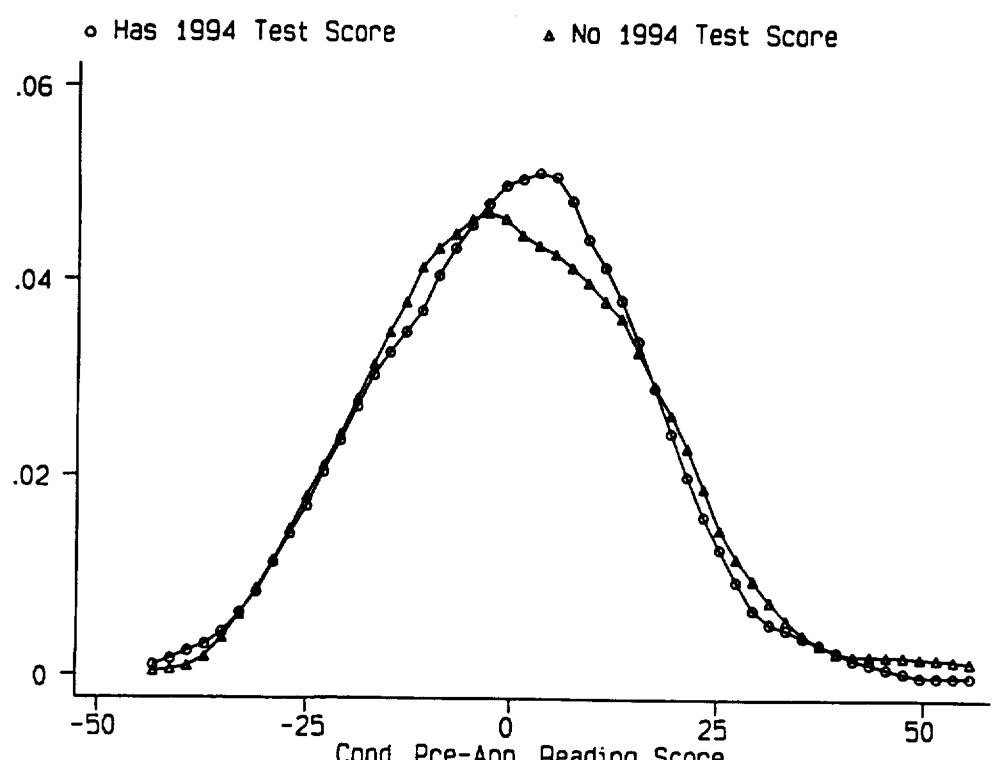

Figure 1d. Kernel Densities for Treatment

I more formally assess the likely effect of missing test scores on the results in Table 7.[32] In this table I employ four alternative methods for handling the missing test score data. In the first column I present OLS estimates using my analysis sample. In column (2) I use a two-step Heckman selection correction in which I exclude the grade of the student in the year of the test from the second stage.[33] The point estimates for both the math and reading scores increase when I use a Heckman selection correction, and in both cases the estimate of λ is negative and significant. One problem with using a Heckman selection correction in this context is that there is not one underlying reason for the missing test scores. As a result, there is no single latent variable that may underlie the first stage probit.

In columns (3) and (4) I attempt to impute a test score for those missing a test score in two ways. In column (3) I first regress the current test score on the pre-application test score and, for the math scores, a dummy indicating whether the current test score was imputed (to account for the fact that the student did not take the entire battery of math tests).[34] I run separate regressions for the treatment and control groups. I then use the predicted test score as the current test score for those with missing data. For reading scores, this imputation does not substantively change the results from those in column (1). However, for math, the point estimate decreases by one-half and becomes statistically insignificant. In column (4) I impute the current test score by computing each student's decile in the (conditional) pre-application test score distribution. I then assign the mean test score from the corresponding decile in the test score distribution for those missing test scores in the current year. The point estimates using this approach are similar to those in column (1) for both reading and

---

[32] For this analysis, I exclude individuals who are missing test scores and are enrolled in either a pre-kindergarten program or in grades 9-12.

[33] I use the grade at the time of the test as it was alleged that in the Milwaukee public schools, the tests were more consistently given in certain grades. First-stage estimates are available from the author upon request.

[34] The "pre-application test score" in this context is the test score from the year of application. I have also tried using higher orders of the pre-application test score with very similar results.

Table 7
OLS Estimates of the Effect of Selection to the Choice Program on Math and Reading Scores: Robustness to Sample Attrition

	Baseline	Two-Step Heckman Selection	Impute Missing by Method 1[†]	Impute Missing by Method 2[††]	Only use those with no missing test score data
	(1)	(2)	(3)	(4)	(5)
			Math Scores		
Selected to attend private school (α'_p)	-0.841 (1.187)	-6.223 (1.238)	0.562 (0.698)	-0.147 (0.710)	-2.083 (1.629)
Number of years since application (δ')	-1.056 (0.399)	-1.416 (0.362)	-0.359 (0.179)	-0.457 (0.182)	-1.366 (0.611)
Number of years since application*Selected (β')	1.509 (0.574)	2.574 (0.569)	0.718 (0.325)	1.222 (0.331)	2.358 (0.822)
λ		-8.999 (0.875)			
R^2	0.098		0.072	0.068	0.128
Number of Observations	3283	6389	6389	6389	1607
			Reading Scores		
Selected to attend private school (α'_p)	1.576 (1.051)	-1.556 (1.148)	0.892 (0.591)	1.984 (0.628)	2.010 (1.388)
Number of years since application (δ')	-0.794 (0.331)	-1.170 (0.328)	-0.506 (0.147)	-0.565 (0.156)	-0.721 (0.489)
Number of years since application*Selected (β')	-0.200 (0.517)	0.424 (0.520)	-0.046 (0.278)	-0.158 (0.296)	0.133 (0.727)
λ		-6.254 (0.899)			
R^2	0.067		0.049	0.049	0.105
Number of Observations	3268	6391	6391	6391	1567

Notes: Standard errors are in parentheses. See note to Table 5b for other regressors.

[*] The grade of the individual at the time of the test and interactions of the grade with whether the individual was randomly selected are excluded from the first stage equation.

[†] "Method 1" regresses the current test score on pre-application test score, and for math whether the current test score was imputed (because the student did not take the entire test battery) separately for the treatment and control groups. The predicted current test score is used to impute the actual test score if the test score is missing.

[††] "Method 2" assigns those missing the current test score their application cohort's mean test score for the decile of the selected or not-selected distribution corresponding to their decile in their application cohort's pre-application test score distribution.

math.[35] Finally, in column (5) I restrict my sample to those without missing test scores (note that the sample size falls dramatically). These point estimates are a little higher than those in column (1) for math, and are now positive, but insignificant, for reading.

These results (together with Figures 1c and 1d) suggest that missing reading test scores are not likely a large source of bias for the estimates of the intent to treat. For math, the evidence broadly indicates that the finding of a significant and positive effect of being selected for the program is robust to the missing test scores, although one should keep the sample attrition in mind when interpreting the results.

In all, being selected to participate in the choice program appears to have increased the math achievement of low-income, minority students by about 1-2 percentage points per year. Given a standard deviation of about 19, this suggests effect sizes on the order of 0.05σ-0.11σ which are quite large for education production functions (see, for example, Greenwald, Hedges, and Laine (in press)). On the other hand, the effects on the reading scores are not as large and are statistically indistinguishable from zero.

B. "Causal" Estimates of the Effectiveness of Private Schools

In this section, I estimate OLS models of the effect of attending private school on test scores, and instrumental variables (IV) estimates using the initial selection as the instrumental variable. Note that I define "attending a private school" as having taken an achievement test in a choice school in the particular year. The first-stage estimates (equation (1)) are presented in Table 8. In all years, having been randomly selected to participate in the choice program increases a student's likelihood of actually attending a private school by about 70 percentage points.[36] The estimate is statistically

[35] Note, however, that I have not adjusted the standard errors to reflect the two-step estimation strategy; the standard errors in columns (3) and (4) are likely understated.

[36] A few students who were not selected initially were enrolled in the choice program in the following spring. They were likely admitted off of a waiting list.

Table 8

Linear Probability Estimates of the Effect of Being Selected on Private School Attendance (First-Stage Estimates)

	Dependent Variable			
	In Private School in 1991	In Private School in 1992	In Private School in 1993	In Private School in 1994
	(1)	(2)	(3)	(4)
Selected	0.761 (0.050)	0.703 (0.038)	0.679 (0.039)	0.681 (0.040)
R^2	0.589	0.529	0.478	0.476
Number of Observations	280	585	668	740

Notes: Standard errors are in parentheses. All regressions include a constant, "applicant pool" fixed effects, a dummy indicating if female, family income, and pre-application math test score, and indicators if income and pre-test score are missing.

significant in all years, as well.

The mean achievement differences between students in the private schools and those in the public schools are presented in Table 9. The first four columns present the OLS estimates, and the last four the IV estimates. For the math scores, in three of the four years, the OLS estimates suggest that those who attended a private school scored lower on the math test than those who were not enrolled in a private school in that year, although these differences are not statistically significant. If, however, there is an achievement effect for those who receive partial treatment (i.e., attend a private school and then leave and return to a Milwaukee public school), the OLS estimate is likely to be an underestimate of the full treatment effect (see equation (5)). In the later years in these data, $\rho' \approx 0.5$-0.65; and, on average, those who have left the choice program received approximately one-half of the treatment as those stayed. Therefore, assuming that $\beta_0=2\psi$, one would expect the OLS estimate to understate the full effect by 30-50% (plus an adjustment for the selection effect). That the OLS coefficient estimates are zero, or slightly negative, suggests that those who received partial treatment may have benefitted as much as, or even slightly more than, those who received full treatment. These OLS estimates could be interpreted as a lower bound on the full treatment effect.

As discussed in Section II, the IV estimate is a re-scaled version of the reduced-form parameter, namely $\beta_{IV}=(\pi_1/\rho)$. Therefore, using the estimates of ρ from Table 8, we would expect that, on average, the IV estimate will be approximately $1.4\pi_1$. And this is the case. For example, in Table 9, the 1991 IV estimate is 2.43 which is 1.85/0.76 (which are the estimates of π_1 and ρ from Tables 4a and Table 8). In general, the IV estimates show that students who were enrolled in the private schools scored approximately 3 percentage points higher than those not enrolled in a private school, although the effect is only significant in 1994. For reading, both the OLS and IV estimates suggest a significant 5-8 percentage point gain for those enrolled in private schools in 1991, although this is not evident in other years. Finally, as noted in equation (8) in Section II, these IV estimates likely provide an upper bound on the full treatment effect because of non-compliance. Again using

Table 9

The Mean "Causal" Effect of Private School Attendance on Math Scores: By Year

	OLS				IV			
	Year of Test Score				Year of Test Score			
	1991	1992	1993	1994	1991	1992	1993	1994
	Math Scores							
Attended private school	-0.759 (2.365)	-0.730 (1.737)	2.581 (1.797)	-0.336 (2.377)	2.433 (3.447)	-0.110 (2.795)	3.238 (2.468)	4.327 (2.681)
R^2	0.312	0.154	0.228	0.287				
No. of Observations	280	585	668	740	280	585	668	740
	Reading Scores							
Attended private school	5.139 (2.115)	1.211 (1.548)	-0.854 (1.316)	0.352 (1.286)	7.822 (3.109)	2.284 (2.511)	-1.135 (2.275)	2.155 (2.354)
R^2	0.265	0.164	0.177	0.225				
No. of Observations	270	573	675	731	270	573	675	731

Notes: Standard errors are in parentheses. See notes to Table 5 for other regressors. The instrument is whether the student was randomly selected for the program.

the estimates of ρ from Table 8, Θ ranges from about 0.5 to 0.7. And, assuming that $\beta_0=2\psi$, the IV estimates likely overstate the full treatment effect by about 25%-35%.[37]

In Table 10a, I estimate basic spline functions by OLS and IV while also assessing the influence of the 1990 cohort on the estimates; in Table 10b, I include individual fixed-effects. Focusing on the IV estimates in Table 10a, in the first specification which includes the entire sample, I estimate that those students who attended a private school had a lower (although statistically insignificant) initial math test score, but that they had significantly higher yearly gains than those not enrolled in a private school. The results in columns (2) and (3) which exclude the 1990 cohort and do not include the 1994 test scores for the 1990 cohort respectively, broadly support the trend derived from the full sample, although both suggest a slower yearly increase in test scores.[38] The results in Table 10b which include individual fixed-effects are quite similar.

The IV estimates for the reading test scores (in the lower panels) indicate that students enrolled in the private schools generally start with higher (but insignificant) scores than those not enrolled in the private schools, although there do not appear to be significantly greater yearly gains.[39]

Taken together, the range of credible estimates of the causal effect of private schools on yearly increases in math scores using the applicant pool control group is about 2 percentage points per year (after adjusting for the upward IV bias), although there is no measurable effect on reading scores.

---

[37] The bias may actually be larger as it appears that in this program, $\psi=\beta_0$ which would imply a bias of about 100%. For evidence, see the OLS results in this table, and the results in Table 11 or Appendix Table 1a.

[38] When I use the Greene, et. al. (1996) sample restrictions, I also find there are significant yearly gains in the test scores, except when I exclude the 1990 cohort or exclude the 1994 test scores for the 1990 cohort. In addition, I do not estimate statistically significant jumps in the linear trend at either the 3rd or 4th years.

[39] I have also tried excluding the 1991 test scores for the 1990 cohort with similar results.

Table 10a
OLS and IV Estimates of the Effect of Private Schools on the Rate of Growth in Test Scores

	OLS			IV		
	(1)	(2)	(3)	(1)	(2)	(3)
			Math Scores			
Selected to attend private school (α'_p)	-2.203	-4.565	-2.046	-1.313	-2.855	-0.277
	(1.153)	(1.490)	(1.204)	(1.290)	(1.638)	(1.307)
Number of years since application (δ')	-0.659	-0.652	-0.814	-1.473	-0.662	-1.052
	(0.346)	(0.438)	(0.363)	(0.521)	(0.637)	(0.521)
Number of years since application*Private School (β')	1.555	2.833	1.627	2.883	2.051	1.432
	(0.621)	(0.899)	(0.671)	(0.975)	(1.302)	(1.010)
R^2	0.098	0.117	0.100			
Number of Observations	3283	2102	3096	3283	2102	3096
			Reading Scores			
Selected to attend private school (α'_p)	0.013	-3.923	-0.086	1.693	-1.578	1.960
	(1.055)	(1.403)	(1.129)	(1.147)	(1.501)	(1.210)
Number of years since application (δ')	-0.744	-0.190	-0.632	-0.810	-0.103	-0.629
	(0.253)	(0.330)	(0.281)	(0.355)	(0.432)	(0.373)
Number of years since application*Private School (β')	0.317	1.688	0.253	-0.016	0.344	-0.535
	(0.528)	(0.800)	(0.602)	(0.878)	(1.200)	(0.940)
R^2	0.066	0.068	0.063			
Number of Observations	3268	2086	3081	3268	2086	3081

Notes: Huber standard errors (that allow for individual correlation) are in parentheses. See note to Table 5b for other regressors. The instruments are whether randomly selected, and whether randomly selected interacted with years since application.

Column Descriptions:
- Column 1: Uses Full Sample
- Column 2: Only Uses 1991-1993 Cohorts
- Column 3: Excludes 1994 Test Scores for 1990 Cohort

Table 10b
OLS and IV Estimates of the Effect of Private Schools on the Rate of Growth in Test Scores
(Individual Fixed-Effects)

	OLS with FE			IV with FE		
	(1)	(2)	(3)	(1)	(2)	(3)
	Math Scores					
Selected to attend private school (α'_p)	-2.724	-4.995	-2.828	-2.126	-3.776	-1.917
	(0.991)	(1.345)	(1.077)	(1.149)	(1.499)	(1.219)
Number of years since application (δ')	-0.496	-0.503	-0.611	-1.204	-0.802	-0.865
	(0.277)	(0.355)	(0.303)	(0.452)	(0.482)	(0.450)
Number of years since application*Private School (β')	1.167	2.439	1.473	2.644	2.737	1.776
	(0.504)	(0.773)	(0.598)	(0.920)	(1.114)	(0.950)
R^2	0.007	0.015	0.007			
Number of Observations	3283	2102	3096	3283	2102	3096
	Reading Scores					
Selected to attend private school (α'_p)	0.379	-2.580	0.372	1.335	-2.110	0.869
	(0.937)	(1.292)	(1.029)	(1.058)	(1.420)	(1.145)
Number of years since application (δ')	-0.431	-0.135	-0.302	-0.364	-0.201	-0.416
	(0.220)	(0.286)	(0.247)	(0.314)	(0.361)	(0.333)
Number of years since application*Private School (β')	-0.565	0.455	-0.766	-1.005	0.492	-0.591
	(0.461)	(0.736)	(0.561)	(0.803)	(1.055)	(0.867)
R^2	0.006	0.005	0.005			
Number of Observations	3268	2086	3081	3268	2086	3081

Notes: Standard errors are in parentheses. All specifications include individual fixed-effects, and for the math scores, an indicator if the total test score was imputed. The instruments are whether randomly selected, and whether randomly selected interacted with years since application.

Column Descriptions:
- Column 1: Uses Full Sample
- Column 2: Only Uses 1991-1993 Cohorts
- Column 3: Excludes 1994 Test Scores for 1990 Cohort

C. *Estimating the Effect of Private Schools Using the Milwaukee Public School Comparison Group*

While these results are partially consistent with those estimated by Greene, et. al. (1996), the math score results conflict with those of Witte, Sterr, and Thorn (1995). Given that the primary difference between the studies is in the choice of control or comparison group, in this section I estimate a basic spline function using the random sample of students from the Milwaukee public schools. One reason why treatment effects may not emerge when using the sample of students from the Milwaukee public schools is that they may not be a representative sample of students eligible for the choice program. They may be too advantaged. As noted in Table 1, the mean family income of the students in the Milwaukee public schools is roughly twice that of students who applied to the choice program. However, if this is the case, even if mean differences in the choice students and the Milwaukee public school students are not statistically different from one another, the choice students may have experienced a faster growth in test scores. Further, including individual fixed-effects would control for any time-invariant unobserved differences between the choice students and the comparison sample.

Using this comparison sample also has some advantages over the applicant pool control group. First, many of the rejected applicants may have left the Milwaukee public schools to attend other private schools. For example, in Table 1 only 28% of non-selected applicants have test scores two years after application compared to 54% of the choice students who were admitted. This could result in a non-representative control group and is particularly a problem for the later years when the sample sizes for the control group are quite small. Second, I could not truly identify the schools to which students applied, and was therefore not completely confident of the application lotteries; the analysis using the Milwaukee public school students does not use such information. Third, I can include the entire population of choice students, not just African-Americans and Hispanics. Finally, the estimated effects of private school are relative to the entire population of public school students, and not just a group of public school students interested in attending private school.

For this analysis, I continue to use my sample of applicants to the choice program. In order to incorporate the students in the Milwaukee public school sample, I need to create an artificial "year of application" for them in order to have a year of reference from which to measure the yearly changes in scores. I do so by considering 1990 the "year of application" for students in the Milwaukee public schools who have a valid 1991 test score. For those without a 1991 test score, I consider their "year of application" to be 1991 if they have a valid 1992 test score, and so forth. This allows me to compare the yearly test score progress of consistent cohorts. Any tests taken before this "year of application" are also included and they allow me to identify the mean effect of attending a choice school while controlling for individual fixed-effects.[40] Finally, for the OLS estimates I continue to control for family income, and also mother's and father's education, and the race of the student.[41] I allow a different intercept for students enrolled in a choice school, those in the choice applicant control group (i.e., those who were not selected and were never enrolled in a choice school), and those who were either selected for a choice school and did not enroll or who were enrolled at one time but had returned to the Milwaukee public schools (i.e., they received partial treatment). I also allow these groups to have different slope estimates. These categories are mutually exclusive in any test year.

Results using OLS (with Huber standard errors) and individual fixed-effects are in Table 11. I present results using only a sample of African-Americans and Hispanics (for comparison with earlier results), and using all students (not missing sex or race). In columns (1)-(4), I find that the students who were enrolled in a choice school had lower math test score intercepts than students in

---

[40] Witte, Sterr, and Thorn (1995) control for the previous year's test score which restricts the yearly growth to be constant each year. I get results more similar to his when I control for the previous year's test score.

[41] I cannot control for the applicant pool fixed effects when using the Milwaukee public school sample. The point estimates are almost identical with and without them for the sample using only the applicant pool (without individual fixed-effects). In addition, I do not include the grade level of the test of the student for comparison with earlier tables. This has very little effect on the results.

Table 11

OLS and Individual Fixed-Effects Estimates of the Effect of Private Schools on the Rate of Growth in Test Scores Using the Milwaukee Public School Sample as a Comparison Group

	Math Scores				Reading Scores			
	Minorities		All Students		Minorities		All Students	
	OLS (1)	FE (2)	OLS (3)	FE (4)	OLS (5)	FE (6)	OLS (7)	FE (8)
Enrolled in Private School	-2.458 (1.150)	-1.935 (1.039)	-2.854 (1.122)	-2.142 (0.999)	-1.029 (1.006)	0.802 (0.962)	-1.468 (0.991)	0.428 (0.935)
Choice Applicant Control Group	-1.779 (1.101)	-3.997 (5.807)	-2.016 (1.097)	-4.351 (5.676)	-0.114 (0.905)	-0.397 (5.364)	-0.761 (0.898)	-0.708 (5.304)
Selected for Choice, Not in a Private School	-0.120 (2.346)	0.636 (2.070)	-0.105 (2.340)	0.659 (2.007)	3.743 (2.147)	3.461 (1.904)	3.483 (2.140)	3.514 (1.869)
Number of years	-0.290 (0.204)	-1.142 (0.157)	-0.263 (0.177)	-0.955 (0.131)	-0.185 (0.166)	-0.760 (0.126)	-0.217 (0.144)	-0.487 (0.107)
Number of years*Private School	0.778 (0.569)	1.887 (0.454)	0.666 (0.544)	1.595 (0.429)	-0.347 (0.472)	-0.339 (0.419)	-0.312 (0.457)	-0.496 (0.402)
Number of years*Choice Control Group	-0.733 (0.473)	-0.014 (0.401)	-0.662 (0.462)	-0.181 (0.377)	-0.898 (0.423)	0.060 (0.365)	-0.838 (0.410)	-0.014 (0.347)
Number of years*Selected for Choice, Not in Private School	-0.043 (0.889)	1.319 (0.743)	-0.027 (0.877)	1.076 (0.717)	-1.819 (0.806)	-0.728 (0.680)	-1.805 (0.796)	-1.057 (0.665)
R^2	0.023	0.013	0.113	0.011	0.032	0.010	0.117	0.005
Number of Observations	8729	8729	11083	11083	8751	8751	11121	11121

Notes: Standard errors are in parentheses. The OLS standard errors allow for correlation "within" an individual. All regressions include a constant, sex, family income, race, mother's and father's education, and indicators if income and parental education are missing; for the math scores, a dummy indicating if the test score was imputed. The fixed-effects estimates include individual fixed-effects, and for the math scores, a dummy indicating that the total score was imputed. The "minorities" sample includes only African-Americans and Hispanics; "all students" includes other students, as well.

the Milwaukee public school comparison group, a difference that is statistically significant. The OLS estimates in columns (1) and (3) also indicate that students who attended a private school experienced greater math test score gains than the Milwaukee public school comparison group, although the increase is not statistically significant and less that one percentage point per year.

On the other hand, the fixed-effects models (in columns (2) and (4)) suggest that the math scores of students in private schools increased an additional 1.6-1.9 percentage points per year, and the effect is significant at the 5 percent level for both samples of students.[42] At the same time, the math test scores of those who received partial treatment increased an additional 1.3 percentage points per year above other students in the Milwaukee public schools although the difference is not statistically significant (the p-value of the difference between those who received full and partial treatment is 0.079 in column (2) and 0.13 in column (4)).

The results for reading, in columns (5)-(7), suggest that students in the choice schools score lower than students in the Milwaukee public schools by about -0.3 percentage points per year, although this difference is not statistically significant. On the other hand, the differences between the applicant control group, those who received only partial treatment, and the Milwaukee public school comparison sample are statistically indistinguishable from zero. Including individual fixed-effects does not make a large difference to the estimated treatment effect of private school attendance, although it does decrease (or eliminate) the negative effect of partial treatment.

V. Conclusion

The results using the quasi-experimental applicant control group and the random sample of students from the Milwaukee public schools (when I include individual fixed-effects) are remarkably similar. On the one hand, I find that the Milwaukee Parental Choice Program and the participating

[42] These results appear robust to alternative methods of dealing with the imputation of the total math score and sample attrition.

private schools likely increased student math scores by 1.5-2 percentage points per year. On the other hand, the results for reading scores were quite mixed with both positive and negative coefficient estimates. While it is not clear how to weigh the math results versus the reading results, when I total the math and reading scores, I estimate that private school students gained approximately 1.3 percentage points per year and the effect has a p-value of 0.063.[43]

Although these results obtain using a variety of estimation strategies and samples, there are at least two caveats to keep in mind. First, I had to impute the total math score for a significant fraction of the Milwaukee public school students. Second, and most importantly, there is substantial sample attrition in the later years. While my analysis suggests that the results are robust to the imputation and sample attrition, my strategies cannot substitute for better data.

The data collection from Milwaukee should be applauded as it allows us to learn more about the effectiveness of this program than from many other reforms. Nevertheless, if one lesson emerges from this effort, it is that future evaluations of reforms of this sort should anticipate high mobility among the students and recognize that administrative data are determined by the needs of the schools and not those of the evaluator. An evaluation design that treats the participants (and control or comparison group) as a survey sample with independent follow-up, though more costly, would avoid some of the data problems experienced here.

Finally, although the experience in Milwaukee suggests that providing vouchers to low income students to attend private schools could help increase the mathematical achievement of those students who participate, it cannot shed light on whether vouchers provide an incentive for the public schools to improve and therefore increase the quality of education provided to all low income children. In addition, the results from one program implemented in one city cannot, and should not, be the only evidence on which important policy regarding the structure of American education is

---

[43] Based on the full sample with all students including individual fixed effects (i.e., such as columns (4) and (8) in Table 11).

based. It is only by piecing together evidence from many places that we will ever really learn whether private school vouchers could increase student achievement.

References

Akin, John S., and Irwin Garfinkel. "The Quality of Education and Cohort Variation in Black-White Earnings Differentials: Comment," American Economic Review 70 (March, 1980): 186-191.

Angrist, Joshua D. "The Draft Lottery and Voluntary Enlistment in the Vietnam Era," Journal of the American Statistical Association, 86, no. 415 (September, 1991): 584-595.

Angrist, J.D., G.W. Imbens, and D.B. Rubin. "Identification of Causal Effects Using Instrumental Variables," Journal of the american Statistical Association, forthcoming.

Coleman, James, Thomas Hoffer, and Sally Kilgore. High School Achievement: Public, Catholic and Private Schools Compared (New York: Basic Books, 1982a).

Coleman, James, Thomas Hoffer, and Sally Kilgore. "Cognitive Outcomes in Public and Private Schools," Sociology of Education, 55 no. 2/3 (April/July, 1982b): 65-76.

Cookson, Peter W. "Assessing Private School Effects: Implications for School Choice" in School Choice: Examining the Evidence, Edith Rasell and Richard Rothstein, editors (Economic Policy Institute: Washington, D.C., 1993).

Coronary Drug Project Research Group. "Influence of Adherence to Treatment and Response to Cholesterol on Mortality in the Coronary Drug Project," New England Journal of Medicine, 33 (1980): 1038-1041.

Behrman, Jere R., and Nancy Birdsall. "The Quality of Schooling: Quantity Alone is Misleading," American Economic Review 73 (December, 1983): 928-946.

Durbin, J. "Errors in Variables," Review of International Statistical Institute, 22 (1954): 23-32.

Efron, B., and D. Feldman. "Compliance as an Explanatory Variable in Clinical Trials," Journal of the American Statistical Association 86, no. 413 (March 1991): 9-17.

Evans, William N., and Robert M. Schwab. "Finishing High School and Starting College: Do Catholic Schools Make a Difference?," Quarterly Journal of Economics 110, no. 4 (November, 1995): 941-974.

Goldberger, Arthur S. and Glen G. Cain. "The Causal Analysis of Cognitive Outcomes in the Coleman, Hoffer and Kilgore Report," Sociology of Education 55, vol. no. 2/3 (April/July, 1982): 103-122.

Greene, Jay P., Paul E. Peterson, Jiangtao Du, Leesa Boeger, and Curtis L. Frazier. "The Effectiveness of School Choice in Milwaukee: A Secondary Analysis of Data from the Program's Evaluation." University of Houston mimeo, August 1996.

Greenwald, Rob, Larry V. Hedges, and Richard D. Laine, "The Effect of School Resources on Student Achievement," Review of Educational Research, 66 (in press).

Gruber, Jonathan. "Cash Welfare as a Consumption Smoothing Mechanism for Single Mothers." National Bureau of Economic Research Working Paper Number 5738 (September 1996).

Hanushek, Eric A. "Conceptual and Empirical Issues in the Estimation of Educational Production Functions," Journal of Human Resources, 14 (Summer, 1979): 351-388.

Hausman, Jerry A. and David A. Wise. "Technical Problems in Social Experimentation: Cost versus Ease of Analysis," in Social Experimentation, Jerry A. Hausman and David A. Wise, editors, (The University of Chicago Press: Chicago, 1985).

Heckman, James J. "Instrumental Variables: A Cautionary Tale," National Bureau of Economic Research Technical Working Paper 185 (September 1995a).

Heckman, James J. "Randomization as an Instrumental Variable," National Bureau of Economic Research Technical Working Paper 184 (September 1995b).

Heckman, James, and Jeffrey Smith. "Assessing the Case for Randomized Evaluation of Social Programs," in Measuring Labour Market Measures: Evaluating the Effects of Active Labour Market Policies, Karsten Jensen and Per Kongshoj, editors (Copenhagen: Ministry of Labour, 1993).

Heckman, James, and Jeffrey Smith. "Substitution Bias in Social Experiments: An Analysis of JTPA Data," University of Chicago mimeo, 1994.

Link Charles R., Edward C. Ratledge, and Kenneth A. Lewis. "The Quality of Education and Cohort Variation in Black-White Earnings Differentials: Reply," American Economic Review 70 (March, 1980): 196-203.

Manski, Charles F. "Identification of Binary Response Models," Journal of the American Statistical Association, 83, no. 403, (September, 1988): 729-738.

Manski, Charles F. "The Selection Problem," in Advances in Econometrics, C. Sims, editor, (New York: Cambridge University Press, 1992).

Manski, Charles F. Identification Problems in the Social Sciences (Cambridge, MA: Harvard University Press, 1995).

Murnane, Richard J. "A Review Essay – Comparisons of Public and Private Schools: Lessons from the Uproar," Journal of Human Resources, 19 (1984): 263-277.

Park, Jin Heum. "Estimation of Sheepskin Effects and Returns to Schooling Using the Old and the New CPS Measures of Educational Attainment." Industrial Relations Section Working Paper #338, December 1994.

Rubin, D. "Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies," Journal of Educational Psychology, 66 (1974): 688-701.

Thorn, Christopher A., John F. Witte, and Troy D. Sterr. "Documentation of Data Used in

The Milwaukee Choice Study". La Follette Institute of Public Affairs mimeo, August 1995.

Wald, Abraham. "The Fitting of Straight Lines if Both Variables Are Subject to Error," Annals of Mathematical Statistics, 11 (1940): 284-300.

Witte, John F. "Private School vs. Public School Achievement: Are There Findings That Should Affect the Educational Choice Debate?," Economics of Education Review, 11 (December, 1992): 371-394.

Witte, John F. "Reply to Greene, Peterson, and Du: 'The Effectiveness of School Choice in Milwaukee: A Secondary Analysis of Data from the Program's Evaluation'", University of Madison mimeo, August 1996.

Witte, John F. and Christopher A. Thorn. "The Milwaukee Parental Choice Program, 1990/1991-1994/1995 [computer file], Madison, WI, 1995.

Witte, John F., Christopher A. Thorn, and Kim A. Pritchard. "Private and Public Education in Wisconsin: Implications for the Choice Debate," University of Wisconsin mimeo, (1995).

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

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

Data Appendix

The data on the Milwaukee Parental Choice Program can be downloaded from http://dpls.dacc.wisc.edu/choice/choice_index_html. I construct the analysis sample according to the following definitions which are similar to those used by Greene, et. al. (1996):

Year of Application

I define the year of application as the first year in which a student applied if the student was either never selected or was selected the first time they applied. If the student applied more than once, I consider the first year they are accepted as the year in which they applied. In addition, in a few cases, the student applied, but was not selected but was nonetheless enrolled in a choice school the following spring (i.e., they were admitted off of a waiting list). I consider their year of the application the year in which they applied and were not accepted, as once a student was enrolled in a choice school they did not participate in the randomization the following years. I only include those who first applied in the years 1990-1993.

For students in the Milwaukee public school sample I consider 1990 the "year of application" for students in the Milwaukee public schools who have a valid 1991 test score. For those without a 1991 test score, I consider their "year of application" to be 1991 if they have a valid 1992 test score, and so forth.

Grade at Application

I only use the grade levels of the Iowa Tests of Basic Skills in determining the grade to which the student applied. As the test is administered in the spring semester, I consider the grade level of the test of the spring following the year of application as the grade to which the student applied. To impute the grade of application (in cases in which the student is missing a test score for the first spring following the year of application), I search backwards and forwards in the test data for a non-missing test grade and add or subtract the appropriate number of years. For example, suppose a student applied in 1990, is missing a 1991 test score, but has a test score for the spring of 1992. I consider the grade at application to be the 1992 grade minus 1. (An alternative method for imputing would be to use the grade of the student from the administrative data (the mastch.dat and chsrdb.dat files). The exact order in which one searches for a valid grade using the test score data and using the alternative method of imputation makes a small difference to the point estimates. (See footnote 18). I only include students who applied to grades K-8 in the analysis.

Race and Sex

I first determine the race and sex of the student from the master choice file (mastch.dat) files. Further, I require that the race and sex of the student be consistent in all years. (Thus, if the student appears to "change" sex or race across years, I consider the race or sex missing.) If the race or sex is missing from these files, I use the race and sex from the choice Student Record Database (SRDB) (chsrdb.dat) file. Again, I require that the race and sex not change over the data set. I only include African-Americans and Hispanics in the analysis.

Test Scores

I use the normal curve equivalent transformation of the math and reading scores. For both, I re-compute them based on the transformation in Thorn, Witte, and Sterr (1995) as several of the test scores appear to have been mis-transformed. I only use current test scores. I impute math scores for students in the Milwaukee public schools who do not have a total score, but do have a score for the problem solving component. I impute using the following equations,

$$\hat{mnce93} = 16.293 + 0.250mnpr93 + 0.023mnpr93^2 - 0.00029mnpr93^3 + 0.0000013mnpr93^4 - 0.079tsyear93$$

$$R^2 = 0.810 \qquad N = 3603$$

$$\hat{mnce94} = 9.778 + 0.265mnpr94 + 0.023mnce94^2 - 0.00031mnce94^3 + 0.0000014mnce94^4 - 0.017tsyear94$$

$$R^2 = 0.814 \qquad N = 3182$$

where mnce9X is the total math score, mnpr9X is the problem solving component, and tsyear9X is the year the test was administered. I estimate these equations using the random sample of students in the Milwaukee public schools (and I include students with test scores from previous years).

The pre-application test score is the test score from the year the student applied. I substitute the mean score for those missing pre-application test scores.

Family Income and Parental Education

Family income is from the survey administered to choice applicants in the fall and spring of each year, as well as to a random sample of the families in the Milwaukee public schools. I average the reported income from the fall and spring of the year after application (to lower the measurement error). If a student applied more than once (and was not immediately selected) and is missing information on family income, I fill in the information with the reported income in earlier years. I convert the measure to 1994 dollars and substitute the average for those missing income.

I construct parental education in a similar fashion. I convert the categorical variable to a continuous education measure according to the mapping suggested by Park (1994): 8th grade or below=8, some high school=10, GED and high school graduate=12, some college=13, 4-year degree=16, and post-graduate work=17.

Choice Students

I consider an individual a choice student who attended private school if they took an achievement test in a choice school. In the analysis, I exclude a few students who only had valid test scores from a Milwaukee public school but who were indicated as enrolled in a choice school in the same spring. There are also 6 students who appear to be enrolled in a choice school in a particular spring and also appear to apply for the first time in the same spring; and there are 2 students who apply in 1990 or 1991, only appear to be enrolled in a choice school in 1993, and who have test scores from both the choice school and a Milwaukee public school in 1993. I include these students, although the results are quite similar when I exclude them.

Appendix Table 1a

A Comparison with Greene, Peterson, et. al.'s Math Score Results

	Math Scores After 1 Year			Math Scores After 2 Years		
	(1)	(2)	(3)	(1)	(2)	(3)
Attended Private School	-0.372 (1.791)	-0.808 (1.743)	-0.766 (1.720)	-0.082 (1.860)	1.205 (1.817)	1.446 (1.811)
Female	0.976 (1.407)	0.863 (1.376)	0.707 (1.295)	0.026 (1.538)	-0.565 (1.541)	-0.133 (1.389)
Selected, but Did Not Enroll in Program			3.609 (3.734)			6.095 (3.784)
Left Private School			3.777 (3.114)			-0.120 (2.359)
R^2	0.155	0.156	0.154	0.098	0.093	0.087
No. of Observations	772	804	887	584	608	752
	Math Scores After 3 Years			**Math Scores After 4 Years**		
	(1)	(2)	(3)	(1)	(2)	(3)
Attended Private School	4.086 (2.535)	4.515 (2.409)	3.934 (2.479)	9.519 (4.653)	10.421 (4.070)	10.277 (3.866)
Female	0.206 (1.919)	0.398 (1.898)	0.207 (1.665)	1.458 (3.957)	1.894 (3.622)	2.391 (2.810)
Selected, but Did Not Enroll in Program			5.462 (4.767)			17.780 (13.570)
Left Private School			1.544 (2.667)			11.274 (4.031)
R^2	0.217	0.201	0.158	0.130	0.172	0.150
No. of Observations	300	313	447	112	123	187

Notes: Standard errors are in parentheses. All specifications include "application pool" fixed effects. "Attended Private School", "Selected, but Did Not Enroll in Program", and "Left Private School" are mutually exclusive dummy variables. "Year" refers to the year of application.

Column Specifications:

Column (1): Replication of Greene, et. al.'s results (i.e., Greene, et. al's sample construction and sample selection rules)

Column (2): Greene, et. al's results using Rouse's sample (i.e., Rouse sample construction and Greene et. al's selection rules)

Column (3): Add "non-compliers", and those who left the Private Schools (and returned to Milwaukee Public Schools) back into the Rouse sample (column (2))

Appendix Table 1b

A Comparison with Greene, Peterson, et. al.'s Reading Score Results

	Reading Scores After 1 Year			Reading Scores After 2 Years		
	(1)	(2)	(3)	(1)	(2)	(3)
Attended Private School	1.212 (1.568)	0.534 (1.547)	0.074 (1.546)	0.875 (1.663)	1.023 (1.593)	1.074 (1.606)
Female	2.287 (1.274)	2.265 (1.258)	2.088 (1.194)	3.069 (1.376)	2.766 (1.355)	2.840 (1.237)
Selected, but Did Not Enroll in Program			-0.615 (3.327)			4.635 (3.537)
Left Private School			4.212 (2.819)			-0.167 (2.075)
R^2	0.121	0.115	0.101	0.080	0.085	0.060
No. of Observations	734	763	846	594	617	763
	Reading Scores After 3 Years			**Reading Scores After 4 Years**		
	(1)	(2)	(3)	(1)	(2)	(3)
Attended Private School	3.445 (2.344)	3.359 (2.213)	2.945 (2.247)	4.603 (3.730)	6.525 (3.388)	6.128 (3.285)
Female	5.485 (1.789)	5.271 (1.760)	4.421 (1.515)	6.847 (3.234)	6.190 (3.056)	5.726 (2.395)
Selected, but Did Not Enroll in Program			1.245 (4.194)			-26.879 (16.015)
Left Private School			0.594 (2.416)			5.648 (3.406)
R^2	0.146	0.148	0.099	0.203	0.228	0.194
No. of Observations	301	314	453	112	123	187

Notes: Standard errors are in parentheses. All specifications include "application pool" fixed effects. "Attended Private School", "Selected, but Did Not Enroll in Program", and "Left Private School" are mutually exclusive dummy variables. "Year" refers to the year of application.

Column Specifications:

Column (1): Replication of Greene, et. al.'s results (i.e., Greene, et. al's sample construction and sample selection rules)

Column (2): Greene, et. al's results using Rouse's sample (i.e., Rouse sample construction and Greene et. al's selection rules)

Column (3): Add "non-compliers", and those who left the Private Schools (and returned to Milwaukee Public Schools) back into the Rouse sample (column (2))

Appendix Table 2

Means and Standard Deviations

	Year of Test Score			
	1991		1992	
	Not Selected	Selected	Not Selected	Selected
Enrolled in Private School	0.016 [0.127]	0.844 [0.364]	0.000 [0.000]	0.789 [0.408]
Math Score	37.661 [20.436]	40.326 [18.806]	37.713 [18.260]	37.194 [18.751]
Reading Score	37.678 [16.988]	42.213 [16.405]	36.442 [17.862]	37.817 [16.440]
Female	0.565 [0.500]	0.523 [0.501]	0.496 [0.502]	0.547 [0.498]
Family Income	12.436 [6.649]	12.786 [6.560]	12.331 [5.973]	12.323 [5.969]
Family Income Missing	0.565 [0.500]	0.394 [0.490]	0.626 [0.486]	0.409 [0.492]
Pre-application Math Test Score	38.532 [14.986]	39.029 [13.536]	37.534 [10.619]	38.036 [11.403]
Pre-application Math Score Missing	0.484 [0.504]	0.514 [0.501]	0.539 [0.501]	0.611 [0.488]
Pre-application Reading Test Score	39.111 [10.227]	38.996 [10.155]	36.997 [9.321]	37.651 [9.700]
Pre-application Reading Score Missing	0.500 [0.504]	0.500 [0.501]	0.548 [0.500]	0.617 [0.487]
Black	0.661 [0.477]	0.863 [0.345]	0.765 [0.426]	0.796 [0.403]
Hispanic	0.290 [0.458]	0.110 [0.314]	0.191 [0.395]	0.181 [0.385]
Grade of Application	3.935 [2.374]	3.670 [2.290]	3.409 [2.438]	3.130 [2.256]
Applied in 1990	1.000 [0.000]	1.000 [0.000]	0.600 [0.492]	0.498 [0.501]
Applied in 1991	NA	NA	0.400 [0.492]	0.502 [0.501]
Number of Observations	62	218	115	470

Appendix Table 2 (continued): Means and Standard Deviations

	Year of Test Score			
	1993		1994	
	Not Selected	Selected	Not Selected	Selected
Enrolled in Private School	0.033	0.742	0.007	0.734
	[0.179]	[0.438]	[0.085]	[0.442]
Math Score	36.255	39.725	39.785	41.582
	[17.814]	[17.494]	[19.401]	[18.976]
Reading Score	36.279	36.242	36.969	37.845
	[17.534]	[15.177]	[18.545]	[15.742]
Female	0.434	0.523	0.485	0.543
	[0.497]	[0.500]	[0.502]	[0.499]
Family Income	12.093	12.269	12.139	12.428
	[6.051]	[6.049]	[5.486]	[6.318]
Family Income Missing	0.533	0.389	0.536	0.385
	[0.501]	[0.488]	[0.501]	[0.487]
Pre-application Math Test Score	37.122	38.789	40.859	38.859
	[12.128]	[11.583]	[15.308]	[10.691]
Pre-application Math Score Missing	0.513	0.612	0.449	0.659
	[0.501]	[0.488]	[0.499]	[0.474]
Pre-application Reading Test Score	36.089	37.715	38.506	37.888
	[10.956]	[9.787]	[11.820]	[9.609]
Pre-application Reading Score Missing	0.493	0.607	0.449	0.658
	[0.502]	[0.489]	[0.499]	[0.475]
Black	0.783	0.769	0.797	0.754
	[0.414]	[0.422]	[0.404]	[0.431]
Hispanic	0.171	0.205	0.145	0.219
	[0.378]	[0.404]	[0.353]	[0.414]
Grade of Application	3.579	2.791	3.239	2.531
	[2.292]	[2.090]	[1.928]	[2.148]
Applied in 1990	0.335	0.364	0.275	0.247
	[0.474]	[0.482]	[0.448]	[0.432]
Applied in 1991	0.250	0.430	0.152	0.311
	[0.434]	[0.495]	[0.360]	[0.463]
Applied in 1992	0.414	0.205	0.348	0.234
	[0.494]	[0.404]	[0.478]	[0.424]
Applied in 1993	NA	NA	0.225	0.208
			[0.419]	[0.406]
Imputed Test Score	0.395	0.110	0.681	0.179
	[0.490]	[0.314]	[0.468]	[0.384]
Number of Observations	152	516	138	602

Note: Standard deviations are in brackets. Based on the sample for math scores.

Appendix Table 3

The Effect of Selection to the Choice Program on Math Scores By Years Since Selection
The Effect of Alternative Imputation Strategies on the Results in Table 4a

	1993 Test Score			
	Impute Missing, No Dummy	Impute Missing	Impute All Controls	Exclude Imputed
	(1)	(2)	(3)	(4)
			1993 Test Score	
Number of years since selection	0.930	1.116	0.681	-0.015
	(0.761)	(0.798)	(0.780)	(0.985)
R^2	0.227	0.228	0.197	0.198
Number of Observations	668	668	668	551
	1994 Test score			
Number of years since selection	1.492	1.685	1.228	1.010
	(0.639)	(0.689)	(0.680)	(1.120)
R^2	0.293	0.293	0.279	0.283
Number of Observations	740	740	740	538

Notes: Standard errors are in parentheses. See notes to Table 5 for other regressors. Columns (2) and (3) include dummy variables indicating if the test score was imputed.