

How Vouchers Could Change the Market for Education

Derek Neal

Vouchers play a small role in the U.S. education system, and there is only a small probability that vouchers will, in the foreseeable future, be adopted on a large scale in any school district or state. At the moment, only three states have adopted publicly funded voucher programs, and all of these programs are limited in scale. Nonetheless, in recent years, over 40 state legislatures have considered proposals to provide some type of voucher program or tuition tax credit for at least some families in a given state or school district, and the Children's Scholarship Fund (see <http://www.scholarshipfund.org>) has helped create or expand privately funded voucher programs in numerous cities.

As vouchers become a more significant part of education policy debates, the time is right to consider what we know and do not know about the likely effects of adopting various voucher schemes. In the balance of this paper, I describe both empirical and theoretical work on education that speaks to this topic. I argue that we cannot confidently predict the outcomes that would result from various voucher schemes, and I also stress that debates over vouchers per se are not informative. Details concerning funding, targeting and discretion in the use of vouchers should greatly affect the outcomes associated with any particular voucher program. Still, empirical evidence on the performance of public versus private schools as well as numerical results from existing simulation models suggest that policymakers should be able to design voucher programs that would be helpful to minority students in large cities. Public schools in large cities often perform poorly in minority neighborhoods, and this is especially true in economically disadvantaged areas.

A starting point for many voucher advocates is the belief that private schools

■ *Derek Neal is Associate Professor of Economics, University of Chicago, Chicago, Illinois, and Research Associate, National Bureau of Economic Research, Cambridge, Massachusetts.*

foster higher levels of achievement and attainment than public schools, and therefore, I begin by reviewing the existing literature on the relative performance of public and private schools. However, achievement and attainment differences among public and private schools are only a small part of a much bigger picture. Many of the most important potential consequences of large-scale voucher programs involve changes in who enters the teaching profession; how teachers and students allocate time and effort to various tasks; and how students sort themselves into schools, classrooms and neighborhoods. These changes are of interest per se and may have important consequences beyond their effects on test scores and graduation rates. After discussing these issues, I conclude by describing agency problems that might arise between taxpayers and schools under vouchers. Opponents of vouchers often contend that private schools cannot be held accountable to taxpayers for how they use public funds. I explore two distinct versions of this hypothesis and discuss their merits.

The Relative Performance of Public and Private Schools

Private Secondary Schools

In 1982, Coleman, Hoffer and Kilgore (1982) created a firestorm in the education research community with the release of a report on the first wave of data collected by the National Center for Education Statistics for its High School and Beyond Survey of 1980 (HSB). The data include more than 10,500 students who were high school seniors in 1980 and over 11,700 students who were sophomores in 1980. Their analyses focus on differences in achievement between public and private school students who come from similar family backgrounds. They conclude that public schools are inefficient and that private schools do a better job of fostering achievement. In a follow-up study, Coleman and Hoffer (1987a) argue that the effects of private schooling on attainment as well as achievement growth are significant, and they focus specifically on the role of Catholic schools within the private sector. Coleman's critics argue that the observed positive relationship between Catholic schooling and academic performance likely reflects the common influences of unmeasured student and family traits that simultaneously affect school choice and academic outcomes. Evans and Schwab (1995) provide one of the most recent and complete analyses of the HSB data. Although their methods differ from those of Coleman, they conclude that the HSB data do provide credible evidence that Catholic schooling raises educational attainment relative to public schooling. Neal (1997b) provides a summary of studies that employ the HSB data.

Two other data sources provide evidence on the relative effectiveness of public and private secondary schools. In Neal (1997a), I use data from the National Longitudinal Survey of Youth 1979 (NLSY79) to examine the link between Catholic schooling and educational attainment. This survey began with more than 12,000 young men and women who were 14–22 years old when they were first surveyed in 1979. The respondents were interviewed annually through 1994 and every other

year since then. My results add an important caveat to the previous findings of Coleman and Hoffer (1987a), as well as Evans and Schwab (1995). I find that attainment gains associated with Catholic schooling are much larger among urban African-Americans and Hispanics than with any other group. Using the public school graduation rate for urban African-Americans and Hispanics as a baseline, my estimates of the effect of Catholic schooling imply at least a 26 percentage point increase in high school graduation rates. Further, while I find modest attainment gains among urban whites, I find little evidence of gains from Catholic schooling among suburban students, regardless of race.¹ I estimate probit models of high school graduation that include controls for student characteristics as well as a dummy variable for attending a Catholic high school. I also estimate bivariate probit models that include both a high school graduation equation containing an indicator variable for Catholic school attendance and a second equation that describes the determinants of Catholic school attendance. I use measures of Catholic school availability and the density of Catholic adherents at the county level as proxies for geographic variation in the cost of attending Catholic schools, and I include these variables in the Catholic school attendance equation, but excluded them from the graduation equation. The reasoning is that transportation costs and the availability of scholarship funds affect the costs of attending a Catholic school, but these factors should not be correlated with individual attainment levels among students with similar backgrounds. I find no evidence from the two-equation model that positive selection into Catholic schools on unmeasured student traits affects the estimated impact of Catholic schooling in the single-equation graduation model.

The National Educational Longitudinal Survey of 1988 (NELS88), also collected for the National Center for Educational Statistics, is another useful data source for public versus private school comparisons. NELS88 follows a group of roughly 25,000 students who were eighth graders in the spring of 1988 through the spring of 1994. The data provide detailed records of academic performance and numerous measures of potential student behavior problems during and before eighth grade. Thus, the data allow analyses of secondary school outcomes conditional on detailed measures of student traits and achievement at the end of elementary school. The public school graduation rate among urban minorities in this sample is 0.75. Given the complete set of controls for family background and eighth grade school records, Grogger and Neal (2000) find that Catholic schooling raises graduation rates among urban minorities by 18 percentage points on average. As in Neal (1997a), this estimated effect of Catholic schooling among urban minorities is quite large compared to the estimated Catholic schooling effects

¹ Evans and Schwab (1995) also find larger gains from Catholic schooling in urban areas, but they did not perform separate analyses on minorities in urban areas. Evans and Schwab's HSB results describe graduation rates conditional on being in school during the spring of tenth grade. The NLSY79 data and the NELS88 data described below show that a number of minority students in urban public schools do not complete tenth grade.

among other demographic groups.² The same pattern holds with respect to estimated effects of Catholic schooling on college attendance. Attainment gains associated with Catholic schooling are much smaller among urban whites, and regardless of race, our results among suburban students do not indicate any consistent pattern of attainment gains from Catholic schooling.

We cannot rule out the possibility that these NELS results for urban minorities in Catholic schools are driven by selection on some yet unmeasured traits of students or families that choose Catholic schools.³ However, readers should not presume that private school students must be positively selected. Some might argue that parents who are willing to pay private school tuition are necessarily revealing something positive about unmeasured aspects of their family that should enhance academic achievement and attainment, but we show that among suburban whites, there is no evidence that students who attend the most expensive private schools actually achieve more, graduate more often or attend college at higher rates than students from similar backgrounds with comparable eighth-grade records. Urban whites who attend elite private schools do not achieve more or graduate more often than do public school students with comparable eighth-grade records, although they do attend college at higher rates than their public school counterparts do.

The National Association of Independent Schools (NAIS) is the largest accrediting agency for private schools. Most NAIS schools are secular in orientation and also more expensive than Catholic schools. The NAIS school students in the NELS88 data have twelfth-grade achievement scores and college attendance rates that are much higher on average than those of white students in either public or Catholic schools. Nonetheless, our analyses provide little evidence that NAIS secondary school students learn more, drop out less or attend college more often than public school students from similar backgrounds who began high school with similar academic records, and among suburban students, the higher attainment and achievement levels observed among students attending NAIS high schools may be attributed entirely to the family and academic backgrounds of NAIS students prior to high school.

These results suggest a puzzle. Why would suburban parents choose expensive private schools that yield no measurable benefits in terms of attainment or achieve-

² These effects are not directly comparable in magnitude to the effects reported in Neal (1997a). Here, the 18 percentage point effect is an average effect taken over all students in the sample. The Neal (1997a) numbers are calculated using the public school graduation rate as a fixed baseline. However, the estimated coefficients on Catholic schooling in the probit models of high school graduation for NELS88 and NLSY79 are almost identical.

³ Coleman and Hoffer (1987a), Evans and Schwab (1995) and Neal (1997a) all use different methods to correct for selection into Catholic schools, and each reports evidence that with respect to traits not captured by HSB and NLSY79 variables, students are negatively selected into Catholic schools. Grogger and Neal (2000) also find weak evidence of negative selection into Catholic schools among urban African-Americans and Hispanics. However, the NELS88 data show that urban minority students in Catholic secondary schools are positively selected on their eighth-grade records, holding constant their family backgrounds. See Altonji, Elder and Taber (2002) for an extensive treatment of methods for addressing selection bias in the estimation of gains from private schooling.

ment? Obviously, NAIS parents may be purchasing a vector of outcomes that is broader than simple attainment and achievement measures. However, it is also likely that wealthy families who send their children to public schools are receiving the very best that public school systems have to offer. Magnet schools with restrictive admissions policies or regular schools with attendance zones that coincide with the boundaries of wealthy neighborhoods function much like elite private schools. Children are more likely to gain admission to magnet schools if they are gifted, and often only the wealthiest families can afford housing in the most desirable school districts or attendance zones.

Voucher Programs and Private Elementary Schools

National survey data provide much less information about the relative performance of public and Catholic school children in elementary schools. However, there is a recent literature on private voucher programs that shows how achievement growth among economically disadvantaged African-American children in cities rises if they are able to attend private schools instead of public schools.

Howell and Peterson (2002) describe the outcomes of several evaluations of privately funded voucher programs. Here, I discuss the results for evaluations of programs in Dayton, Ohio, New York City and Washington, D.C. Each of the studies followed the same basic research design. Survey teams drew treatment and control samples from the lists of families who applied for vouchers and were eligible to receive them. Eligibility rules involved city residence requirements and financial need calculations. Both the treatment and control samples took baseline achievement tests prior to group assignment, and the survey teams collected baseline information about the families, as well. In Dayton, researchers conducted two yearly rounds of follow-up testing and surveys. In New York and Washington, D.C., the survey teams collected three years of data. The baseline testing began in 1997 in New York and in 1998 in Washington, D.C., and Dayton. The teams collected three years of follow-up data in New York and Washington, D.C., but only two years of follow-up data in Dayton.

Howell and Peterson (2002) present separate analyses for white, Hispanic and African-American students based on responses to baseline questionnaires. All families who identify themselves with a Spanish-speaking group (for example, Dominicans) are included in Hispanic category. They find little evidence that private schooling helps or hurts white or Hispanic students, although it is worth noting that the sample of African-American children is the only sample that contains significant numbers of students in all three cities. The vast majority of white students in the samples are in Dayton, and almost all of the Hispanic students are in New York. They do find evidence that African-American children in these cities benefit from access to private schools. Pooling African-American children at all grade levels in all three cities, Howell and Peterson find that private schooling enhances one- and two-year total achievement gains by 3.9 and 6.3 percentiles, respectively. Further, in New York and Washington, D.C., the comparable three-year gain estimate is 6.6 percentiles.

These estimated effects reflect gains on a composite math and reading score, and they are large by any metric. The three-year gain figure is roughly equivalent to one-third of the overall test score gap between African-American and white students on national achievement tests. In total, the Howell and Peterson (2002) results compliment the findings of Neal (1997a) and Grogger and Neal (2000) concerning the performance of Catholic secondary schools. Since most voucher recipients who use their vouchers are attending Catholic elementary schools, the Howell and Peterson results provide additional evidence that African-American students in cities benefit from access to Catholic schools.⁴

Yet, I have some reservations about the Howell and Peterson (2002) data. The magnitude of the estimated effects varies considerably among grade levels and among cities at a point in time. Table 6-4 of Howell and Peterson provides estimated private school effects for African-American students that are broken out by grade level. For the one-, two- and three-year gain measures, there are considerable differences among grades. As an example, the estimated three-year gains range from 2.9 percentiles in fourth grade to 12.4 percentiles in fifth grade. There are also differences in estimated effects over time within cities. Aggregating all grades together but treating the samples from each city separately, Howell and Peterson find that the estimated gain from private schooling among African-American children in Washington, D.C., is quite variable over time. The one-, two- and three-year estimated gains are -0.09 , 9.2 and -1.9 percentiles, respectively. Only the 9.2 figure is statistically significant. This pattern raises questions about data collection, and two aspects of the survey design are noteworthy. To begin, there is considerable attrition from the samples, and the levels of attrition vary among cities. Year two and three response rates were roughly constant within cities, but varied among cities with rates for New York, Washington, D.C., and Dayton of roughly 66, 60 and 50 percent, respectively. Even though the levels of attrition are high and variable, Howell and Peterson do show that on most dimensions, attrition is approximately random with respect to the baseline characteristics of the treatment and control groups.

However, attrition is not random with respect to past achievement growth. In New York, achievement growth during the first two years lowers the probability of attrition in the treatment sample, but raises the probability of attrition in the control sample. A similar pattern holds for year-one achievement growth and year-two response patterns in Dayton. Nonetheless, these correlations are weak, and

⁴ An earlier set of studies debates the effects of access to secular private elementary schools during the initial years of the Milwaukee voucher program, which in its early years did not permit the use of vouchers in Catholic or other religious schools. Rouse (1998) provides a fairly complete analysis of the data and a balanced assessment of the competing claims in previous studies. She concludes that access to private schools in Milwaukee led to notable improvements in math but not reading achievement. Nationwide, only a tiny fraction of urban minority students in private schools attend secular schools.

Howell and Peterson conclude that systematic attrition likely has small effects on their results.⁵

Can the Urban Catholic School Results be Widely Replicated?

Although the literature on the performance of public versus private schools does not indicate that existing private schools outperform public schools across the board, there is evidence that public schools in cities may perform poorly compared to their private school neighbors. The most compelling evidence that private schools yield real benefits comes from data on the experiences of minority students in cities, especially African-American students, who gain access to Catholic schools. Still, this evidence does not directly imply that large-scale voucher plans would generate broad-based improvements in achievement and attainment among urban minorities.⁶

It is possible that estimated gains from Catholic schooling arise from peer effects and not from better school performance. When a student with a voucher or a scholarship attends a private school, this student may enter classrooms that contain very different peers than she knew in public school. Such a student could experience gains in achievement and attainment because she has better peers to emulate or because the new peers are less disruptive in class. If peer effects are key to the success of urban Catholic schools, then the effects of large-scale voucher plans on educational outcomes would be determined completely by how such plans change the sorting of students into schools and classrooms. I return to the issue of sorting below, but for now, I note that there is mixed evidence in the attainment and achievement literature concerning whether or not peer effects are important, and in chapter 5 of their book, Howell and Peterson (2002) found only minor differences between public and private school students in the reported characteristics of their friends at school. This is true at baseline and in follow-up surveys.

Yet, even if we stipulate that urban Catholic schools in minority communities perform better than their public school counterparts, we cannot be sure that vouchers will generate a large number of high-performing private schools in these communities. If a specific city adopted a large-scale voucher plan, many families and students who do not match well with either the religious culture or educational philosophy of Catholic schools would likely receive a voucher. These students may not benefit from access to Catholic schools, and we cannot predict what other types

⁵ See Appendix A of Howell and Peterson (2002) for details. The links between achievement growth and future attrition raise some concerns because they are expected given the survey design. For treatment group members, the main benefit of showing up for testing is maintaining the option to keep using the voucher. For controls, the main benefit is remaining eligible for future voucher lotteries. Thus, holding baseline characteristics constant, treatment group members who find a private school that suits them well have a relatively strong incentive to show up for testing. Among control group members, those who see their public school options grow worse after the baseline period also have relatively strong incentives to show up for testing.

⁶ There is some evidence on the effects of large-scale vouchers in other countries. Angrist et al. (2001) provide results for Columbia. McEwan (2001) provides an analysis of the program in Chile.

of private schools would emerge or how well these schools would perform.⁷ Even if we could assume that most urban minority students would benefit individually from access to the existing stock of Catholic schools, we still would not know what it would cost to replicate these schools on a large scale, because we do not really know why urban Catholic schools succeed, and we know little about the supply of teachers and administrators who are willing and able to work in Catholic schools. In sum, while the data clearly indicate that urban minority students who attend Catholic schools benefit from the experience, we cannot predict how particular voucher systems would change the overall distribution of education outcomes in cities. Nonetheless, economists are gaining insights concerning the likely qualitative effects of vouchers, at least on some dimensions.

Who Would Teach?

Arguments over vouchers sometimes proceed as if their only effect would be to redistribute students among a very similar group of teachers. But some of the most important potential outcomes of adopting vouchers involve likely changes in the labor market for teachers. These changes may have important impacts on who enters the teaching profession as well as how teachers allocate their efforts among the multiple tasks that comprise the job of teacher.

We all know from our own experience as students that some teachers are clearly and persistently better than others, and this is true both among and within schools. This observation is backed up by a literature on “teacher effects” summarized in Hanushek (2002). For example, Rivkin, Hanushek and Kane (2001) examine panel data that follow cohorts of children through Texas public schools. These data provide multiple observations on achievement for individual students as they pass through different schools and work under different teachers. Rivkin, Hanushek and Kane find that the identity of a child’s teacher in a given year is an important determinant of achievement growth even among students in the same school. The “teacher effects” literature supports the notion that a crucial task in managing a school is identifying, retaining and motivating talented teachers.

However, personnel systems that govern the pay and promotion practices of public schools are quite bureaucratic and rigid. While public employees in all areas of government demand employment regulations to protect them from supervisors who may be motivated by nepotism or political considerations, the personnel systems that govern compensation among public school teachers are inflexible even by government standards. Ballou and Podgursky (2002) point out that the Federal General Schedule (GS) system, which governs pay and promotion for federal workers, is much more flexible than the personnel systems used by public school districts. This is true even within the GS level that corresponds to the skill and

⁷ See Grogger and Neal (2000) for a more complete discussion of this point.

education level of teachers. In earlier work, Ballou and Podgursky (1995, p. 55) summarize public school personnel systems this way: “Virtually all public school districts use salary schedules to determine teacher compensation. A schedule is essentially a grid specifying salary as a function of experience and education (degrees or credits). All teachers employed in a district, regardless of grade level or subject matter, are paid on this basis.”

Thus, all teachers in the same school district who began teaching in the same year and have the same number of total credits make the same wage. This holds true regardless of their past performance, whether or not they teach math or English, or whether they teach five year-olds or teenagers. Such pay schedules are striking because teaching a different subject or students in a different age group often requires very different skills and knowledge.

Although wage compression alone does not imply inefficiency, a number of empirical results in the literature on teacher labor markets suggest that public school personnel policies are inefficient and, further, that large-scale vouchers might create important changes in the practices that govern hiring, promotion and pay among teachers.⁸ First, it is common for school districts simultaneously to face persistent shortages of applicants in some fields while rationing jobs in others. Consider a school district that, given its own hiring standards, has trouble finding qualified applicants for math and science positions, and further assume that this district faces queues for positions in the humanities. It is reasonable to suspect that the district could increase the quality of its math and science faculty while sacrificing little or nothing in terms of the quality of its humanities faculty by raising wages for math and science teachers while simultaneously reducing wages for humanities teachers.

Hoxby (2002) provides strong suggestive evidence that charter schools and private schools do deviate from public school salary schedules in just this manner. (Charter schools are public schools set up by a group of teachers and parents. They accept students by application or lottery and are allowed freedom to experiment with different educational formats.) These schools hire faculty with more math and science credits than do neighboring public schools, and in these sectors, wages are more positively related to an individual teacher’s background in math and science. These results are stronger for charter schools than for private schools, and the comparison between public and charter schools is quite compelling, since public schools and charter schools within the same districts often receive comparable funding per student.

⁸ Wage compression is not proof, in and of itself, that salary schedules are inefficient. Imagine an education production function that specifies the final human capital of a high school graduate as a Cobb-Douglas function of the quality of instruction in each of his classes. If this production function contained equal coefficients on each input, the optimal policy for schools would be to spend the same amount on teacher quality in each classroom. If “good” math teachers cost more than “good” English teachers on some efficiency units scale, the optimal policy given a Cobb-Douglas production function of this type would involve hiring better English teachers than math teachers and paying them all a constant salary.

Further, given the results from the “teacher effects” literature, it is quite striking that public schools pay teachers in a given subject the same wage, conditional on seniority and credentials, regardless of past job performance. Hanushek (2002) makes three points about the literature on teacher quality that are important when considering the absence of pay for performance among public school teachers. First, as I note above, a number of studies find considerable differences among teachers in their ability to foster achievement among students. Second, observed characteristics of teachers that might be found on a resume or in a school database explain little of measured performance differences among teachers. Finally, principals are typically aware of the quality differences among their teachers. Without running any regressions, principals are still able to discern which teachers are doing the best job of fostering achievement.

These three results do not surprise me. Among economics departments in colleges and universities, teaching performance varies greatly among faculty members who look quite similar on paper, and these performance differences are often well understood by students and faculty alike. What is striking is the fact that persistent individual differences in teaching performance do not affect compensation among public school teachers even when principals are aware of these differences.

In contrast, Ballou and Podgursky (1995, chapter 6) present evidence that private schools employ a much more flexible wage structure. Even when private schools have a stated pay schedule that resembles a public school salary system, deviations from these schedules are common. Further, while it is not possible to determine whether or not private schools are actually paying more to high-performing teachers, salaries in private and charter schools do vary with the few measured characteristics of teachers that are correlated with measured performance. Hanushek (2002) reports that teachers with stronger academic records and test scores do perform better in the classroom, and Hoxby (2002) shows that compared to salaries in the public sector, salaries within the private and charter school sectors vary more with the SAT scores of teachers and the quality of a teacher’s undergraduate institution. Hoxby also shows that charter and private school teachers come from more selective colleges and report higher test scores than teachers in neighboring public schools. Taken in isolation, these hiring patterns do not tell us whether or not teachers with excellent academic records are less likely to apply for teaching jobs in the public sector, or whether they are less likely to be hired in public schools than private and charter schools given the pool of other applicants, or both. However, Ballou and Podgursky (1995) and Ballou (1996) provide evidence that public schools give relatively little weight to the overall quality of an applicant’s academic record when making hiring decisions.

The rigid wage structures among public schools also raise questions about teacher retention. Murnane and Olsen (1989) and (1990) as well as Stinebrickner (2001) analyze separation rates among public school teachers. All three studies find that teachers with higher test scores and better college records leave their jobs at higher rates. This evidence is not enough to prove that public schools are failing to

retain their best teachers. Indeed, it is very difficult to make firm claims concerning whether or not public schools are attracting and retaining persons who should be teachers from the perspective of market efficiency. We have no data on the potential performance of people who are not teaching, and teacher turnover studies do not contain complete measures of job performance. Nonetheless, it is clear that charter and private schools make distinctions among teachers that are not permitted within many public school pay systems and also that public schools lose a disproportionate share of teachers who have strong academic records.

Teachers, like attorneys, accountants and other professions, limit entry into their profession by establishing training and certification requirements. However, professional organizations for teachers stand alone as the only professional groups that become involved in wage setting. This involvement may arise because teaching is the only large profession consisting almost entirely of public employees, but in any event, the outcome is that teachers in public schools are paid according to personnel systems that resemble the systems governing pay and promotion for blue collar union members. It is reasonable to suspect that if a voucher system could create active competition for teachers among independently governed schools, it might have significant and positive effects on the composition of the teaching profession.

Accountability Systems as Substitutes for Markets

Defenders of the status quo may note that if public school principals and superintendents had the ability to set salaries for individual teachers, they might abuse such discretion. Nepotism and political affiliation might well influence pay setting in such a system. Without a means of holding principals directly accountable for performance in their schools, there are good reasons to be cautious about granting them discretion over the use of public funds in setting salaries for individual teachers.

In recent years, several states have adopted accountability systems that provide financial rewards for principals and teachers that are tied directly to school-level performance on standardized tests. These programs are politically appealing because test scores are commonly accepted as objective measures of school performance, and some might conjecture that these team-based incentive systems provide a substitute for the performance incentives that would be inherent in a competitive market. Accountability systems create bonuses and tournament prizes that could generate incentives for principals to hire the best teachers as well as incentives for teachers to teach well.

However, the early returns from research on accountability systems are mixed. There is evidence that accountability systems raise scores on the particular tests used to determine rewards and sanctions, but most of these measured achievement gains do not carry over to other tests that cover the same subject matter (Koretz,

2002; Klein et al., 2002).⁹ A forthcoming issue of the *Journal of Human Resources* contains several papers about the design of incentive systems for government agencies that provide basic education, vocational training or job search assistance. A common theme in several of these papers is that straightforward incentive systems based on objective standards are often problematic because objective performance measures are often easy to game (Baker, 2002; Dixit, 2002; Heckman, Heinrich and Smith, 2002; Koretz, 2002). In the case of test-based accountability systems, gaming takes the form of diverting instructional time away from valuable material that might not be covered on a standardized test toward materials that will be tested. Although the achievement gains associated with existing accountability systems in some states are noteworthy, we have no idea what costs might be involved in terms of achievement losses in subjects that are not tested as part of the accountability systems. Further, there is also evidence that test-based accountability systems can create considerable amounts of fraud and cheating.¹⁰

A common response to this complaint is that states or the federal government should simply fund external testing agencies that could develop a battery of tests that are “worth teaching to.” However, existing research provides little guidance concerning the cost of such a testing program. In terms of cost-benefit calculations, we do not know whether accountability systems built around such a testing agency would be more or less cost effective than other proposed reforms, such as changing class size.¹¹ To establish a test-based accountability system that does not invite unwanted distortions of classroom instruction, educators would have to incur substantial development costs and devote considerable student time to test-taking that might be used for additional instruction.

Further, I am willing to assert that the vast majority of parents want their children to learn many things in school that are hard to quantify. One of the first items in early parent-teacher conferences is whether a child is learning to work and play well with others, and I assume that I am not alone in desiring that my children develop intellectual curiosity, self-confidence and the ability to think critically. The process of identifying teachers who foster these outcomes necessarily involves subjective evaluations, and one of the most important potential outcomes of large-scale voucher systems is the establishment of an approximately competitive market for schools that provides strong incentives for principals to make these subjective evaluations carefully. Greater competition among schools might provide

⁹ Hanushek (2002) provides a summary of papers from a recent conference on state accountability systems. In addition, Kane and Staiger (2001; this issue) show that accountability systems in some states are so poorly designed that the formulae used for determining bonuses create a mechanical relationship between school size and the probability of winning a bonus payment. This arises because the formulae adopted by states make no allowance for the fact that average achievement gains are measured less precisely among small schools.

¹⁰ Koretz (2002) discusses this issue. Jacob and Levitt (2002) provide strong suggestive evidence that in response to recent rounds of high-stakes testing in Chicago, some teachers, principals or other school officials changed a significant number of student answer sheets.

¹¹ See Krueger (1999), Angrist and Lavy (1999) and Hoxby (2000) for recent work on the effects of changing class sizes.

a catalyst for putting principals in place who would have the talent and incentive to identify and to retain teachers who perform well on many tasks that are both hard to quantify but are highly valued by parents.

Persons in professions other than teaching rarely receive incentive pay based on some simple formula tied to an objective performance standard. Rather, they receive raises, bonuses and promotions based on the outcomes of ongoing and often subjective evaluations. The pairing of incentive pay and subjective evaluations is necessary in complex jobs where output is multidimensional and difficult to measure (Baker, 1992; Holmstrom and Milgrom, 1991), but it is difficult to imagine how this practice can become the norm in teaching as long as the vast majority of teachers are civil servants employed by school systems that face limited competition.

Sorting

From the perspective of an individual student, voucher systems are likely to affect not only the composition of teachers but also the composition of classmates.¹² We do not have data from large-scale voucher experiments in the United States. Further, even if we did have data on student sorting in a particular large-scale experiment, such data might be of limited use in determining how sorting would be different under voucher plans with different funding levels or targeting rules. Thus, I must rely heavily on results from general equilibrium theory or numerical results from computable general equilibrium models to describe expected outcomes under various voucher plans. I do not attempt to provide an exhaustive menu of results, but rather focus on two important themes.

First, there are good reasons to believe that families who now live in wealthy school districts or attendance zones with high-quality public schools would suffer welfare losses under most voucher plans. Because most voucher plans break the link between residential choice and the quality of publicly funded schools, families that now enjoy access to the best public schools might see their housing wealth, which is presently linked to local school quality, fall substantially under vouchers. Second, it is not helpful to ask how vouchers per se would affect the sorting of students among schools. Voucher plans that differ with respect to funding, targeting, tuition restrictions or regulations on admission procedures may yield quite different outcomes.

Nechyba (1999) develops a model that is quite useful for considering the possible outcomes that might occur under various voucher systems. He uses the model in several other papers, including Nechyba (2000, 2002), to produce simulation results that characterize the outcomes associated with various voucher experiments. The basic ingredients of Nechyba's model are as follows. Households

¹² Here, I will not directly address the effects of peer quality on student outcomes, but I take as given that parents care about the characteristics of their children's classmates and that the sorting of students among schools by income, ability, race and other characteristics is of interest per se.

take on discrete types that reflect their endowments of income, wealth and ability. Public school districts are divided into neighborhoods that have homogenous housing stocks. Districts differ with respect to the average quality of houses in their neighborhoods. Owning a house in a particular neighborhood yields both housing services and the option of attending public schools in the neighborhood's district. Parents care about the quality of schools, and they perceive quality as being dependent on the per pupil spending and the average ability level in their child's school. Household income and ability are assumed to be positively related. Per pupil spending in public schools is determined by majority voting at the district level. Private schools operate as clubs. In equilibrium, they consist of homogenous students and charge tuition optimally given the preferences and incomes of the families in their school.

Families choose a residence, whether or not to attend public or private schools and their vote concerning local property taxes. These choices determine the distributions of public school qualities, private school qualities and housing qualities consumed by families. Nechyba shows that there exists an equilibrium set of taxes rates, housing prices and private school fees such that no one wants to change schools, change houses or change their vote given the choices everyone else has made. This set of prices and taxes supports the equilibrium allocation of families to schools and houses in Nechyba's model. The most attractive feature of this model is that Nechyba is able to calibrate the model to existing data from several metropolitan areas, and the model is able to replicate observed patterns of residential sorting, housing prices and per pupil spending among public school districts.

Nechyba has conducted numerous simulation exercises to evaluate the impacts of various voucher schemes, and for the most part, the results appear to be sensitive to details concerning the size of vouchers, rules for financing them and any targeting provisions. However, one result is almost ubiquitous. Wealthy families in the best public school districts usually lose from the introduction of vouchers. Vouchers break the link between school quality and location, and breaking this link implies a capital loss for those who own houses in the best school districts. Even though residents of these districts still enjoy close proximity to good schools, vouchers should create other good schools in the central city that give parents who work in the city the option of quality schools for their children and a shorter commute at the same time. This type of increase in the supply of good schools can only decrease the demand for housing in the best suburban school districts. Unless voucher access is restricted to families who have lower incomes than the poorest families who currently reside in suburban districts with high-quality schools, housing prices in these districts should fall after the introduction of vouchers. This insight may help explain why suburban residents have voted against vouchers in state ballot initiatives.¹³

¹³ Brunner, Sonstelie and Thayer (2000) describe voting patterns from a 1993 California ballot initiative for statewide vouchers. Voters from the best public school districts were less likely to vote for the initiative than similar voters from other districts, and they conclude that concerns about drops in

While suburban households in good public school districts would incur losses under most voucher plans, it is important to remember that, in many cases, the details of particular voucher programs matter greatly when trying to predict outcomes. For example, consider a program that would provide modest vouchers for residents of a low-income urban district, and assume that voucher levels would fall well short of per pupil spending in public schools. This type of program is often debated in many states, and the likely outcomes of such a program vary greatly with the targeting provisions that might be included. Consider two scenarios: Vouchers could be targeted to all persons who live in a particular low-income district or just to low-income families who reside in the low-income district. The expected effects under these two voucher schemes are quite different. A plan targeted to all residents of a low-income district might well increase per pupil spending in local public schools because middle-income or wealthy families might move into the district and start their own private schools. Such families would not use resources in the public schools, but they would pay taxes and raise property values.¹⁴ On the other hand, the existence of private schools in the low-income district might lead to an exodus of the middle class from the district's public schools, which could lead to adverse impacts on public school peer quality. In contrast, a voucher targeted only to low-income residents of the low-income district might actually raise peer quality in public schools, but such a voucher would yield smaller gains in per pupil spending for the public schools because it would not attract wealthy families from other districts. This comparison is just one example of a general theme that emerges from the simulation exercises that Nechyba presents. It is a mistake to talk about *the* effects of adopting vouchers. The details of targeting and financing matter for predicting likely outcomes.

Admissions rules and tuition policies also matter, and these elements of voucher design are important determinants of student sorting among types of schools under vouchers. Epple and Romano (1998) show that vouchers may exacerbate the cream skimming phenomena that is alleged to exist in private schools already. When parents care about peer quality, they will be willing to bribe able students to attend school with their child. Private schools price discriminate in an effort to improve the average ability level of their student body, and because private schools can price discriminate and public schools cannot, vouchers might create a brain drain from public schools that would lead to an increase in total stratification by ability among schools. However, in more recent work, Epple and Romano (2002) show that it is possible to design voucher systems that increase school efficiency but do not lead to increased ability stratification. Consider a voucher plan in which schools are required to accept the exact amount of vouchers

property values contributed to this pattern. According to press accounts, suburban voters also opposed vouchers in a recent Michigan ballot initiative. However, I have not found similar academic studies of the Michigan voting patterns.

¹⁴ I assume that these new families are politically marginal in the sense that they cannot change the political equilibrium that determines education spending levels.

as tuition, and voucher amounts vary inversely with student ability. Epple and Romano show that it is possible to design a plan of this type that would result in two school sectors. A sector comprised of voucher schools would replace public schools, and each school in this sector would be completely homogenous with respect to student ability. The private school sector that existed in the absence of vouchers would continue to exist in the same form.

This result deals with an extreme case. In practice, it is not possible to implement a sliding voucher schedule based on Epple and Romano's (2002) concept of student ability. Measured academic ability is in part determined by school choice, because measured ability reflects not only a student's endowments of family background and innate talent, but also the outcomes of past schooling experiences. Nonetheless, the Epple and Romano insight is quite important. Voucher plans that are currently under consideration and others that will be proposed in the future include regulations concerning what schools may charge for tuition and also admission policies, including rules for assigning voucher students among schools when some schools face excess demand. In many large cities, urban public schools are already highly segregated by income, ability and race, and there may be several ways to implement voucher systems that would yield less overall stratification on income and ability than we observe now. Voucher plans could specify that voucher schools facing excess demand must select students randomly from their applicant pools, or voucher plans could dictate that families with less educated parents receive more generous vouchers. These regulations and others could mitigate stratification under vouchers without eliminating many efficiency gains that should arise from greater competition among schools.¹⁵

While it is difficult to predict the outcome of any large-scale voucher experiment, voucher systems targeted toward large cities with a history of public school failure may have the greatest potential for yielding large benefits. Cities like Chicago and New York have large numbers of minority students from economically disadvantaged families who attend schools where achievement and future attainment are quite low. The existing literature comparing public and private school performance suggests that many of these students may benefit from access to private schools. Further, based on the work of Nechyba (1999) and Epple and Romano (2002), I conjecture that it is possible to design targeted voucher systems that would result in better outcomes for most if not all students in large urban school districts like these.

To this point, I have ignored another important aspect of sorting. Sorting on religion receives considerable attention in education policy discussions, but little treatment in formal models of schooling markets. However, Ferreyra (2002) estimates a version of Nechyba's model that does include preferences for religious instruction. Ferreyra's results suggest that, among voucher plans that might be targeted at large metropolitan areas, plans that permit Catholic and other religious

¹⁵ See Hoxby (2001) for further discussion of voucher system design.

schools to receive vouchers could well lead to larger expansions of the private school sector and more segregation of students among schools by religious affiliation. This work and future work in this area are of particular interest given the Supreme Court decision in June 2002, in the case of *Zelman vs. Simmons-Harris*, upholding the right of parents to use vouchers in religious as well as secular schools.

Conclusion

Voucher opponents often argue that private schools cannot play a large role in reforming publicly funded education because private schools cannot be held accountable to taxpayers concerning how they use public funds.¹⁶ There are two distinct versions of this argument. The first asserts that voucher schools will waste public money either through incompetence or outright fraud. Underlying this claim is the belief that competition among schools cannot guarantee efficient use of voucher funds because parents are not able to monitor school performance effectively. However, this claim begs an important question. If one assumes that parents are not able to discern the quality of schools, should one believe that parents would get higher-quality schools from publicly employed providers who enjoy a local monopoly on service provision or from private contractors who compete with each other? The relevant question is not whether some voucher schools would misuse public funds, but instead whether vouchers would lead to an increase or decrease in the flow of public funds to ghost payrolls, ill-advised construction projects, excess support staff and other forms of waste or abuse that already plague many public school systems, especially in large cities.

Even if some voucher schools operate as for-profit firms, the answer to this question is not obvious. The returns from undetected malfeasance might be particularly high for residual claimants who operate for-profit voucher schools, but parents with imperfect information and a voucher might be more effective monitors than parents who have no means of rewarding or punishing the schools that serve their children. No existing models of agency or organizational structure provide clear guidance concerning whether misuse of public funds would increase or decrease under vouchers,¹⁷ but political interest in vouchers has clearly arisen in part from the fact that audits of large public school districts often demonstrate that some school officials do waste and even steal significant amounts of public money. A recent and striking example comes from Dallas, Texas, where an FBI

¹⁶ This theme is common in policy papers produced by the American Federation of Teachers (AFT) and the National Education Association (NEA). The broader literature on school choice contains related concerns about charter schools. Fiske and Ladd (2001, p. 77) write: "Accountability is the single greatest challenge for the U.S. charter school movement. . . . The high potential costs of a quality review process provide an additional argument for keeping charter schools on the fringe of the traditional education system."

¹⁷ Hart, Shleifer and Vishny (1997) deal with related issues surrounding the operation of prisons.

investigation into misuse of school district funds resulted in numerous criminal charges and convictions.¹⁸

A more important accountability concern arises not from the fear that voucher schools would deceive parents, but rather from the concern that voucher schools would give parents exactly what they want.¹⁹ Constitutional restrictions on actions of public employees as well as the role of median voters in determining outcomes of school board elections limit the range of political and religious views that shape educational practice in public schools, but large-scale voucher plans would give many different groups of like-minded parents the opportunity to form schools that closely match their desires and aspirations for their children. While voucher supporters cheer this expansion of parental choice and opportunity, voucher opponents argue that such a great expansion of parental control over education could create significant costs for society as a whole.

Under vouchers, some new schools might create significant social costs. Regulations and oversight could limit but not eliminate the possibility that some voucher parents would use public money to form a school that in practice, if not in name, operated as Eastside Ku Klux Klan Academy, and one can imagine other examples of potential voucher schools that might create more social costs than private benefits. Vouchers are appealing to many because they provide a means for giving parents greater control over their children's schooling. However, granting this control to *all* parents involves some risks for society as a whole. Thoughtful support for vouchers requires the judgment that these risks are worth taking.

■ *I thank Gary Becker, Dennis Epple, Eric Hanushek, James Heckman and Thomas Nechyba for helpful discussions. I thank Sarah Hamersma for excellent research assistance funded through a grant from the Institute for Research on Poverty at the University of Wisconsin-Madison.*

¹⁸ During the past decade, audits and investigations have revealed mismanagement and fraud in many large urban school districts. See (<http://economics.uchicago.edu/dneal>) for a summary of news accounts surrounding these episodes.

¹⁹ Parents who use existing voucher programs to move their children from public to private schools report higher levels of satisfaction with their new schools. Also, compared to randomly selected families with children in urban public schools, voucher families with children in urban private schools give their schools much higher performance ratings. See Howell and Peterson (2002, Chapter 7).

References

- Altonji, Joseph G., Todd Elder and Christopher Taber.** 2002. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." Mimeo, Northwestern University.
- Angrist, Joshua D. and Victor Lavy.** 1999. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics*. May, 114:2, pp. 533–75.
- Angrist, Joshua D. et al.** 2001. "Vouchers for Private Schooling in Columbia: Evidence From a Randomized Natural Experiment." NBER Working Paper 8343.
- Baker, George P.** 1992. "Incentive Contracts and Performance Measurement." *Journal of Political Economy*. 100:3, pp. 598–614.
- Baker, George P.** 2002. "Distortion and Risk in Optimal Incentive Contracts." *Journal of Human Resources*. Forthcoming
- Ballou, Dale.** 1996. "Do Public Schools Hire the Best Applicants?" *Quarterly Journal of Economics*. 111:1, pp. 97–133.
- Ballou, Dale and Michael Podgursky.** 1995. "Recruiting Smarter Teachers." *Journal of Human Resources*. Spring, 30:2, pp. 326–38.
- Ballou, Dale and Michael Podgursky.** 1997. *Teacher Pay and Teacher Quality*. Kalamazoo, Mich.: W.E. Upjohn Institute for Employment Research.
- Ballou, Dale and Michael Podgursky.** 2002. "Seniority, Wages and Turnover Among Public School Teachers." *Journal of Human Resources*. Forthcoming.
- Brunner, Eric, Jon Sonstelie and Mark Thayer.** 2001. "Capitalization and the Voucher: An Analysis of Precinct Returns from California's Proposition 174." *Journal of Urban Economics*. November, 50:3, pp. 517–36.
- Bryk, Anthony S., Valerie E. Lee and Peter B. Holland.** 1993. *Catholic Schools and the Common Good*. Cambridge, Mass.: Harvard University Press.
- Coleman, James S. and Thomas Hoffer.** 1987a. *Public, Catholic, and Private Schools: The Importance of Community*. New York: Basic Books.
- Coleman, James S. and Thomas Hoffer.** 1987b. *Public and Private High Schools: The Impact of Communities*. New York: Basic Books.
- Coleman, James S., Thomas Hoffer and Sally Kilgore.** 1982. *High School Achievement: Public, Catholic and Private Schools Compared*. New York: Basic Books.
- Dixit, Avinash.** 2002. "Incentives and Organizations in the Public Sector: An Interpretative Review." *Journal of Human Resources*. Forthcoming.
- Epple, Dennis and Richard E. Romano.** 1998. "Competition Between Private and Public Schools, Vouchers, and Peer-Group Effects." *American Economic Review*. March, 88:1, pp. 33–62.
- Epple, Dennis and Richard E. Romano.** 2002. "Educational Vouchers and Cream Skimming." Mimeo, Carnegie Mellon University.
- Evans, William N. and Robert M. Schwab.** 1995. "Finishing High School and Starting College: Do Catholic Schools Make a Difference?" *Quarterly Journal of Economics*. November, 110:4, pp. 941–74.
- Ferreira, Maria.** 2002. "Estimating the Effects of Private School Vouchers in Multi-District Economies." Unpublished Dissertation, University of Wisconsin.
- Fiske, Edward B. and Helen F. Ladd.** 2001. "The U.S. Charter School Movement: Lessons from New Zealand's Experience with Self-Governing Schools and Parental Choice," in *Charters, Vouchers & Public Education*. P.E. Peterson and D.E. Campbell, eds. Washington, D.C.: Brookings Institution, pp. 59–79.
- Grogger, Jeffrey and Derek Neal.** 2000. "Further Evidence on the Effects of Catholic Secondary Schooling," in *Brookings-Wharton Papers on Urban Affairs*. W.G. Gale and J.R. Pack, eds. Washington, D.C.: Brookings Institution, pp. 151–93.
- Hanushek, Eric.** 2002. "Publicly Provided Education," in *Handbook of Public Finance*. A. Auerbach and M. Feldstein, eds. Amsterdam: North-Holland, forthcoming.
- Hanushek, Eric and Margaret Raymond.** 2002. "Lessons and Limits of State Accountability Systems." Presented at *Taking Account of Accountability: Assessing Policy and Politics*, Harvard University, June.
- Hart, Oliver, Andrei Shleifer and Robert W. Vishny.** 1997. "The Proper Scope of Government: Theory and an Application to Prisons." *Quarterly Journal of Economics*. November, 112:4, pp. 1127–161.
- Heckman, James J., Carolyn Heinrich and Jeffrey Smith.** 2002. "The Performance of Performance Standards." *Journal of Human Resources*. Forthcoming.
- Holmstrom, Bengt and Paul Milgrom.** 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design."

Journal of Law, Economics, and Organization. 7, pp. 24–51.

Howell, William and Paul Peterson. 2002. *The Education Gap*. Washington, D.C.: Brookings Press.

Hoxby, Caroline M. 2000. “The Effects of Class Size on Student Achievement: New Evidence from Population Variation.” *Quarterly Journal of Economics*. November, 115:4, pp. 1239–285.

Hoxby, Caroline M. 2001. “Ideal Vouchers.” Mimeo, Harvard University.

Hoxby, Caroline M. 2002. “Would School Choice Change the Teaching Profession?” *Journal of Human Resources*. Forthcoming.

Jacob, Brian A. and Steven D. Levitt. 2002. “Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating.” Mimeo, University of Chicago.

Kane, Thomas and Douglas Staiger. 2001. “Improving School Accountability Measures.” NBER Working Paper No. 8156.

Koretz, Daniel. 2002. “Limitations in the Use of Achievement Tests as Measures of Educators’ Productivity.” *Journal of Human Resources*. Forthcoming.

Klien, Stephen P. et al. 2002. “What do Test Scores in Texas Tell Us?” Issue paper, Rand, April.

Krueger, Alan B. 1999. “Experimental Estimates of Education Production Functions.” *Quarterly Journal of Economics*. May, 114:2, pp. 497–532.

McEwan, Patrick J. 2001. “The Effectiveness of Public, Catholic, and Non-Religious Private Schools in Chile’s Voucher System.” *Education Economics*. August, 9:2, pp. 103–28.

Murnane, Richard J. and Randall Olsen. 1989. “The Effects of Salaries and Opportunity Costs

on Duration in Teaching: Evidence from Michigan.” *Review of Economics and Statistics*. May, 71:2, pp. 347–52

Murnane, Richard J. and Randall Olsen. 1990. “The Effects of Salaries and Opportunity Costs on Length of Stay in Teaching: Evidence from North Carolina.” *Journal of Human Resources*. Winter, 25:1, pp. 106–24.

Neal, Derek. 1997a. “The Effect of Catholic Secondary Schooling on Educational Achievement.” *Journal of Labor Economics*. January, 15:1, pp. 98–123.

Neal, Derek. 1997b. “What Have We Learned About the Benefits of Private Schooling?” *FRBNY Economic Policy Review*. March, 4:1, pp. 79–86.

Nechyba, Thomas J. 1999. “School Finance Induced Migration and Stratification Patterns: The Impact of Private School Vouchers.” *Journal of Public Economic Theory*. January, 1:1, pp. 5–50.

Nechyba, Thomas J. 2000. “Mobility, Targeting, and Private-School Vouchers.” *American Economic Review*. March, 90:1, pp. 130–46.

Nechyba, Thomas J. 2002. “Introducing School Choice into Multi-District Public School Systems,” in *The Economics of School Choice*. C. Hoxby, ed. Chicago: University of Chicago Press, forthcoming.

Rivkin, Steven G., Eric A. Hanushek and John F. Kain. 2001. “Teachers, Schools, and Academic Achievement.” NBER Working Paper No. 6691.

Rouse, Cecilia Elena. 1998. “Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program.” *Quarterly Journal of Economics*. May, 113:2, pp. 553–602.

Stinebrickner, Todd. 2001. “A Dynamic Model of Teacher Labor Supply.” *Journal of Labor Economics*. January, 19:1, pp. 196–230.