Among the many reforms proposed for K–12 education are changes in governance that would increase the power of parents to choose schools and thereby make the education system function more like a market. Within this set of reforms, which also includes offering greater choice among public schools and the opportunity to establish public “charter” schools, school voucher programs are particularly controversial because they would permit parents to use public funds to secure education not only at public schools, but also at private schools. Proponents and opponents disagree about the effects of voucher programs on student achievement, on the social and racial segregation of students and on disadvantaged students. In addition, they differ on the importance of maintaining the separation between religious private schools and the state.

School voucher programs currently exist only on a small scale in the United States. The main publicly funded voucher programs are in Milwaukee, Cleveland and Florida. In addition, small privately funded programs provide vouchers for low- and moderate-income students in cities such as New York City, Dayton, Ohio, and Washington, D.C. Another privately funded program, the Children’s Scholarship Fund, operates at the national level.

Recent studies based primarily on U.S. evidence typically conclude that the data are insufficient to draw clear conclusions about the net effects of vouchers on academic achievement, access to schools, racial integration and civic education (for example, Ladd and Hansen, 1999; Levin, 2001; Campbell and Peterson, 2001; Gill et al., 2001). In the light of the limited U.S. experience, some authors support investments in large-scale voucher experiments as a way to generate more definitive information on their effects. However, before making such investments, it would

Helen F. Ladd is Professor of Public Policy Studies and Economics, Sanford Institute of Public Policy, Duke University, Durham, North Carolina. Her e-mail address is (Hladd@pps.duke.edu).
behoove U.S. researchers and policymakers to pay more attention to the evidence from large-scale programs in other countries such as Chile and New Zealand. Chile, for example, has had a universal school voucher program since the early 1980s that has been subject to careful evaluation (Hsieh and Urquiola, 2002; McEwan and Carnoy, 2000; McEwan, 2000a; Gauri, 1998). In addition, New Zealand introduced in the early 1990s what some observers have referred to as a universal quasi-voucher system. Parents can choose any school within the public sector, which has included religious schools since the 1970s, and school funding is based largely on student enrollment (Fiske and Ladd, 2000).

In this paper, I marshal the available evidence, including the international evidence, to show that contrary to the claims of its proponents, a large-scale universal voucher program would not generate substantial gains in overall student achievement and that it could well be detrimental to many disadvantaged students. The case for a more narrowly targeted means-tested voucher program is stronger, but even with careful attention to its design, such a program should at most serve as one element of a broader strategy designed to provide more options and better education for disadvantaged students.

**Relevant Characteristics of the U.S. Education System**

Four characteristics of the U.S. education system are especially relevant to the voucher debate. First, the existing system biases parental choices toward the public sector. Second, many middle- and upper-income families currently have much more choice among schools than do low-income families. Third, K–12 education is compulsory. Fourth, parents judge the quality of schools in part by the characteristics of the students in the school.

**An Educational System with a Strong Bias Toward Public Production**

Because all families have access to free tax financed schools at the K–12 level, they face strong financial incentives to choose public over private schools. Nationally, 12 percent of students are enrolled in private schools, the bulk of which have a religious affiliation. A generous universal voucher program could potentially eliminate this bias toward the public schools.

In many other countries, privately owned and operated schools receive large amounts of direct financial assistance from the state (Plank and Sykes, forthcoming). In contrast, this country’s commitment to the separation of church and state, as embodied in the establishment clause of the U.S. Constitution, has historically kept public funding from being used to support private schools. That situation could change now that the U.S. Supreme Court has decided that there is no constitutional barrier to the use of indirect funding for religious schools through school vouchers in *Zelman v. Simmons-Harris*, a July 2002 decision relating to Cleveland’s voucher program. However, widespread extension of voucher programs to religious schools must still overcome the hurdles of establishment clauses in many state constitutions and strong political opposition in many state legislatures. If one believes that the separation of church and state in connection with
elementary and secondary education has provided substantial gains to the United States in the form of a more pluralistic and open society, then weakening this separation through the extension of voucher funding to religious schools could potentially generate large social costs. However, this tradeoff is one to which economists—or at least this economist—do not have much to contribute.

There’s another tradeoff as well, this one for the private schools themselves. The international experience indicates that widespread public funding of private schools typically brings with it greater regulation (Plank and Sykes, forthcoming). Such regulation is a natural way for the government to assure that taxpayers’ dollars are being used to promote the ends that justify public funding. To the extent that a universal voucher program in the U.S. context were accompanied by additional state control over areas such as curriculum, assessment and school admissions policies, many existing private schools could choose not to participate.

**Choice Among Schools is Available, But Limited**

Within each school district, a child typically attends the school assigned to children in that neighborhood. Hence, a family’s decision of where to live plays a large role in where a child attends school. To the extent that middle- and high-income families congregate in well-to-do suburbs, their children tend to be grouped in suburban schools, financed in part by taxes from large local property tax bases. Conversely, students whose families are restricted by their low income or by their race to economically or racially isolated areas of central cities will also end up grouped in particular schools, often those with high concentrations of disadvantaged students, insufficient resources, low average achievement and high dropout rates.

Families with sufficient income can move among suburban school districts to increase the quality of the education their children receive (Hanushek, Kain and Rivkin, 2001) or else consider private schools. Moreover, the capitalization of differential school quality into housing prices, so that homes in school districts with higher-quality schools cost more, exacerbates the effects of income differentials (Black, 1999; Figlio and Lucas, 2000). While children from low-income families in big cities do switch schools quite frequently within districts, those moves are largely driven by the vagaries of the low-income housing market and typically do not result in higher-quality schools and better educational outcomes (Hanushek, Kain and Rivkin, 2001). Hence, a carefully designed school voucher program targeted to low-income families could potentially provide low-income families with some of the freedom of school choice now available only to higher-income families.

Recent theoretical work has highlighted another potential benefit of a school voucher program, namely, the reduction in residential segregation by income (Nechyba, 1999, 2001). The notion is that if families had greater access to private schools, they would be less likely to separate their residences by income. Based on a theoretical model calibrated to New Jersey data (a state with many small school districts and over 20 percent of its children in private schools), Nechyba’s (2001) simulations indicate that access to private schools reduces the variation in incomes across his three stylized districts by more than half and substantially increases the variation in income within districts, particularly within the poorest district.
Unfortunately, the model does not incorporate racial preferences that could influence both school and residential choices. Moreover, school voucher programs would be a rather indirect policy tool for reducing residential segregation. Other options would include zoning and housing policies, as well as the expansion of parental choices among public schools so as to break the link between where families live and where their children go to school. In addition, any benefits of lower residential segregation achieved through a school voucher program would need to be weighed against less desirable outcomes that emerge from Nechyba’s stylized model. These include greater economic homogeneity within schools and lower-quality public schools in the poorest district.

Compulsory Schooling

School attendance through age 16 in the United States is compulsory. This public commitment to schooling, which reflects perceptions of large social benefits above and beyond the private benefits to the students themselves, need not require public ownership of the schools and, thus, is not inconsistent with a voucher program. Nonetheless, it does imply a much higher level of public interest and responsibility than is the case in other sectors to which K–12 education might be compared, such as higher education.

The contrast with higher education is important given the apparent success of the U.S. higher education system with its mix of public and private universities and colleges competing for students. The elite—and expensive—public and private institutions serve many students of high ability very well. At the same time, the system is highly stratified, with large differences in the academic ability of students across institutions and significant homogeneity along various dimensions, such as religion, within institutions. Moreover, the sector falls far short of serving all college-age students. In 1998, less than 40 percent of those between the ages of 18–21 were enrolled in postsecondary education institutions, and as of 1999, only 25 percent of the population aged 25 and over had completed four years of college (National Center for Education Statistics, 2001, Tables 8 and 39). At best, the U.S. system of higher education can serve as a model for the top of the ability and wealth distribution of K–12 students.

The compulsory aspect of K–12 education also distinguishes it from the private sector of the economy. Competition works in that sector in part by the expansion or replication of successful firms and the shutting down of unsuccessful firms. In K–12 education, successful schools have few incentives to expand, especially when expansion entails admitting more costly-to-educate students. Moreover, the establishment of new schools takes time and, as has become clear from the U.S. experience with charter schools, requires upfront funding for planning and capital facilities. At the same time, in a country with compulsory education, failing schools can be shut down only if there are adequate places for their students in other schools.

Role of Peer Groups in Family Choices

Parents behave as if the peers of their children matter. Evidence from studies around the world indicates that parents exercising choice seek to move their children to schools in which the average socioeconomic characteristics or nonmi-
nority share of the students is higher than it would be in their original or assigned school. This phenomenon has been documented in systems as diverse as New Zealand (Fiske and Ladd, 2000; Ladd and Fiske, 2001a), Chile (McEwan and Carnoy, 2000), Scotland (Willms and Echols, 1993) and Chicago (Cullen, Jacob and Levitt, 2000).

This behavior is consistent with many parental motivations, including the quest for better schools. For example, parents might use the socioeconomic level of the parents of other children in the school as a proxy for school quality, based on the well-documented observation that the average achievement of students within a school is highly correlated with the socioeconomic and racial composition of the student body. Not only are levels of achievement higher in such schools, so are educational gains in each grade. Data from North Carolina and other states show, for example, that the schools with larger gains in test scores are those with higher proportions of white and higher proportions of nonpoor students (Ladd and Walsh, 2002; Clotfelter and Ladd, 1996).

The positive correlation between the socioeconomic composition of a school and the performance of its students largely reflects what happens at home rather than at school. However, school-related factors may also help to explain why schools serving more affluent and nonminority students tend to exhibit larger gains in test scores than those serving more disadvantaged students. Schools serving more affluent students may benefit from positive peer or spillover effects from one student to another; they can more easily maintain educational processes, such as assigning homework; they are more able to attract high-quality teachers; and they typically have access to more resources in the form of both budgetary resources and those provided by parents in the form of contributions and volunteer activities (Fiske and Ladd, 2000).

This observation that the “customer mix” matters to parents has four substantial implications for all educational systems, including systems financed by vouchers. First, it generates a hierarchy of schools. In the context of an abstract model in which parents judge school quality in part by the average ability of the students in the school and in which private schools charge tuition, Epple and Romano (1999) show that students with the lowest ability and lowest family income end up concentrated in public schools at the bottom of the hierarchy. Other students are distributed among a set of private schools that differ from each other by the ability and income of their students, with the private schools at the top attracting the most able and most affluent students. Empirical work based on U.S. data generally supports the model’s predictions (Epple, Figlio and Romano, 2000).

Second, when the characteristics of the school’s student body are an important determinant of the school’s quality, no simple programs or educational strategy can make a school with a large proportion of disadvantaged or low-performing students look effective. In many instances, the best strategy for such schools is to try to raise the quality of their student intakes, a strategy that cannot work in the aggregate.

Third, successful schools will be reluctant to expand if doing so requires lowering the average socioeconomic or ability level of their students. In New Zealand’s experience with full parental choice and self-governing schools, successful
schools in urban areas had no desire to expand their enrollment. To the contrary, they did everything they could to maintain the mix of students that made them attractive to parents and students in the first place (Fiske and Ladd, 2000).

Finally, schools with large concentrations of disadvantaged students have difficulty competing for students (Ladd and Fiske, 2001a). This observation does not, by itself, rule out vouchers as a policy tool. For policymakers concerned about equity, however, it raises some warning flags. As discussed further below, it casts serious doubt on the proposition that competition will improve the schools serving students who attend schools at the bottom of the distribution.

Effects of a Voucher System on Achievement and Productivity

Improving student achievement is typically the single most important goal of current education reform efforts. A large-scale voucher program could potentially affect student achievement through three interrelated mechanisms (Hsieh and Urquiola, 2002). First, it would shift students from the public sector to the private sector. Provided the private sector were more productive than the public sector in generating student achievement, this sectoral shift would increase the productivity of the education system. Second, such a program is likely to generate greater socioeconomic and racial polarization of students among schools as students seek to improve the quality of their peers. This greater polarization may increase overall achievement, decrease it or have no effect depending on how one’s peers affect the achievement of different groups of students. Third, the introduction of a voucher system would increase competition for students. Such competition, proponents argue, would increase achievement by forcing the public schools to become more effective. The following sections look at the issues and evidence related to these mechanisms.

The Sector Effect: A Shift to a More Productive Private Sector

The relative productivity of public and private sector schools is the most well studied component of the three mechanisms. The two major strands of the research, however, have relatively little to say directly about a large-scale voucher program.

One strand of the research was initiated by James Coleman and others in the early 1980s, using national data sets such as High School and Beyond or the National Educational Longitudinal Survey. In their seminal study, Coleman, Hoffer and Kilgore (1982) conclude that students in private high schools, most of which were Roman Catholic, outperformed their public school counterparts. However, that study did not fully account for differences in who enrolls in Catholic schools, and so many subsequent researchers have reexamined the issue paying close attention to the problem of self-selection.

To that end, researchers have had to grapple with the statistical challenge of finding an appropriate instrument, that is, a variable that is correlated with a family’s decision to choose a Catholic school, but is not a direct determinant of
educational outcomes. Many authors have used some function of whether a student is Catholic or the proportion of the county that is Catholic for this purpose. Such studies tend to show that Catholic schools appear to have at most small effects on student achievement as measured by test scores, but somewhat larger positive effects on the probability that students will graduate from high school and will attend college (Evans and Schwab, 1995; Neal, 1997; Grogger and Neal, 2000). In general, the benefits seem to be largest for urban minorities.

Other researchers have criticized this statistical approach on the grounds that Catholicism may be a direct determinant of educational outcomes, and they have proposed alternative measures. In their study of the relative productivity of all private schools, for example, Figlio and Stone (1999) use a set of instruments related to state labor laws, maintaining that such policies affect residents’ perceptions of the differences in performance between the public and private sectors and hence their inclination to use the private schools. Although Figlio and Stone report somewhat smaller positive overall effects on educational attainment than those that emerge from the studies of Roman Catholic schools, they do find higher test scores in math from attending religious schools for the subgroup of African-Americans in big cities.

A second strand of the research uses the recent voucher experiments to investigate the productivity of private elementary schools compared to that of public schools. Most of the knowledge about the effects of private schools at the elementary level emerges from evaluations of the publicly funded Milwaukee Parental Choice program and of the privately funded programs in Dayton, Ohio, Washington, D.C. and New York City. All of these voucher programs serve only a small fraction of the eligible students. Until the late 1990s, the Milwaukee program served no more than 1.5 percent of the district’s 90,000 students. During the time covered by the studies, the New York City program offered scholarships to only 1,300 students, the Dayton program to 515 public school students (and to another 250 already enrolled in private schools) and the Washington, D.C., program to 460 students (Peterson et al., 2000).

Results from the publicly funded Milwaukee program have been the most controversial, in large part because that program was not set up as a true randomized experiment and so researchers had to exercise judgment in choosing the appropriate control group to which the voucher students would be compared. Witte, Stern and Thorn (1995) compared the performance of voucher students to that of a random sample of other students in the Milwaukee public schools and concluded there were no significant achievement gains for voucher students. Greene, Peterson and Du (1998) compared voucher students to potential users of

---

1 The instruments are state “duty to bargain,” or “right to work,” laws and various interactions of these variables with median income in the country and the socioeconomic status of the student’s family. Subsets of the instrumental variables meet the criteria of being highly correlated with a student’s selection of public or private school and yet are likely to have no direct affect on a student’s achievement.

2 Evaluations of other programs in San Antonio, Indianapolis and Cleveland, with a less useful study design, can be found in Peterson and Hassel (1998).
vouchers who were not admitted to their preferred schools and concluded that by
the third and fourth years of the program, voucher students exhibited significant
gains in both math and reading (of the order of 0.1 to 0.5 standard deviations).
Looking at the same data a third time, Rouse (1998) was able to build on and
improve upon the research methodologies of the two previous studies; for example,
she took into account the fact that many students who were given vouchers did not
exercise their right to use them and also paid attention to nonrandom attrition of
students from the sample over time. Rouse’s sensitivity to the statistical problems
provides some confidence in her finding that the program generated small gains
for students in math, but none in reading.

Recent studies of privately funded voucher programs in Dayton, Ohio, Wash-
ington, D.C., and New York City provide additional information on how voucher
programs—and hence private schools—affect the achievement of elementary and
middle school students. In contrast to the Milwaukee program, each of these
programs was set up as an experiment with random assignment of children.
Participating families filled out baseline surveys of background information and, in
principle, all children in both the treatment and control groups were tested
Annually. In contrast to Milwaukee, the private funding of these programs made it
possible for students to use their vouchers in religious as well as secular schools.
The programs limited eligibility to families with low income (New York City) or low
and moderate income (Washington, D.C., and Dayton). Voucher amounts were

Based on three years of the voucher programs in New York and Washington,
D.C., and two years in Dayton, Howell and Peterson (2002) find no evidence of a
general achievement difference between the public and the private schools. In no
year and in no individual city (other than the second year in Washington, D.C.) was
there evidence that students who shifted to private schools achieved at higher
average levels than students who remained in the public school system. Further,
when the analysis was disaggregated by the race of the students, no differences
emerged for either white or Hispanic students.

Only for the subgroup of African-American students did positive differences in
achievement emerge. Even for this group, however, the differences were consistent
across neither cities nor grades. For example, African-American students in Wash-
ington, D.C., who shifted to private schools achieved at far higher levels in the
second year of the program, but their gains were negligible in years one and three.
Although the New York study generated a more consistent average pattern of
achievement over time, the positive differential emerged clearly and consistently
only for students in the fifth grade (Howell and Peterson, 2002, Table 6.2,
Table D.1).

\[ \text{In practice, there was a significant drop off over time in the students who returned for testing. In}
\text{Dayton and Washington, D.C., return rates ranged from 40 to 60 percent. In New York City, 82 percent}
\text{of the students returned for testing the first year and about 66 percent in years two and three (Howell}
\text{and Peterson, 2002, Table 2.4). The authors weighted their regressions to adjust for observed differ-
\text{ences in the characteristics of respondents and nonrespondents.} \]
Based on Howell and Peterson’s (2002) preferred estimates, which disproportionately weight the relatively stable New York results, African-Americans who switched to private schools scored about 3.9, 6.3 and 6.5 percentile points higher than did comparable students in the control group in the first three years of the program. These effects are based on the national distribution of percentile rankings on the Iowa Test of Basic Skills, and, for perspective, the gains are about two-thirds the size of the differences that emerged in another intervention that helped African-American students—the Tennessee experiment that reduced class sizes.

These effects for African-Americans represent estimates not of the offer of a voucher per se, but rather of the shift to a private school. This distinction is important given that only 53 percent of the students offered vouchers in New York City and less than 29 percent in Washington, D.C., were still in private schools three years into the program and that the students who used the voucher to attend private schools were not a random sample of those offered a voucher. In New York City, for example, having a mother on welfare reduced the probability that a student would accept a voucher by 8 percentage points and the probability of remaining in a private school for two years by another 7 points, and having a learning disability reduced the probabilities by 14 and 13 percentage points. Working in the other direction, being religiously observant significantly increased the probabilities (Howell and Peterson, 2002, Table 3.12). Thus, the positive differential effect for African-Americans attending private schools applies not to a random sample of low-income African-Americans, but rather to a particular subset of this group.

Furthermore, it is not at all clear that this positive effect of private schools can be extrapolated to an expanded voucher program, even one targeted at a similar group of African-American students. The issue turns on the explanation for the observed achievement gains. If the apparent success of private schools for African-American students reflected the autonomy and absence of bureaucracy in such schools, then, following the logic of Chubb and Moe (1990), an expanded private school sector could, in principle, generate comparable gains. However, the observation that the differential achievement gains emerged only for African-Americans argues against this explanation. Alternatively, if the success of the private schools reflected a better match between the needs of African-American students and the offerings of particular schools, then expanding the private sector would once again

---

4 The Tennessee class size experiment generated gains in comparable percentile rankings of 9–10 points for black students shifting to small classes within public schools (personal correspondence with Alan Krueger). This comparison of effects on percentile rankings is preferred to comparisons based on the more common measure of “effect sizes” defined in terms of standard deviations in that the latter comparisons require that the relevant standard deviations come from similar distributions of test scores.

5 The New York take-up rate is based on information provided in Howell and Peterson (2002, p. 235, footnote 19) and is well below the 70 percent figure they report in Table 2.3. The latter figure is the percentage of the voucher recipients who attended the third year follow-up testing session who used a voucher to attend a private school during the three years. Similarly, the 29 percent figure reported in their Table 2.3 for Washington, D.C., also exceeds the true take-up rate in the third year, but the true rate is not provided.
continue to generate comparable positive benefits, but only if the factors that parents were looking for could be replicated in newly established private schools. The problem is that the measured success of the African-American students who shifted to private schools likely reflects in part the more disciplined student bodies in the private schools, especially in the Catholic schools, or the more advantaged and motivated group of students whose parents were willing to pay the private school tuition. These school characteristics cannot be replicated as easily in new private schools. This explanation for the positive findings in the Howell and Peterson (2002) research cannot be ruled out, since the authors had no data on the mix of students in the receiving private schools.

More explicit evidence about the sector effect for a large-scale voucher program emerges from Chile, where the universal voucher program generated a large number of new for-profit secular private schools that operated alongside the more established and somewhat better resourced Catholic schools. Careful analysis of fourth-grade achievement data in Chile indicates that, compared to the traditional public schools, Catholic schools generated higher achievement in Spanish and math while the new secular schools produced marginally lower achievement in Santiago and even lower achievement outside the capital city (McEwan and Carnoy, 2000, p. 227). This observation is important to the U.S. debate, where advocates of vouchers tend to use evidence from some types of private schools, namely Catholic high schools, or a combination of existing Catholic and other schools in the voucher experiments, to generalize to an expanded private sector that would inevitably include many types of private schools, not all of which would be able to attract the same mix of students as in the existing private schools.

Even if private schools do not generate positive differences in student achievement, lower costs could still make them more productive than public schools. Unfortunately, it is difficult to compare the true economic costs of education at private and public schools. One complicating factor is that public schools typically serve a greater proportion of students who need costly services such as special education or vocational education. Another is that tuition payments, or even total expenditures by private schools, do not represent their true costs (Levin, 1998). Private schools, especially religious private schools, receive resources in many forms: special fees, church subsidies, teachers working at below-market wages and donations of money, time, land and buildings.

If private schools could operate more cheaply than public schools, one would expect for-profit education firms like Edison or Tesseract to make a profit from taking advantage of those cost efficiencies. The fact that such firms have not been making profits suggests that cost efficiencies from private production are illusory. Further suggestive evidence comes from the voucher experiments. For example, the private schools in New York City serving the African-American voucher students offered significantly smaller class sizes than did the public schools (Howell and Peterson, 2002, p. 100). Given the large share of education costs attributable to teachers and classrooms, this observation suggests that the true resource costs of educating students in private schools could well exceed those of the public schools serving comparable students.
Thus, one should expect neither higher overall achievement nor lower resource costs as a result of a shift of students from public to private schools. At most, there are likely to be small achievement gains for a selected group of African-American students. Furthermore, a universal voucher program could possibly require the government to spend more public funds on education, because some of the voucher funds would undoubtedly go to families who would otherwise have paid all of the cost of putting their children in private schools.

The Peer Effect

Because many parents use the social and ethnic composition of a school’s students to judge a school’s quality, a large-scale voucher program—or, more generally, any unrestricted educational choice program—is likely to increase the racial and socioeconomic stratification of schools. Other aspects of voucher programs would also contribute to stratification to the extent that they placed low-income families in a less favorable position to exercise choice than higher income families. For example, low-income families would be disproportionately affected if the government did not pay for transportation to the chosen schools, if voucher schools were allowed to charge fees and tuition in addition to the amount of the voucher, if schools were allowed to select their students or if low-income families have less access to information than did high-income families. While a voucher program could be designed to mitigate these effects, the more basic pressure for stratification would remain as long as the voucher program were not restricted to low-income households.

The question here is how increased stratification is likely to affect the overall productivity of the education system. A variety of studies in which the authors have carefully addressed various thorny statistical problems have found evidence of pure peer effects in the sense of spillovers within the classroom from one student to another (Hoxby, 2000b; Hanushek et al., 2001). Positive peer effects also emerge from other studies, but not always consistently, as in Henderson, Miezkowski and Sauvageau (1978), Willms (1986), Zimmer and Toma (2000) and Bryk and Driscoll (1988).

If peer effects were positive and linear, the gains in achievement for the students who move out of the public schools in search of higher-quality peers would be exactly offset by the losses to other students, either those in the schools left behind or those in the destination schools. However, if the magnitude of the peer effects were greater for students with low socioeconomic status, then the movement of such students into schools with more affluent peers could potentially increase overall achievement. Alternatively, however, if the students who left the public schools were the more able and more motivated students, their gains in achievement could be more than offset by the loss in achievement of the students in the schools left behind, thereby reducing overall achievement.

Such asymmetry in peer effects is quite plausible. Students whose internal motivation to learn is reinforced by an educationally rich home environment, as is true for many students with high socioeconomic status, are likely to do relatively well in most academic settings. In contrast, the performance of students from more
educationally impoverished backgrounds could depend more heavily on the school. The presence of unmotivated fellow students and other associated features such as low expectations in the classroom, poor teachers and limited resources may well take its toll in significantly lower learning for such students relative to how they would perform in a school with stronger peers and higher levels of expectations, teacher quality and resources.

However, evidence on whether peer effects are asymmetric in this way is limited and inconclusive. Focusing on the spillovers from having high-ability peers, Hanushek et al. (2001) find that peer effects are positive throughout the range of student test scores, and they find no evidence of nonlinear effects. Using gender and race to define peer groups, Hoxby (2001) finds at most only limited evidence of nonlinearities. Evidence related to a more general form of peer effects that include neighborhood effects emerges from recent analysis of an experimental program of the Department of Housing and Urban Development in which families were moved out of high-poverty areas into more economically mixed neighborhoods. That analysis provides evidence that children in elementary schools who had access to schools with higher average test scores and more affluent neighborhoods achieved at higher levels than children in the control group. It provides no support, however, for the presence of a nonlinear relationship (Ludwig, Ladd and Duncan, 2001).

On the other side, consistent with an asymmetric effect are findings that school-level measures of socioeconomic status have stronger effects on the performance of black than on white students (Coleman et al., 1966) and that racial isolation had negative impacts on student performance in North Carolina schools (Mickelson, 2001). This lack of clarity about how peer effects differ among groups rules out any clear predictions about whether a voucher program would be likely to increase or decrease the overall productivity of the education system through the mechanism of peer effects.

**The Competition Effect on School Productivity**

Even with a large, unrestricted voucher program, the majority of students are likely to remain in traditional public schools. In Chile’s universal voucher program, for example, the percentage of students in private schools was still below 50 percent more than ten years into the program. Hence, crucial to the argument that a universal voucher program would increase overall student achievement is that voucher-induced competition from private schools would pressure traditional public schools to become more productive and force the weaker schools to close.

There are reasons to question these predictions. One possibility emphasized by McMillan (1999) is that the greater availability of private schools may reduce parental involvement in the public schools and thereby reduce one important positive contributor to student achievement in those schools. Moreover, in contrast to the private schools, which typically have significant leeway to select their students and, therefore, to offer focused, coherent programs designed to meet the needs of those students, competition may force the public schools to offer a diverse and unfocused education program as they struggle to be attractive to all comers (Fiske
and Ladd, 2000). Finally, the notion that the unproductive public schools will go out of business and that new and more effective public schools will replace them is far easier imagined than done.

The strongest claims that voucher programs have succeeded in making public schools more productive are based on two empirical studies: Jay Greene’s study of the Florida voucher program and Caroline Hoxby’s analysis of fourth-grade achievement in the Milwaukee voucher experiment. However, in both cases there are good reasons to question whether the point has been proven.

In a highly publicized study for the Manhattan Institute, Greene (2001) studied the effects of Florida’s voucher program on achievement in the public schools. Under the Florida system, schools are given grades of A, B, C, D or F. If a school receives two Fs within four years, its students are eligible for vouchers to attend private schools. Greene reports that after the first year of the program, the schools that had one F, and hence were subject to the threat of a voucher, raised their achievement significantly more than comparable schools not subject to such a threat and that their greater gains remained even after he adjusted for the statistical problem of regression to the mean.

However, with Florida data alone, Greene is unable rule out an alternative and, in my view, more convincing explanation for his findings, namely, that the improvement in the state’s low-performing schools was a response to the state’s grading of schools, rather than to the small voucher component of that program. Support for this alternative interpretation emerges from other states such as North Carolina, which, like Florida, rank schools, but, unlike Florida, do not have voucher programs. The patterns of gains in student performance in the low-ranked schools in North Carolina after the first year of that state’s program, for example, were almost identical to those found in Florida (Ladd and Glennie, 2001). The comparison between Florida and other states such as North Carolina strongly suggests that the increased scrutiny, shame and additional assistance associated with being labeled a low-performing or “failing” school is a more likely explanation for the improvement of the bottom schools in Florida than the threat of a voucher.

Hoxby (2001) carried out an analysis of fourth-grade achievement in Milwaukee public schools before and after the expansion of that city’s voucher program in 1998. She reports that the annual increase in student achievement was higher in the Milwaukee public schools most subject to competition from private schools (those with 75 percent or more of their students eligible for vouchers) than in other Milwaukee schools and even higher than in a control group of twelve other Wisconsin elementary schools. In math, for example, the annual increase in test scores of 7.1 percentile points in the schools subject to competition exceeded the increases of 5.3 and 3.7 percentile points in the other two groups, and similar patterns emerged for science and language.

However, Hoxby’s (2001) interpretation overstates the potential gains from a voucher program because she was unable to control for the changing mix of students in her treatment and control groups. The Milwaukee program was limited to low-income families. Detailed analysis of an earlier version of the Milwaukee’s means-tested voucher program indicated that the average test scores of the voucher
applicants were well below those of other students in the Milwaukee system (Witte, 1999, chapter 4). If provided that pattern continued with the 1998 expansion of the Milwaukee program, we would expect to see a movement of relatively low performing students out of the treatment group of schools and a potentially large corresponding increase in the average achievement of students remaining within those schools even with no change in the productivity of the public schools.

In addition, it appears that the mix of students in Hoxby’s (2001) twelve control schools from the rest of the state may have been changing as well. My own crude analysis based on data from the National Center for Education Statistics at the district level shows that the percentages of low-income students increased on average during the 1997–2000 period in the four districts in which most of the 12 control schools were apparently located at the same time that the percentage declined by 4.7 percentage points in Milwaukee. Those increases in the proportions of challenging-to-educate children thus provide an alternative explanation for the sluggish growth in average achievement in Hoxby’s control schools relative to the schools subject to competition in Milwaukee.

Thus, it would be premature to conclude, based on the Greene (2001) and Hoxby (2001) studies, that voucher programs have unleashed strong positive impacts on the public schools. In the absence of longitudinal data for individual students that would allow the researcher to isolate impacts on achievement of the same group of students over time, any conclusions about competitive impacts are highly suspect.

Other potentially relevant sources of evidence on this question are studies of how competition from private schools and charters have affected the traditional public schools in the United States and studies of how the voucher-induced private schools have affected public schools in Chile.

The U.S. evidence from private schools provides evidence of at most small positive impacts of private schools’ competition on academic achievement in the public schools. In a comprehensive review, Belfield and Levin (2001) report that well over half of the 94 estimates in 14 studies were statistically insignificant and that any positive impacts were either substantively small or subject to question based on subsequent studies. A handful of estimates, including those by McMillan (1999),

---

6 In 1994, for example, the test scores for voucher applicants were 14 percentile points below those for all Milwaukee public school students in reading and 19 points below in math. Compared to a random sample of low-income students in Milwaukee public schools, the test scores for voucher applicants were 5 points lower in reading and 10.5 points lower in math (Witte, 2000, Table 4.6).

7 I do not know the precise set of schools that Hoxby (2001) included in her analysis. However, by identifying all Wisconsin primary schools in which more than 25 percent of the students were black, I was able to narrow the relevant set of districts. I relied on district-level averages of students receiving free and reduced price lunches, since that information was not available at the school level. Between 1997 and 2000, the percentage of students on free and reduced price lunch rose in Racine by 0.3 percentage points, rose in Kenosha by 1.5 percentage points, rose in Beloit by 3 percentage points and declined by 0.8 percentage points in Madison at the same time that it fell by 4.7 percentage points in Milwaukee. These district-wide averages could well translate into even larger increases in the schools within each district that Hoxby included in her sample because of their relatively high proportions of minority and low-income students.
who incorporated parental involvement, suggest that competition from private schools may have a negative impact on public schools.

With respect to the effects of competition from charter schools, evidence is only now beginning to emerge in states such as California, Arizona and Michigan, where charter schools are now common. Anecdotal and interview data suggest that some school districts or schools have been responding to the establishment of charter schools in positive ways (Rofes, 1998; Gresham et al., 2000; Hess, Maranto and Milliman, 2000). For example, some school districts have set up after-school or all-day kindergarten programs, established new magnet schools, changed curriculum, empowered teachers or changed principals. In addition, some principals appear to have promoted experimentation in teaching or pursued other forms of behavior that could be viewed as positive. However, many school districts have not responded at all to the new charter schools.

Only a few studies examine impacts of charter schools on outcome measures such as achievement. Hoxby (2001) finds small positive impacts, but her estimates in this study are subject to the same sorting bias that emerges in her study of vouchers in Milwaukee. I find more methodologically persuasive a study of Michigan schools by Bettinger (1999) because of the attention it pays to various statistical problems—including the possibility that the location of charter schools may be influenced by the performance of the public schools. To deal with this simultaneity problem, the author took advantage of the role of public universities in founding charter schools to develop an exogenous instrumental variable. The study relies on data only through 1996, but it finds no impact of charter schools on public school performance.

Potentially more reliable information emerges from Chile, which has 20 years of experience with a large-scale voucher program. If competition increases achievement, one would expect the public schools in Chilean municipalities with large increases in private enrollment shares to exhibit greater gains in achievement than those subject to less competition from the private schools. McEwan’s (2000a) study of Chilean schools uses panel data and a “differences-in-differences” methodology to sort out the effects. This statistical strategy involves comparing changes in test scores in one period to changes in test scores in the previous period. The advantage of this approach is that it controls both for unobserved determinants of achievement that are constant over time for individual schools and for unobserved time trends in each school’s achievement. McEwan’s conclusions are mixed. His preferred estimates suggest that 15 years of competition led to modest gains in achievement of about 0.16 to 0.2 standard deviations among some public schools in Santiago, Chile’s capital, but small negative effects in the rest of the country, which is home to three-quarters of the country’s population. He concludes (p. 152) that the results “neither refute nor provide strong support for the view that competition will lead to improvements in the quality of public schools.”

Net Impact of the Three Mechanisms on Student Achievement

The evidence suggests that the overall effect of vouchers on student achievement is likely to be small at best. Studies from the United States and other countries
provide no compelling evidence in support of the view that the private sector is generally more productive than the public sector, except possibly for a subset of African-American students, nor that there are significant gains to be had from competition. Nor is there clear evidence of the asymmetric peer effects that could affect overall productivity. Empirical support for at most a small overall effect emerges from Chile, where Hseigh and Urquiola (2001) estimated the net effect of the three mechanisms based on that country’s 20-year experience with vouchers, but found only small and statistically insignificant effects of the voucher program on student achievement.

The observation that the productivity argument for a universal voucher program is weak does not rule out other types of benefits, such as those that would accrue to families who used school vouchers to achieve a better match between their values, including their religious values, and the values of the schools their children attend. It does imply, however, that the debate about voucher programs should revolve around the desirability of benefits of that type rather than around their alleged contribution to student achievement.

**Impacts of a Voucher Program on Disadvantaged Students**

Many supporters of vouchers believe they will improve the welfare of educationally disadvantaged students. The question of whether a voucher program will help or harm such students turns in part on whether the program is a large-scale universal program or a smaller program targeted specifically toward disadvantaged students.

**A Large-Scale Universal Voucher Program**

Several theoretical models of vouchers predict that the students at the bottom of the distribution in the public school sector will end up worse off under a universal voucher program (Epple and Romano, 1998; Nechyba, 1999, 2001). The main reason the students at the bottom are harmed is peer effects. With the introduction of vouchers, the more able or motivated students leave public schools for private schools. Their departure reduces the quality of public schools and renders students remaining in those schools worse off than they would have been without the voucher program. Note that this outcome requires only that peer effects be positive, not that they have differential effects on different groups of students. Moreover, in the Epple and Romano (1998) model, even some students who switch to private schools may end up worse off. That outcome can occur when private schools are allowed to charge more than the voucher. In that case, some students will switch to private schools primarily to avoid the decline in the quality of the public schools. Even if their achievement is higher in the private school, they could be worse off because they now must pay tuition.

Some might discount the predictions of these models because the authors explicitly assume that a voucher system would have no positive impact on the productivity of the education system. However, the review of the empirical evidence in the previous section suggests that assumption is reasonable.
The data on educational outcomes from Chile’s universal voucher program generally supports the predictions of the theoretical models. Chile’s voucher program did induce the higher-income and -ability students out of the public sector schools. Cross-sectional quantile regressions for that country, which allow one to look at the performance of different parts of the student distribution, show that voucher-induced expansions of private schools widened the variation in educational outcomes across students (Hsieh and Urquiola, 2002). New Zealand’s experience with universal public school choice was similar, although conclusions about achievement are harder to confirm because of the limited achievement data in that country. However, there is little doubt that the expansion of choice in that country exacerbated the problems of the schools at the bottom of the distribution and reduced the ability of those schools to provide an adequate education (Fiske and Ladd, 2000).

Of course, some disadvantaged students in both countries were made better-off because vouchers in Chile and public school choice in New Zealand gave those students access to some schools that previously were outside their financial reach. This observation is important and complicates any discussion of voucher programs. From an ethical perspective, it is hard to justify denying schooling options to such children because their families are poor or because their departure may reduce the quality of education of those who remain behind. Providing additional choices to such families is a desirable goal.

The adverse effects of large-scale voucher programs on the students left behind highlights the need to shape voucher policies in ways that could minimize those effects. One starting point might be to adjust the voucher amount both to the characteristics of the students (with costly-to-educate students getting significantly larger vouchers) and possibly to the characteristics of the schools (for example, students attending economically integrated schools would receive larger vouchers that those in homogeneous schools). Such schemes, however, are likely to be politically contentious and difficult to implement. A second approach is to place restrictions on the use of vouchers, such as prohibiting the participating private schools from charging additional tuition or fees above the voucher amount (a prohibition that is part of Chile’s program) and requiring that oversubscribed private schools select students randomly (as is the case in Milwaukee and Cleveland). With such constraints, public funds would support education only in schools that were available to all students, and all students would continue to be guaranteed an education at no out-of-pocket expense, whether it be in a public or private school. These constraints would make voucher-financed private schools similar to charter schools. Indeed, one prominent supporter of charter schools argues that these two characteristics of charter schools—equal access and no fees—are what make charter schools preferable to voucher-financed private schools (Hassel, 1998).

However, even if a universal voucher program were modified in these ways, the fundamental problem facing the U.S. education system with respect to disadvantaged students would remain. That problem is the significant concentrations of difficult-to-educate students in some schools. The challenge is to find a way to
expand the educational choices available to families, while at the same time reducing those concentrations. That challenge can best be met with some form of controlled or managed choice among public schools, as forcefully advocated in a recent Task Force Report from the Century Foundation (2002). Under such a system, families would specify their preferred schools. Students would then be assigned to schools based on those preferences, but with attention to the mix of students in each school, either the racial mix (as was long the case in the managed choice program in Cambridge, Massachusetts) or the economic mix as advocated by the Task Force (and recently introduced in Cambridge). Any such system of controlled choice would require that the public authorities ultimately have the power to assign students to schools. In addition, it would require targeted investments focused on teaching and learning in the schools that were not successful in attracting students so as to promote healthy competition among all schools. Since private schools are likely to be reluctant to participate in a program of that type, vouchers would not be a logical component of such a strategy.

A Means-Tested Voucher Program Limited to Low-Income Families

If the goal were to use a voucher program specifically to assist low-income families, at a minimum, the program would have to be means tested. Evidence from Milwaukee suggests that a means-tested voucher program can be successfully designed to serve low-income families, especially those whose children are unsuccessful in the public schools (Witte, 1999, p. 196). The main advantage of such programs is that parents tend to be more satisfied with their new schools than with their assigned public schools. In the fifth year of the Milwaukee voucher program, for example, more than three-quarters of choice parents gave their child’s school a grade of A or B. Similarly, in the three urban privately funded voucher programs, 40 percent of the private school parents gave their schools an A compared with 14 percent of the control group (Howell and Peterson, 2002, p. 174). Moreover, satisfaction levels among private schools parents were higher with respect to all the major components of the school: the academic program, school safety, parental involvement and class size (Howell and Peterson, 2002, Table 7.1, p. 173).

Of course, some of this increased satisfaction may reflect not the specific policies of the schools but rather a different, more congenial or more motivated set of peers in the new schools. To the extent that families opt for schools in which their children will have peers with higher socioeconomic status, their behavior complicates the policy discussion, because not all families can achieve that end. Nonetheless, it is hard to argue that low-income families should be denied opportunities to benefit from such choices simply because they are poor.

Even for means-tested programs, however, design matters. For example, the Milwaukee program required that schools accept the voucher in lieu of all tuition and fees and that oversubscribed schools select students randomly—a design that helps reduce the negative impact of vouchers on disadvantaged students. In contrast, consider the design of the Children’s Scholarship Fund (CSF), a privately funded national school voucher program that provides scholarships for children from low- to moderate-income families (up to 270 percent of the poverty level) to
attend private schools, with the scholarships scaled to the income of the family. Although the program reached large numbers of African-American low-income and minority households, only 30 percent of all families offered vouchers ended up using them. Forty-five percent of the decliners said that they could not afford their preferred school, 10 percent said no space was available, and 8 percent cited transportation problems (Campbell, West and Peterson, 2001). Thus, when tuitions are not limited to the amount of the voucher and transportation is not provided, many families will not be able to benefit from the voucher program.

Additional concerns arise from the high rates of attrition from voucher programs over time. In the early years of the Milwaukee program, dropout rates were very high: 54 percent of the voucher recipients did not return to their private school after the first year of the program (Witte, 1999). The high attrition rate in that program might be explained in part by the poor quality of some of the participating private schools, given that religious schools were not permitted to participate and the fact that one major private school closed. However, high dropout rates also emerged in the privately funded New York City voucher program. By the end of the third year of that program, 38 percent of the voucher users had left their voucher-subsidized private school (Howell, 2002, footnote 21). Multivariate analyses of those who were offered vouchers, used them and then dropped out indicate that while African-Americans were more likely than other racial groups to accept an offer of a voucher, they were also far more likely to drop out each year (Howell, 2002, p. 18). This finding raises additional questions about the ability of even a means-tested voucher program to meet the needs of low-income African-American students.

Given the serious educational challenges facing disadvantaged students, particularly those living in areas of concentrated poverty, it is hard to argue against any policy, including means-tested vouchers, that might improve the educational experiences of some students. In some situations, means-tested vouchers could play a useful role in providing additional options to students whose schooling choices would otherwise be severely constrained. Even such voucher programs, however, would need to be embedded in a larger strategy of education reform that focused on teaching and learning in the public schools and one that provided greater choice within the public school system.

Conclusion

Contrary to the claims of many voucher advocates, widespread use of school vouchers is not likely to generate substantial gains in the productivity of the U.S. K–12 education system. Any gains in overall student achievement are likely to be small at best. Moreover, given the tendency of parents to judge schools in part by the characteristics of the students in the school, a universal voucher system would undoubtedly harm large numbers of disadvantaged students. Although small means-tested voucher programs might provide a helpful safety valve for some children, policymakers should be under no illusion that such programs will address
the fundamental challenge of providing an adequate education to the large numbers of disadvantaged students in many of our large cities. At the same time, there are good arguments for giving families, especially those who are economically disadvantaged, more power to choose the schools their children attend. The challenge for policymakers is to find ways to expand parental choices without excessively privileging the interests of individual families over the social interests that justify the public funding of K–12 education.

The author appreciates the helpful comments of David Figlio, Patrick McEwan and the editors of this journal, particularly Timothy Taylor.

References


This article has been cited by:


7. Gerhard Glomm, B. Ravikumar, Ioana C. SchiopuThe Political Economy of Education Funding 4, 615-680. [CrossRef]

8. Alejandra Mizala, Florencia Torche. 2010. Bringing the schools back in: the stratification of educational achievement in the Chilean voucher system#. International Journal of Educational Development . [CrossRef]


22. Simon Burgess, Carol Propper, Deborah Wilson. 2007. THE IMPACT OF SCHOOL CHOICE IN ENGLAND. *Policy Studies* 28:2, 129-143. [CrossRef]

23. WeiLi Ding, Ming Lu. 2007. Are equity and efficiency irreconcilable goals in education?. *Frontiers of Education in China* 2:2, 163-180. [CrossRef]


32. Miguel Urquiola . 2005. Does School Choice Lead to Sorting? Evidence from Tiebout Variation. *American Economic Review* 95:4, 1310-1326. [Citation] [View PDF article] [PDF with links]


