CHAPTER 8

Education and Education Reform

Education is widely viewed as an effective path to social mobility. By getting a good education, disadvantaged children can make up for their disadvantages. Yet, one of the ways in which high-poverty areas typically differ from other parts of the metropolitan area is that they have lower quality schools. Children in inner-city poverty neighborhoods often attend schools that are markedly inferior to those attended by children in well-to-do suburban communities.

It is widely believed that many of the inner-city schools are broken. Critics claim that students are rewarded for "seat time," that children who simply stay in school are rewarded with meaningless diplomas even if they cannot read or add and subtract.

If one of the reasons that children from high-poverty neighborhoods do poorly in life is that they attend low quality schools, then it is natural to ask whether there are ways to reform schools so that poor children do better.

There is general agreement that while money can be helpful, increasing spending on schools is not sufficient to ensure reform.[1] Many struggling urban school systems spend more per pupil than successful suburban school systems. The suburban school systems are able to create environments in which it is less costly to educate children, perhaps because much of the social education of the children has been done at home.

As a consequence there have been a number of movements to reform the public schools in the United States. Many of these concern curriculum or school management issues that are designed to substitute for the advantages of suburban school system. Some experts believe that highly structured school reform programs such as "Success for All" can be used at the level of the individual school to achieve reform.

Discussion of curricular, administrative or structural reform with schools is beyond the scope of this chapter. Instead, we will focus on

[1]Hanushek, Eric A., *Making Schools Work: Improving Performance and Controlling Costs*, Washington, DC: Brookings Institution Press, 1994.

reforms that are being considered or implemented at the district, state or federal level. In particular, we look at

- Mandatory testing
- Vouchers
- School choice
- Charter schools

However, before we do so, we begin with a discussion of the relation between earnings and education. Is it really true that getting an education is the path to upward mobility?

1. Education and Earnings

People with college educations, on average, earn more than those with only high school diplomas who, in turn, earn more than high school dropouts. By this point in the book, I hope you are saying, "But that doesn't mean that education raises earnings. Maybe the type of person who gets a lot of education also tends to be the type of person who would earn a lot anyway."

Economists have grappled with this question. It is not easy to determine whether education raises earnings, but there have been a large number of attempts to address this question, and almost all find that at least part of the relation between education and earnings is causal, that is that getting more education raises earnings.

Rather than attempt a full review of this vast literature, I will take it for granted that the relation is a least partially causal and discuss why education might raise wages. There are three principal explanations: human capital, sorting and credentialism.

The *human capital* story is simple. School teaches students valuable skills: mathematics, reading, how to get along with other people. Employers observe these skills and are willing to pay more to workers who have them.[2] In its extreme form, the *sorting* model says that education does not provide skills, but more able workers get more education to signal that they are more able. Employers cannot determine directly whether a worker has the skills it wants, but knowing that more able workers get more education, they use education as a screen, hiring more educated workers for jobs requiring more ability.

For example, firms are willing to pay a premium to hire smarter workers. If smarter workers get more education, they will pay a premium to more educated workers if they cannot observe "smarts" directly. All workers will recognize that more educated workers get paid

[2]The classic book in this area is Becker, Gary S., *Human Capital*, 2nd ed., Chicago: Chicago University Press, 1975.

more, but those who find education more costly will get less education than those who find it relatively low cost. If smarter people find education less painful (and thus less costly), they will get more education. In equilibrium, more able workers get more education because it is rewarding to them and less costly. Firms pay more for more educated workers because more educated workers are more able.[3]

Clearly both mechanisms are at work. Firms can use interviews, recommendations and other information to learn about worker productivity without relying solely on education, but it is likely that they also infer information from the worker's educational attainment. Firms care about some, but not all, of the skills learned in school. Economists have developed hybrid models that account for both roles of education.[4]

Education also provides credentials. *Credentialism* is superficially similar to sorting in that it maintains that employers use education to sort workers. However, the mechanism is somewhat different. Credentialism maintains that the actual productivity differences associated with education are often small. However, employers are often forced to pay a wage that is independent of the worker's quality either because the wage is set by bargaining, government regulation (especially in government jobs) or is advertised.

In such cases, the firm will hire the best applicant. But figuring out who the best applicant is can be difficult. So the firm uses a shorthand, education, to narrow the pool of candidates from which it chooses.[5]

2. Testing

One of the biggest changes that has occurred in the United States has been a movement to create standardized tests at the state level. Many states use the tests as a partial basis for grade promotion or graduation. Students who do not do sufficiently well on the test are not eligible for promotion or graduation. Tests that have consequences for students, schools or school districts are referred to as high-stakes tests. This section examines the effects of high-stakes testing for students.

[3]Spence, Michael. "Job Market Signaling," *Quarterly Journal of Economics*, 87 (1973): 35-374.

[4]Lang. Kevin, "Does the Human Capital/Sorting Debate Matter for Development Policy?" *American Economics Review,* 84 (March 1994). Weiss, Andrew, "Human Capital vs. Signalling Explanations of Wages," *Journal of Economic Perspectives*, 9 (Fall 1995): 133-54.

[5]Credentialism is usually presented as irrational as in Berg, Ivar E., *Education and Jobs; The Great Training Robbery*, New York: Praeger Publishers, 1970. The discussion here develops it in the context where it is profit maximizing for the firm.

Later in this chapter we will examine the effects of making the tests high-takes for schools or school districts.

Supporters of high-stakes testing for students usually make two arguments:

- Standards for students are very low in many schools. Students graduate on the basis of "seat time" rather than accomplishment. A student who shows up to school and behaves will graduate regardless of whether he can read or add and subtract. By requiring a minimum standard of achievement, testing ensures that students must not only put in seat time, but also that they put in effort.
- Potential employers have little information about the performance of high school graduates. They cannot rely on the fact that a job applicant has graduated from high school as proof that she has acquired basic skills. By making students pass a test in order to graduate, we ensure that potential employers know that high school graduates have at least mastered some set of skills measured by the test.

It is important to recognize that, while somewhat different, both of these arguments are based on the idea that it is important to raise standards, at least at some schools. The theory below focuses on the idea that testing raises standards.

2.1. Theory. High-stakes testing only makes sense if employers cannot observe the productivity of workers directly. If they can quickly figure out how good a worker is, then the results of the test provide them with no additional information, and there is no reason that they should pay any attention to the test. Therefore, high-stakes testing only makes sense if we are in a world described by either the sorting model or a hybrid model in which employers use educational attainment to infer information about worker productivity.

We will begin by distinguishing two extreme cases and then discuss what happens if we combine them. In the first case, workers do not really care about the skills tested on the exam. They want employees who can answer algebra questions because these students will generally be good at arithmetic and other reasoning skills but do not expect their employees to use algebra in their jobs. We call this the pure sorting model. In the second case, employers care about the content of the test. Workers who have acquired the skills on the test are more productive than those who have not acquired the skills regardless of any general reasoning or mathematical skills. We call this the human capital case

although since employers use the test to learn about productivity, it also contains an element of sorting.[6]

2.2. Pure Sorting. In our first extreme case, education does not increase human capital. Instead, employers use information from the test to screen workers for characteristics that will make them more productive but not to ensure that the workers know the actual tested skills. For example, in Iran at least part of the national exam consists of "fill in the blank questions" where the only correct answer is the word used in the textbook. The ability to memorize the textbook sufficiently to use the exact word found in the text may be indicative of other cognitive and behavioral traits that are useful to employers, but it is unlikely that memorizing the textbook is directly useful on the job.

Thus the actual skills are of no productive use but the types of students who can graduate high school with less effort are also the types who will do well in the job market. In reality the relation is not perfect. Many of us know high school dropouts who went on to be very successful in the job market. Again, however, just to keep the story simple, we will assume that the relation is perfect.

Firms will pay workers who pass the test more than they will pay workers who do not pass the test because they expect workers who pass the test to be better workers. If, on average, high school dropouts who apply for jobs are worth \$7.00/hour, then, in the absence of other information, employers will be willing to pay them that amount. On average, high school graduate applicants might be worth \$10.00/ hour and would be paid that even if they were somewhat better or worse than the average applicant. Of course, over time, employers will learn about which high school graduates are particularly good workers and which are not and wages will adjust, but just to keep the story simple, let us assume that this learning takes long enough that most starting workers care only about what they will earn initially.

The average productivity of students who pass the test and the average productivity of workers who do not pass the test will each depend on the ability cutoff at which workers pass the test. The higher the ability cutoff at which students pass the test, the higher the average ability of **both** students who pass the test **and** those who do not. In case this is not clear, consider an example with six workers who produce 1, 2, 3, 4, 5 and 6, respectively. If only the person who produces 6 passes

[6]These models draw on Betts, Julian R. and Robert M. Costrell, "Incentives and Equity under Standards-Based Reform" in Diane Ravitch, ed., *Brookings Papers on Education Policy: 2001*, Washington, DC: Brookings Institution, 2001.

the test, the average of those who pass the test is 6 and the average of those who do not is 3. If the three most productive people pass the test, their average productivity is 5, and the average productivity of the remainder is 2. If everyone except the person producing 1 passes the test, the average productivity of those passing the test is 4 and the average productivity of the one person who does not is 1.

In the example above, the wage differential between those who pass the test and those who do not is always 3. That is the result of the specific example. In general, increasing the proportion of people who graduate high school (pass the test) can either increase or decrease the differential, but to keep things simple, we will ignore this effect in the discussion below.

What happens if we raise the standard for high school graduation by adding a high-stakes test? Two groups have no reason to change their behavior. Those who would have dropped out anyway will find it even more difficult to pass the test and will have no reason to work harder in order to graduate. Those who already meet the new standard also have no reason to work harder. However, those who met the old standard but not the new standard must decide whether to increase their effort and meet the new standard or dropout. Those close to meeting the new standard will increase their effort. Those who had difficulty meeting the old standard will drop out.

To summarize, there are four groups who respond differently to the increased standard. Low-skill students who would have dropped out anyway do not change their behavior. Low-moderate-skill students respond by putting in less effort and dropping out. High-moderate-skill students respond by increasing their effort in order to graduate, and high-skill students who would have met the standard anyway do not change their behavior.

Perhaps surprisingly, the groups that are most clearly helped by the increased standards are the two that did not change their behavior. Since they would have dropped out anyway, low-skill students do not increase their effort. However, because the set of dropouts now includes the low-moderate-skill students, employers are more willing to take a chance on a dropout, and their wages rise. Wages for high school graduates also rise since they are now a more select group. For those who do not have to work harder to meet the new standard, this is clearly a benefit.

The effect on the other two groups is more complex. For some of the individuals who work harder to meet the new standard, the higher wage for high school graduates offsets the added cost of meeting the higher standards. For others it will not. Perhaps surprisingly, some

TESTING IN A PURE SORTING MODEL: EXAMPLE								
	Original Standard				Higher Standard			
Productivity	Required Effort	Actual Effort	Wage	Net Wage	Required Effort	Actual Effort	Wage	Net Wage
1	3.2	0.0	1	1.0	6.4	0.0	2	2.0
2	2.4	2.4	4	1.6	4.8	0.0	2	2.0
3	1.8	1.8	4	2.2	3.6	0.0	2	2.0
4	1.2	1.2	4	2.8	2.4	2.4	5	2.6
5	0.6	0.6	4	3.4	1.2	1.2	5	3.8
6	0.0	0.0	4	4.0	0.0	0.0	5	5.0

Figure 1:

of the new dropouts are helped and some are hurt. Those who were almost indifferent between meeting the old standard and dropping out are better off because the opportunities available for dropouts have improved. Those who are almost indifferent between meeting the new standard and dropping out are probably worse off.

Table 1 continues our earlier example. The left panel shows the equilibrium with the "original" standard. I have chosen the effort required of each type so that all except the lowest type choose to meet the standard. As discussed above, the wage for those meeting the standard in this case is 4 and the wage for those not meeting it is 1. Because it costs type 1's 3.2 to meet the standard, they choose not to. For all other types, the cost is less than 3, and they choose to meet the standard. The fourth column in the table shows their wage net of the cost of meeting the original standard.

The second panel in table 1 shows the equilibrium with a higher standard. Now, workers with productivity 1, 2 or 3 choose not to meet the standard while those with productivity 4, 5, and 6 choose to meet it. Wages for those meeting the standard and those not meeting the standard are higher under the higher standard.

However, as discussed more generally above, there are winners and losers when we take into account changes in the decision to meet the standard and changes in effort. Comparing the last columns in the two panels, we see that the person who would not meet either standard is better off because he does not change his effort level but benefits from being grouped with better workers. The same is true for the person who

meets the new standard effortlessly. Among those who choose not to meet the higher standard but would choose to meet the lower standard, the less productive worker is better off and the more productive worker worse off. The opposite is true among the workers who choose to raise their effort to meet the new standard.

2.3. Human Capital. In contrast with the previous subsection, we now assume that firms care about the content of what is taught in school. We call this the human capital model because as in the human capital model, the education makes workers more productive. It is, in fact, a hybrid model because a test is required to determine whether or not the worker has acquired the human capital.

The first thing to note about this case is that (ignoring for the moment students going on to college), students will either put in the effort they need to pass the test or not put in any effort at all. There is no benefit to learning some of the material covered in the test even if employers value it because they only know if the student has passed the test.[7]

Each has to decide whether it is worth making the effort to graduate high school in order to earn the additional money. Some will find graduating high school very easy and will put in the small amount of effort they require. Others will find high school so difficult that they could not pass with any reasonable level of effort and will drop out. Most will weigh the costs and benefits. Some will decide that they would have to put in too much effort and will dropout while others will decide that the effort is worthwhile.

Employers know that students who pass the test put in exactly the effort needed to pass and are worth whatever firms are willing to pay for workers with that level of knowledge. Workers who have not passed the test have put in no effort and will be paid as workers lacking those skills.

Suppose now that the standards for the diploma are increased. Because the skills measured by the test have increased, firms will be willing to pay more for workers with a high school diploma, and the benefit will increase. However, it is now also more costly to acquire the skills needed to graduate. The ability level at which individuals decide to acquire the necessary skills to graduate may increase or decrease.

Put differently, the graduation rate can rise or fall. If the additional cost of acquiring the new skills is low and the benefit in terms of market

[7]If passing the test is to some extent a random event or if there is some information about performance beyond merely whether the student passed or failed, the argument becomes more complicated but is similar.

productivity high, then more students will put in the effort to get a high school diploma. In this scenario, all students are better off or at least no worse off with the higher standards. Those who would have graduated anyway will have increased earnings that more than outweigh their additional costs. Those who do not graduate under either standard are unaffected by the change.[8] Finally, those who now choose to graduate but would not under the lower standard are better off since they are now choosing graduation when they could have dropped out and had the same earnings as under the old standard.

On the other hand, if the additional market productivity from the increased standard is low and the added cost of acquiring the skills is high, the situation is more complicated. Again those who would not graduate under either standard are unaffected. Those who would have graduated under the old standard but do not under the new standard are worse off. In addition, at least some of those who continue to graduate under the new standard are worse off. They are getting little benefit from the increased standard but pay a high cost to meet it.

So far we have ignored the possibility that some workers with very low costs of learning will acquire the skills need for college entrance. Of course, just as students with high costs of education weigh the benefit of a high school diploma against dropping out, students with low costs weigh the benefits of college admission against the cost of obtaining the additional skill required for admission.

The increased standards for graduation will lower the wage differential between graduating high school and attending college but will also lower the cost differential. Therefore, in general, we cannot predict whether raising standards increases or decreases college attendance. We should note, however, that the conditions under which higher standards favor college attendance are the same as those that increase the drop out rate. If the higher standards raise the cost of graduation a lot and thus make it similar to the cost of reaching the college entrance requirements and if they do not raise the earnings of high school graduates very much, they will encourage college attendance. If they raise high school graduates' earnings but not the cost of meeting the graduation standards, they will discourage college attendance.

[8]Since the higher standards increase the supply of skilled relative to unskilled labor, this may change their relative prices. We ignore this effect as small at least in the short run since the stock of unsklled and skilled labor is large relative to the flow.

Thus we would expect that increased standards would either increase the high school dropout rate or reduce college attendance. However, this latter outcome is not necessarily bad since it reflects the increased value of a high school diploma.

2.4. More general models. It is possible to make apparently minor modifications to the models that change some of the conclusions about the effect of raising graduation requirements by requiring high-stakes tests. However, these simple models capture the basis for much of the debate around high-stakes testing. In general, there are likely to be winners and losers. Some of the debate around testing concerns how large each of these groups is and how much it wins or loses. Some reflects differences in how concerned people are about the different groups. While it would be comforting to say that there should be no winners and losers in education, the reality is that we make choices whatever policies we choose.

There is a second source of debate not covered by this model. High-stakes tests raise standards only in those areas that are actually tested. If the test has only multiple-choice questions, then students may divert their efforts from learning to write good essays to acquiring the information needed to answer multiple-choice questions. If the test favors five-paragraph essays, students will not learn as much about writing term papers.

In the jargon of economics, this is known as the multitask principal agent problem.[9] Management scientists refer to it as "the folly of paying for A while hoping for B,"[10] Teachers call it "What you test is what you get." In the economics version of the problem, I want you to perform a job for me and that job requires two tasks, one of which I can observe and the other of which I cannot. Depending on the exact nature of the technology, it may or may not be a good idea for me to pay you on the basis of the task that I observe. For example, suppose I want you to produce as many units of high-quality output as possible but cannot easily determine whether the quality is high. If I pay you piece-rate (i.e. on the basis of the quantity you produce), I encourage you to work fast and produce a lot of output. If the quality you produce is very sensitive to the speed at which you work, you will produce a lot of low quality output which is not what I want. If the quality you produce is

[9]Holmstrom, Bengt and Paul Milgrom, "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," J*ournal of Law, Economics, and Organization.* 7 (Special Issue 1991): 24-52.

[10]Steven Kerr, "On the folly of rewarding A, while hoping for B." *Academy of Management Journal*, 18 (December 1975): 769-783.

relatively insensitive to the speed at which you work, paying piece-rate may be a good idea.

In the case of testing, the issue is whether schools "teach to the test." Whether teaching to the test is good will depend a great deal on the degree to which the test does a good job of covering the full range of skills that we want students to develop. If we test only English and math, schools may neglect science and social studies. If we test science, they may not teach how to conduct experiments if it is not easily tested and does not prepare them for the exam. On the other hand, testing can promote good teaching if the best way to prepare students for the exam is by having them perform experiments even if the ability to do experiments is not tested directly.

2.5. Evidence. It is difficult to obtain good evidence on the effects of testing. John Bishop and Ferran Mane[11] found that individuals from states with high-stakes tests (in 1992) were less likely to graduate high school and more likely to have passed a high school equivalency exam (GED). The increase in the number of students passing the GED was about equal to the decrease in the number graduating high school. However, since the GED is generally viewed as inferior to a high school diploma,[12] this suggests that the high-stakes tests decrease educational attainment at the lower end of the distribution.

Bishop and Ferran also found that individuals in states with high-stakes test have higher earnings for a given level of education as predicted by the theoretical models. However, this effect operates through working more rather than through higher hourly wages which is not predicted by the model.

The main difficulty with the Bishop and Ferran study should be familiar by now. Why do some states but not others choose to have high-stakes tests? If states with big dropout problems are less likely to have high-stakes tests, then the effect of testing on the dropout rate will appear to be smaller than it really is. On the other hand, if high-stakes tests are more politically acceptable when high school graduation is less common, we might falsely attribute the higher dropout rate in these states to the high-stakes test.

Therefore we will focus on the case of Texas which has implemented a series of reforms since the mid 1980s. We choose Texas for two

[11]Bishop, John H. and Ferran Mane, "The Impacts of Minimum Competency Exam Graduation Requirements on High School Graduation, College Attendance and Early Labor Market Success," *Labour Economics*, (May 2001): 203-22.

[12] Cameron, Stephen V. and James J. Heckman, "The Nonequivalence of High School Equivalents," *Journal of Labor Economics*, 11 (January 1993): 1-47.

reasons. First, it has been studied extensively. Second, it is a large state which makes it easier to get data with which it can be compared to other states. However, while the Texas reforms included high-stakes testing, Texas also initiated a number of other reforms from requiring a minimum GPA in order to be eligible to play sports to establishing an accountability system with rewards and punishments for schools and school districts. It is difficult if not impossible to separate the effects of the high-stakes testing from the effects of the other reforms. In particular, since the graduation rate is one element of the accountability system, schools may try to reduce the dropout rate even if the direct effect of high-stakes testing increases it.

The most recent reform in Texas involved the creation of the Texas Assessment of Knowledge and Skills (TAKS). This replaces Texas Assessment of Academic Skills (TAAS) which was the subject of much study and controversy. Beginning with the class of 2005, all students will be required to pass 11th grade TAKS exams in English Language Arts (reading/writing), mathematics, science and social studies. Because achievement generally rises rapidly in the first few years that a test is administered, Texas will use lower requirements for attaining a passing grade in 2005 and 2006.[13] The TAKS mathematics exam now has a very small number of open response questions (one in spring 2004) while the ELA test has a small number of questions requiring a brief essay response.

We will focus on the TAAS since the TAKS has been implemented too recently to permit any evaluation. While we will treat high-stakes testing as beginning with the TAAS, Texas had a statewide test prior to the TAAS. However, the test was viewed as easy and not a real hurdle for graduation.

The TAAS was quite different. It was viewed as difficult. The fall 1990 pretest of the TAAS revealed that the passing rates would be in the 40-60% range on TAAS (with pass rates for Black and Hispanic students on the math portion of TAAS falling to the 27-33% range). The test was first administered as a high-stakes test to 11th graders in fall 1991. Starting in spring 1993, the test was given to 10th graders. Students had up to 10 opportunities to pass the 10th grade exam which they were required pass in order to graduate.

Starting in 1994, the TAAS Reading, Mathematics and Writing tests were administered to students in grades 4, 8 and 10 in the spring of each year. The focus of test-based accountability in Texas was on

[13]Texas Education Agency, *Student Guide to Graduation*, http://www.tea.state.tx.us/student.assessment/resources/grad/grad_broch.pdf

the TAAS tests of reading, mathematics and writing (there are also TAAS tests of social studies and science and end-of course tests in some high school subjects). The TAAS tests were mostly multiple-choice in format with the exception of the writing exam.

Following its implementation, scores on the TAAS improved dramatically. Stephen Klein and his coauthors at the Rand Corporation[14] compared the dramatic improvement on the TAAS with improvements on the National Assessment of Education Progress (NEAP). While TAAS scores shot up, 4th grade scores on the NEAP rose only somewhat, not dramatically, faster in Texas than elsewhere and gains at 8th grade were similar to the rest of the country. Their findings raise serious questions about the meaning of these test score improvements. It appears that the high-stakes nature of the test creates strong incentives for improvement on the TAAS but which are not reflected in other measures of performance. Their other major conclusion is that while the performance gap between minority and white students has narrowed on the TAAS, if anything, it increased on the NEAP.

David Grissmer and his colleagues,[15] also at Rand, reached a somewhat different conclusion. They also studied the results of the NAEP and determined that test scores grew more rapidly in Texas than in the United States as a whole. They suggest that the Texas accountability program, including TAAS, is a plausible source for this faster growth.

On the basis of these studies, it is reasonable to conclude that the Texas education reform did increase performance on standardized tests. It increased performance on the assessment instrument (TAAS) significantly more than on other standardized tests but since some of what is measured on the TAAS is similar to the NAEP, some of the improvement carried over to that test.

What is not at all clear is whether the reform improved performance on the types of skills that are not measured on such tests.

As discussed above, if raising standards has a small effect on the cost of attaining the standards and a large effect on the value of the diploma, we expect raising standards to decrease the dropout rate. If the increase in costs is high and the benefit low, we expect the opposite. In neither case does the change in the dropout rate demonstrate that the policy reform is good or bad. Reforms almost definitely have winners and losers.

[14]Klein, Stephen P. et al "What Do Test Scores in Texas Tell Us?" *Rand Organization Issue Paper No. 202*, 2000.

[15]Grissmer, David W., et al, *Improving Student Achievement: What State NAEP Test Scores Tell Us*, Santa Monica, CA: Rand Corporation, 2000.

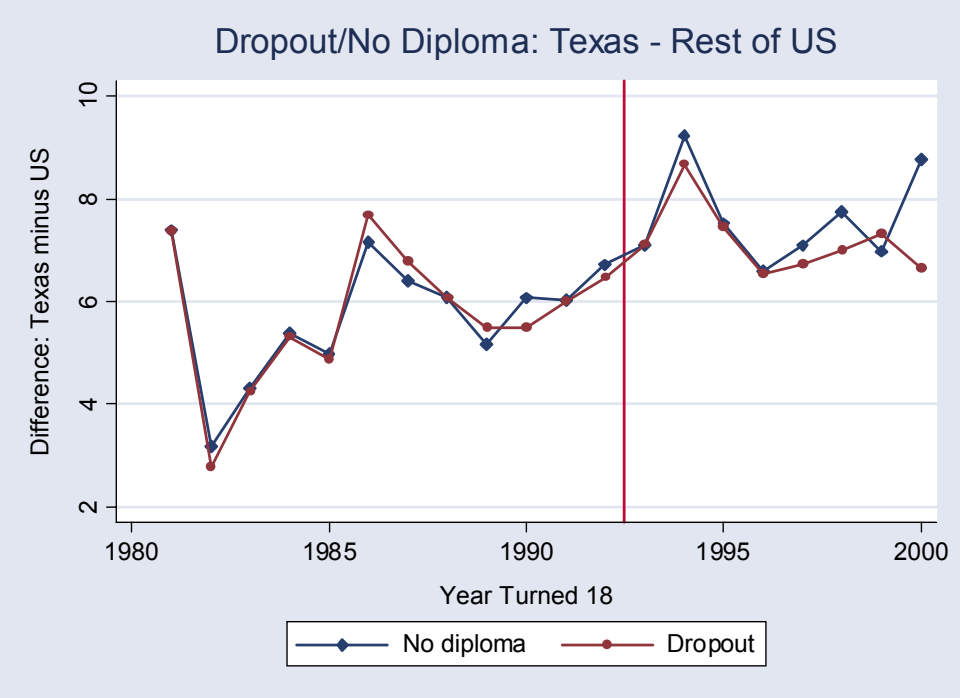

Figure 2:

Texas has periodically changed the way it records the dropout rate so that it is difficult and perhaps impossible to use official Texas data to determine the effect on the dropout rate.[16] We can look at educational attainment by age group in Texas and the rest of the United States to see if we can detect an effect on the dropout rate.

I use the Current Population Surveys from 1992 through 2002 to calculate the proportion of individuals with less than a 12th grade education (dropouts) and the proportion who do not have a high school diploma or GED.[17] The horizontal axis is the year in which the individual turned eighteen. The vertical axis is the difference between Texas and the rest of the United States. Thus we see that in 1981 both the dropout and no diploma rates were between 7 and 8 percentage points higher in Texas than in the rest of the United States. The vertical line divides the cohorts affected by the graduation requirement from the cohorts that were not covered by the requirement.

[16]Haney, Walt, "The Myth of the Texas Miracle in Education," *Education Policy Analysis Archives*, 8 (August 19, 2000): No. 41.

[17]The sample is limited to those 19-30 years old at the time of the survey.

Figure 2 shows some of the difficulty of deciding whether the dropout rate increased in Texas as a result of the introduction of the TAAS. Dropout and no diploma rates are about 1.5 to 2 percentage points higher to the right of the vertical line than to the left of it. On the other hand, if we ignore 1981, it is easy to see the post 1993 pattern as a continuation of a pre-1993 trend towards higher dropout rates in Texas relative to the rest of the country. And just to make things complicated, it might be argued that the pre-1993 trend reflects the effect of earlier education reforms in Texas.

Further confounding the problem is that the Current Population Surveys group individuals who received high school equivalency by passing the GED exam with those receiving a regular diploma. There is evidence that after the TAAS graduation requirement was established, the use of the GED increased in Texas relative to the rest of the United States.[18]

Thus there is no evidence that the TAAS and accompanying education reforms reduced the dropout rate, but the evidence for an increase in the dropout rate is somewhat "in the eye of the beholder."

In summary, the Texas education reform with its focus on raising standards and accountability through the TAAS

(1) Improved student performance on those aspects of education measured by the TAAS
(2) Had an unknown but certainly smaller and possibly negative effect on aspects of education not measured by the TAAS
(3) Probably somewhat increased the dropout rate.

3. Decentralization and School Quality

While so far we have focused on high-stakes testing for students, much of the focus of education policy is on how to make schools better. Some of these, notably charter schools, vouchers and school choice are designed to change the link between residence and where students attend school. Before we examine these policies, we must examine the theory that supports the traditional decentralization of the U.S. public school system.

Historically, education in the United States has been mostly the responsibility of local government. State government played a supporting role, and the federal government was even less involved. Most of the funding for education came from the local community and was usually raised via taxes on property.

[18]Haney, *op cit.*

By international standards this is a highly unusual way to fund education. It has advantages and drawbacks. The major advantage is that since communities paid for education, relative to most developed countries, they were free to offer education in the way that they favored. The state might require a minimum number of days of school, but the community was free to offer more days. It could choose class size, teacher pay and many aspects of the curriculum. It could choose to put money into the football program or into performing arts.

The economist Charles Tiebout[19] showed that, under some restrictive assumptions, the competition among communities worked to create a market for goods like education. To see how this works, suppose that all families have exactly one child who goes to school. Suppose that there is just one varying element of school quality, for example class size. Once it has chosen class size, then the community has established how much it will cost to educate each student. A class size of 30 might involve a cost of $4,500 per student while a class size of 20 might cost $6,000 per student.

In order to enroll their child in a school, parents must pay two prices. They must pay a fee to become part of the community. This fee is the cost of a house (or the value of the land on which the house is located). The other price is the taxes that residents must pay.

Let us consider the taxes first. Since people prefer not to pay taxes, the community will never set the tax above the cost of providing the education. On the other hand, if it sets the tax below the cost of providing the education, the community will not be able to pay the teachers and other costs. Therefore, the tax will be exactly the cost of education per child.

Now let us think about how much people will have to pay to buy a house. If two communities offer the same school quality, but one charges higher taxes, people will prefer the community with the lower tax. Therefore the price of housing will rise in the low-tax community and the price of housing will fall in the high-tax community until people are just indifferent between paying the higher cost of housing and lower taxes, on the one hand, and paying less for the house but higher taxes, on the other. Members of a community know that if their community is inefficient and pays too much for a given level of school quality, their housing values will fall. Therefore, competition among communities creates an incentive for communities to produce school quality efficiently.

[19]Tiebout, Charles M. "A Pure Theory of Local Expenditures," *Journal of Political Economy*, 64 (October 1956): 416-424.

But the competition among communities does more than just provide an incentive for efficient production. It also pushes communities to provide the levels of school quality that people want. Suppose that everyone is willing to pay an extra $2,000 per year to have twenty students per class instead of thirty students per class. If one community offers classes of thirty and taxes of $4,500 and another offers classes of twenty and taxes of $6,000, people will prefer to live in the community with smaller classes. They will bid up the price of houses in the community with the smaller classes and bid down the price of houses in the community with the larger classes. Because they care about the value of their houses, owners will pressure their communities to have classes of twenty rather than thirty.

Of course, not everyone has the same tastes. Some prefer to pay $4,500 and have large classes while others prefer to pay $6,000 and have small classes. If too many communities offer large classes, then people who want small classes will have difficulty finding houses in the communities they like and will bid up their price. On the other hand, if too many communities offer small classes, prices will be bid down in those communities. Residents, interested in maximizing the value of their homes, will respond to the market incentives. Some communities will offer high-quality education and require high tax payments while others will offer low-quality education but require only low tax payments.

In contrast with countries in which decisions about education are made by the central government, the U.S. system allows people to choose among communities with differing school qualities. Those who value school quality a great deal will live in communities with good schools while those who place less value on education will live in communities with weaker school systems. Moreover, since school quality is not one dimensional, families can choose the types of school that they like. One family may choose to move to a school with a strong traditional curriculum while another prefers the school with more experimental programs.

The real world is much more complicated than this simple story suggests. The taxes that people pay in their communities are not some fixed amount but typically depend on the value of their house. In the same community, people with expensive houses pay more taxes than people with inexpensive houses. Some people have more than one child while others have none. Since we base taxes on the value of property not on the cost of educating the family's children, some people more than pay for the cost of educating their children while others do not cover this cost.

Still, local control of education has two distinct advantages.

(1) It creates competition among communities so that they attempt to provide education efficiently and in a manner that responds to the preferences of the community. Communities that do this effectively will have higher housing values.
(2) And, it maintains public support for funding education – even people who do not have children in the public schools may support funding for public education because better schools raise property values.

Does the market for school quality work the way that the theory suggests? We will examine the relations among school quality, competition and housing values. The question of political support and school funding is complex and is not central to the study of poverty.

3.1. Competition and the Efficiency of the Provision of School Quality. Caroline Hoxby[20] has studied the relation between the extent of choice among school districts in a metropolitan area and the performance of the schools in that area. More district choice has a mixed relation to student performance, raising student performance along some measures and lowering it along others.

However, the number of districts in a metropolitan area is, to some extent, a matter of choice. Politicians might respond to dissatisfaction with school districts by merging them or by breaking them up. In the former case, poor performance would lead to fewer districts.

Therefore Hoxby looked for physical features of the metropolitan area that influence the number of districts. Historically, rivers were major barriers to travel and continue in many cases to slow movement between areas. Hoxby finds that metropolitan areas with more rivers have more school districts. She predicts the number of school districts based on the number of rivers and other factors. She then looks at the relation between this predicted school district concentration and student performance. She finds that measures of student performance are consistently higher when there is more competition among districts.

This does not by itself prove that the schools are more efficient. It could be that 60% of residents prefer low taxes and low school quality and 40% prefer high taxes and high quality. When there is only one district, the majority votes for low taxes and quality. When there are two districts, those who want low taxes and low quality live in one district and vote for low quality and low taxes as before, and those

[20]Hoxby, Caroline M., "Does Competition among Public Schools Benefit Students and Taxpayers?" *American Economic Review*,.90 (December 2000): 1209-38.

who want high quality and taxes live in the other district and vote to provide these. The move from one district to two districts improves student performance but does not necessarily increase the efficiency of the school districts.

However, Hoxby goes on to look at school spending. She finds that increased competition among districts lowers school spending while reducing class size. Thus the improved student performance cannot be attributed to increased spending. Moreover, she finds that increased choice is associated with fewer students attending private school.

Hoxby's results are certainly consistent with the expectation that more choice will push schools to be more efficient and to respond more to the preferences of the community. However, more choice may also be associated with more community control and involvement. When school districts are more compact, schools may be more effective regardless of competition.

3.2. School Quality and Housing Prices. There is an extensive literature on the relation between house prices and school quality. In general, neighborhoods with better schools also have higher house prices. However, this relation is not informative. People with higher incomes will tend to spend more on housing. Because of the advantages associated with being in a richer family, their children will also tend to do better in school. It is not surprising that wealthy areas both have schools with high test scores and expensive homes. Put differently, a neighborhood with expensive single-family homes will have children who will, on average, do better in school compared with children in a neighborhood with small rental apartments even if the children in the two neighborhoods attend the same school.

Sandra Black[21] reduces this problem by looking at districts with houses located on the boundaries between school attendance districts. Within school districts such boundaries are often drawn in very arbitrary ways and divide neighborhoods. The houses on one side of the street are likely to be similar to those across the street. The houses and neighborhoods are therefore very similar except that the children on different sides of the street attend different schools. She finds that a 5% increase in test scores increases the price of housing by 2.5%.

[21]Black, Sandra E., "Do Better Schools Matter? Parental Valuation of Elementary Education," *Quarterly Journal of Economics.* 114 (May 1999): 577-99.

Katharine Bradbury and her coauthors[22] used a somewhat different approach. They looked at the effect of a tax limitation measure on housing prices. Massachusetts Proposition $2\frac{1}{2}$ constrained some but not all communities to spend less on education than they would otherwise have chosen. The authors found that property values rose less rapidly in communities that were forced to restrain their spending, particularly their education spending.

3.3. School Choice. People who can afford to live in the suburbs often have considerable choice about where to attend school. If they are unhappy with their current school, they can move. For the middle class, the Tiebout model may apply well. Moreover, voters in suburban school districts may recognize a strong link between their housing values and school spending.

In contrast, poor people living in high-poverty areas often have little choice. The school choice movement aims to give poor people a range of choices. By promoting competition among schools, the movement hopes to imitate the benefits of Tiebout competition. Of course, such choice will not replicate the link between housing values and school quality. Indeed, as we will see, it may weaken or sever the link.

There are a number of proposals for enhancing school choice:

- Residential neighborhood can be separated from school. Within a district, students may be able to apply to more than one school rather than simply attending a school assigned to them on the basis of neighborhood boundaries. The district then uses parents' preferences when assigning students to schools. Neighborhood may or may not play a central role in this assignment. The school choice may also be extended beyond districts.
- Vouchers provide more choices by allowing students to attend private (and in some cases, parochial) schools. In a typical program, if a student is accepted by a private school and chooses to attend, the school accepts the amount of the voucher as full payment.
- Charter schools are privately run public schools. They may be run as for profit or as non-profit organizations. Since they are privately run, they are typically exempt from some state regulations and are expected to be more experimental than

[22]Bradbury, Katharine L. and Christopher J. Mayer and Karl E. Case, "Property Tax Limits, Local Fiscal Behavior, and Property Values: Evidence from Massachusetts under Proposition 2 1/2," *Journal of Public Economics*, 80 (May 2001): 287-311.

regular public schools. Usually, the charter school receives a direct payment from the state for the students it enrolls and is not permitted to accept private students.

We examine each of these approaches in turn.

3.4. Interdistrict and Intradistrict School Choice. Intradistrict school choice has been primarily associated with racial desegregation of schools rather than with an attempt to increase competition among schools. There are some exceptions including a mixed neighborhood/choice program in New York City. We discuss it only briefly since the major focus of more recent reforms is on choice across districts.

In a typical intradistrict school choice plan, parents are asked to list in order the schools in the district that they would like their children to attend. The district then assigns the children to schools on some basis that partially reflects parent preferences. For example, the district might decide that the first criterion is that children have the right to attend schools where they have siblings. After assigning all children whose first choice school is attended by a sibling, the district might then look at the racial, sex and ethnic balance of the schools. It would then assign students to their first choice school if it improved the balance along these lines. Students are not guaranteed to be placed in their preferred school or even in one of their top choices.

Because public schools are required to serve all the children in a district, intradistrict school choice has little effect on competition across schools. School districts typically try to avoid having excess capacity. If the number of children attending public school in the district is approximately equal to the number of openings in the district, then all schools will be full. Even if a school is placed last by every parent, the school district will still assign some students to that school. Of course, the district may respond by changing the principal, giving the school more resources or in some other way, but the staff does not automatically suffer if their school is not highly regarded by parents.

We can see that intradistrict school choice gives some people more choice and others less. When school assignment is based on neighborhood, anyone with sufficient resources can choose her children's school by buying or renting housing in the neighborhood. School choice eliminates the guarantee that she can choose that school.

On the other hand, housing will generally be more expensive in the neighborhoods with schools that many people find attractive. For poor people, living in such neighborhoods may not be a realistic option or

may involve considerable sacrifice. A poor person living in one neighborhood can select a school in a different neighborhood and have some chance of having her child enrolled in that school.

Whether this increased choice helps poor children depends on a number of factors including how important peers are in influencing school quality. Upper income parents are likely to respond to their failure to get their first choice school by leaving the district, either by moving or sending their children to private school. This not only removes more advantaged students from the schools, but may weaken the public schools politically. When there are already lots of alternatives for upper income families, intradistrict school choice is therefore less likely to be effective in helping poor children. Intradistrict school choice is much more attractive in large school districts such as New York City where many of the children of the well-to-do already attend private schools.

3.4.1. *Intradistrict Choice in New York City*[23]. Community School District 4 is one of thirty-two districts within the New York City public school system, the largest in the United States. It is located in East Harlem, a ghetto and barrio neighborhood. In the 1970's, the district initiated a reform program which gave schools considerable power to reform themselves and to open new schools. Schools were distinguished from school buildings, and one junior high school building could house more than one independent junior high school. There was considerable expansion of the number of separate schools.

Starting at the junior high school level, parents were allowed to select which school their children attended. They ranked their top choices and were considered for admission by the first school on their list. If they were turned down, they were considered by the second and then, if necessary, by the third. If they were not admitted to one of the top three choices (5% of applicants in 1992), they met with administrators to agree on a placement.

During this period student performance soared. More students gained placement in the city's elite high schools and standardized test scores went from being the lowest of the 32 districts to being well above average. After the initial gains, there was some slippage. As of 2002, District 4 scores on the state 8th grade math and English tests placed

[23]The following is drawn heavily from Cookson, Peter W., Jr., and Sonali M. Shroff *School Choice and Urban School Reform*, Teachers College, Columbia University, December 1997, http://eric-web.tc.columbia.edu/monographs/uds110/, and their "Recent Experience with Urban School Choice Plans," *Eric Clearinghouse on Urban Education Digest*, Number 127, (October 1997), http://eric-web.tc.columbia.edu/digests/dig127.html.

it 29th and 22nd among the system's 32 regular districts as measured by the proportion getting a 3 or 4 (proficiency or better) on the tests.[24]

It is apparent that the reform of District 4 was a success although perhaps less dramatic than is sometimes claimed in political circles. What is less obvious is how much of a role choice played and how much must be attributed to the role of inspired education reformers and the transformation of the programs, particularly shifting to smaller schools. At the same time, it is questionable whether the spirit of reform could have been fostered without the opportunity for choice.

3.5. Interdistrict School Choice. Interdistrict school choice differs from intradistrict school choice in two important ways:

(1) There is a home district that has responsibility for educating all of the students in the district. Districts may accept students on a space-available basis. Students are not guaranteed a spot in a district other than their own.
(2) In all such programs funds flow to the district that receives out-of-district students and in most programs, funds are subtracted from the district from which those students are sent. As a result, it is possible for the receiving district to benefit financially and the sending district to be harmed.

Interdistrict school choice simultaneously increases and decreases competition among school districts. On the one hand, if a school district provides high-quality education, it can attract students from other districts. If it costs less to educate those students than the district receives for educating them, the district receives a fiscal windfall. It can provide more services to its residents or reduce taxes.

On the other hand, the Tiebout model we discussed earlier relied on the fact that people have to obtain housing in a district in order to send their children to its schools. If parents can send their children to a school without living in the district, this will reduce demand for housing in districts with good schools and raise demand in districts with poorer schools. If the relation between housing prices and school quality is weakened, this may reduce support for funding local schools.

Minnesota was the first state to initiate statewide school choice. Under the Minnesota Open Enrollment Option, any student can apply to any public school in the state. If the district believes that, after serving its own students, it has sufficient space, it accepts the student. If it has sufficient space for some but not all of the out-of-district students who apply, it must choose among them using a lottery. The district is

[24] http://www.nycenet.edu/daa/test_results/

responsible for transportation from the district boundary to the school, but parents (with some state support for poor families) are responsible for transportation between home and the district boundary. State funding for the pupil is transferred from the student's home district to the district where he attends school. State funding is considerably less than the average cost of education but may or may not be greater than the marginal cost of educating an additional student.

Unfortunately although the program has been in effect for more than a decade, there is no clear evidence of its effectiveness. However, there have been two studies of the effect of the program on property values.[25] Both find property values increased in districts that lost students while they decreased in districts that accepted students from outside. In fact, Reback calculates that the increased property tax from the higher property values was about what would be needed to offset the lost revenue in communities losing students. Thus the interdistrict choice program seems to have, if anything, reduced the incentive of school districts to compete efficiently for students.

3.6. Vouchers. Voucher programs provide students with vouchers that can be used to pay for tuition at private, and, in some cases, parochial schools. The voucher programs that have been tried in the United States have been oriented towards poor children. In principle, everyone could receive a voucher. We will discuss universal voucher systems later since any such discussion is largely speculative.

William Howell and Paul Peterson[26] have evaluated voucher experiments in New York, Dayton and Washington, DC. All three were privately funded voucher programs that provided subsidies for poor children to attend private or parochial school. In each case more children applied to the program than could be served and vouchers were allocated on the basis of a lottery. They found that, after two years, African-American students offered a voucher did significantly better on

[25]Reback, Randall, "Capitalization under School Choice Programs: Are the Winners Really the Losers?" National Center for the Study of Privatization in Education Teachers College, Columbia University, Occasional Paper No. 66, 2002, and Hoyt, William E, ""Public School Choice, Property Values, and Investment in Public Education", University of Kentucky, unpublished, 1997.

[26]Howell, William G. and Patrick J. Wolf and Paul E. Peterson and David E. Campbell, "Test-Score Effects of School Vouchers in Dayton, Ohio, New York City, and Washington, D. C.: Evidence from Randomized Field Trials," Paper presented at the annual meetings of the American Political Science Association, Washington, D. C., September 2000. See also Howell, William G. and Paul E. Peterson, *The Education Gap: Vouchers and Urban Schools*, Washington, DC: Brookings Institution Press, 2002.

the Iowa Test of Basic Skills than did African-American students who were not offered a voucher. The implied effect of using a voucher was about 6 national percentile points. In contrast, there was no significant difference between members of other race/ethnic groups who did and did not receive a voucher.

The New York City experience has been studied most extensively[27] and has been the basis for considerable controversy. For this program, a private foundation offered vouchers worth $1,400/year each to over 1,200 students selected randomly from over 20,000 applicants. Of those offered a voucher, 74% used it the first year, 62% the second and 53% the third. Inevitably, experiments are imperfect. Some families offered vouchers do not make use of them. Researchers lose track of some of these families and of some families that did not receive vouchers. Or they are unable to collect some information for some families. Moreover, there is no agreed upon definition of race. The results of the New York City experience are somewhat sensitive to decisions about how to deal with these difficulties. A fair summary of the debate is that most choices of research strategies lead to the conclusion that the vouchers had a modest positive effect on the recipients. However, there are choices that lead to the conclusion that there were no statistically significant effects and others that the effects were large.

For the most part, the voucher programs that have been tried in the United States have been quite modest both in terms of the level of the voucher and the number of students involved. The major exception is the Milwaukee Parental Choice Program which in 2002-2003 served over 11,000 students. Participating schools received up $5,783 per student in the program.[28] Participation is limited to 15% of enrollment in the Milwaukee public schools and only families with incomes no higher than 1.75 times the poverty rate are eligible. Approximately half of the participating schools are parochial.

[27]Peterson, Paul E. and William G. Howell, "Efficiency, Bias, and Classification Schemes: Estimating Private-School Impacts on Test Scores in the New York City Voucher Experiment Latest Results from New York City School Voucher Research – African American in Private Schools Score Higher, *Program on Education Policy and Governance Working Paper PEPG/02-15 - 22,* 2002, and Krueger, Alan B. and Pei Zhu, "Another Look at the New York City School Voucher Experiment," *NBER Working Paper No. 9418*, 2002.

[28]http://www.dpi.state.wi.us/dpi/dfm/sms/choice.html

Cecilia Rouse[29] (1998) tried to evaluate the effect of the Milwaukee program on the participants. Her evaluation is based on the early years of the program when spaces were very limited. Perhaps the most striking result of her evaluation was the difficulty she had in finding a comparison group even early in the program when participation was much more limited. Although places were allocated on the basis of a lottery, most of those who failed to win the lottery subsequently left the Milwaukee public schools anyway, either because they won a subsequent lottery, used their own resources to pay for tuition or left Milwaukee. This raises the question as to the degree to which vouchers are largely a subsidy to families that would not have attended the inner-city public schools in any event.

Thus there is room for those with strong feeling on each side to find data to support their views. However, in some sense the debate over the New York City voucher experiment misses the point. The arguments for the traditional system of local control in the United States and for new types of competition are both based on the assumption that parents can make intelligent choices on behalf of their children. It would be surprising to both groups if large numbers of parents chose to send their children to schools that did not improve their educational outcomes. If they did, we would have to conclude either that parents care about different outcomes from the ones measured by the tests used to evaluate the voucher program or that they could not assess school quality.

The real issue is whether the competition generated by the existence of vouchers improves education for students who do not use vouchers. On this point there is almost no evidence.

3.7. Charter Schools. Charter schools are public schools that are permitted to operate with considerably less regulation and direct supervision than is typical for public schools. While the precise rules vary among states, in general, a charter school must establish a clearly defined set of objectives and a set of measures that will help determine whether it has met those objectives. Schools that meet their objectives may continue; those that fail to meet their objectives will be closed. Charter schools are generally free from having to hire unionized teachers and may have exemptions from teacher certification requirements. In principle, the greater flexibility afforded charter schools should lead to more innovation.

[29]Rouse, Cecilia E., "Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program," *Quarterly Journal of Economics*. 113 (May 1998): 553-602.

Setting up charter schools is different from giving students vouchers because the state (or other authority) has the power to close the school if it is not proving effective. The school cannot accept private tuition-paying students.

In other ways, the charter and voucher movements are similar. Each is designed to create more choice for parents and their children. Each is designed to foster competition that is intended to improve education in the regular public schools.

There are two questions that we ask about charter schools:

(1) Do charter schools improve the performance of the students who attend?
(2) Does competition from charter schools improve the performance of the regular public schools?

3.7.1. *Charter Schools and Student Performance.* One of the states that has made the most extensive use of charter schools is Michigan. Eric Bettinger[30] compared eighteen Michigan charter schools with 546 regular public elementary schools and 178 regular public middle schools located within five miles of one of the charter schools. He used results from Michigan's statewide tests to determine student performance.

How can we tell whether the Michigan charter schools improved student performance? One way we might consider answering this question is by comparing students who attend charter schools with those who attend regular public schools. Student scores two years after the founding of the charter schools were substantially worse at the charter schools than at the regular public schools. The proportion of fourth graders getting a score of satisfactory was 17 percentage points lower on the math exam and 12 percentage points lower on the reading exam. Among seventh graders, the proportions were 18 percentage point and 13 percentage points lower in math and reading. There are symmetric differences among the proportions getting a low grade. Students in charter schools were much more likely to receive a low score two years after the charter was founded (see table 3).

However, charter schools are likely to attract students from low-performing public schools. These students will be disproportionately disadvantaged, and it will not be surprising if they do not perform, on average, as well as the average regular public school student. On the other hand, students in charter schools may come from particularly motivated families. If we do a good job of controlling for differences

[30]Bettinger, Eric P, "The Effect of Charter Schools on Charter Students and Public Schools," unpublished paper, Case Western Reserve University, 2002.

DIFFERENCES BETWEEN CHARTER AND OTHER PUBLIC SCHOOLS IN PROPORTION OF STUDENTS SCORING "SATISFACTORY" AND "LOW" ON MICHIGAN EDUCATIONAL ASSESSMENT PROGRAM				
	Math		Reading	
	4th Grade	7th Grade	4th Grade	7th Grade
Levels				
Satisfactory	-17	-18	-12	-13
Low	12	17	10	11
Changes				
Satisfactory	1.9 (5.2)	-0.2 (6.7)	-4.1 (5.0)	* (5.9)
Low	-7.0 (4.5)	-3.4 (6.9)	0.1 (4.7)	-2.0 (5.7)
Changes Adjusted for Initial Scores				
Satisfactory	-10.5 (3.4)	-0.6 (4.2)	-8.9 (4.1)	0.3 (2.9)
Low	7.4 (4.3)	6.2 (4.3)	6.6 (3.2)	-0.2 (2.7)

*Less than 0.05

Source: Bettinger, Eric P, "The Effect of Charter Schools on Charter Students and Public Schools," unpublished paper, Case Western Reserve University, 2002.

Figure 3:

in family background, we may find that students in charter schools do better simply because we cannot measure parental motivation.

Bettinger therefore looks at whether the test scores of charter school students improved more or less than the test scores of students in comparison schools. When he looks at changes in scores, there is much less evidence of a difference, and the results, if anything, suggest that charter schools may outperform regular public schools. None of the differences is statistically significant, but, for three of the four tests, charter schools saw a greater decline in the fraction scoring low and there is essentially no difference between charter and other public schools on the fourth test. While we would not want to draw strong conclusions from statistically insignificant results, these findings would certainly push us to explore further the possibility that charter schools improve the performance of the lowest performing students.

However, recall that when we discussed training programs, we raised the concern that the people who enter programs often have had unusually low earnings before entering. People who have had decent jobs and then suffered long-term unemployment may be more likely to enter training programs and appear to "improve" as a result of the training program. Just as sick people often get better on their own, people who are "down on their luck" often recover through their own efforts and changes in their luck.

This same phenomenon may apply to charter schools. Students whose performance in the public schools is below what they anticipate may apply to a charter school. Good students who have had a few years of bad teaching or bad behavior may apply. Many of these students would have had improved experiences in regular public schools. Bettinger attempts to address this by looking at changes controlling for initial scores. His results suggest that, at least among fourth graders, the charter schools did worse than the regular public schools. Controlling for their poor initial level of performance, over two years, the charter school fourth graders showed less improvement than did fourth graders in regular public schools.

Caroline Hoxby and Jonah Rockoff[31] studied seven charter schools in Chicago. Because the schools are over-subscribed, students are admitted by lottery. Overall they found no difference in outcomes between lottery winners and lottery losers. When they broke the students down by the grade for which they applied, among those applying for kindergarten or first grade, lottery winners did somewhat better in reading but not in math. Among those applying for second or third grade, winners did better in math but not in reading. They found no difference between winners and losers applying for the fourth and fifth grades or the sixth through eighth grades.

We should not read too much into these studies. After all, together Bettinger's and Hoxby and Rockoff's samples are limited to twenty-five charter schools in two states. Charter schools may have performed better in some states with different charter school regulations. Moreover, it is undoubtedly true that just as there are good and bad regular public schools, there are good and bad charter schools. Some charter schools are founded by talented and creative leaders. Others are led by principals who lack the skill to run a high-quality school. Two years is a short period in which to create a first-rate school. What this research does show is that we should not equate charter school with good school.

[31]Hoxby, Caroline M. and Jonah E. Rockoff, "The Impact of Charter Schools on Student Achievement," unpublished, Harvard University, 2004.

This conclusion is reinforced by a Rand study of charter schools in California.[32] The study found similar levels of performance and improvement in performance between charter schools and regular public schools once they controlled for differences in student backgrounds. However, it found considerable variation in performance among charter schools. In particular, it found that charter schools with a significant out-of-school instructional component performed less well. The authors are careful to note that such schools may attract very different students from those who attend traditional public and other charter schools and that therefore the poorer performance may not reflect worse school quality.

3.7.2. *Effects on Regular Public Schools.* Perhaps the most innovative aspect of Bettinger's[33] research on Michigan charter schools is his attempt to find their effect on other public schools. This is a difficult task. People are likely to want to found charters and charters are likely to be granted in areas where regular public schools are viewed as weak. If we compared public schools in areas with lots of charters with those in areas with no charters, we would almost definitely find that public schools were weaker where there were charters. This would not prove that charters hurt public schools, only that charters were founded where public schools were weak.

Bettinger recounts that the governor put great pressure on public colleges and universities in Michigan to found charters. As of 1999, 150 of 170 charter schools in Michigan were founded by colleges and universities. Since institutions generally open charter schools near their own location, in Michigan charter schools were, on average, located in areas where school test scores were slightly better than average.

Bettinger looks at changes in the performance of public schools and finds no consistent evidence than public schools located near charter schools exhibited smaller or greater improvement that schools located farther from charter schools. He is able to establish that any effect, whether positive or negative, was not large.

3.8. Universal Vouchers. In principle, a system of vouchers could replace the present system of neighborhood schools. Under such a system, schools would play much the same role that health maintenance organizations (HMOs) do in our public health programs. Just as HMOs receive a payment from Medicare or Medicaid for each eligible person

[32]Buddin, Tichard and Ron Zimmer, "Academic Outcomes," in Zimmer, Ron et al eds., *Charter School Operations and Performance: Evidence from California,* Santa Monica, CA: Rand Corporation, 2003.

[33]Bettinger, *op cit.*

they enroll, the school would receive a fixed voucher payment for each pupil which might be based on pupil characteristics. In some cases, the fixed payment might be supplemented for particularly expensive special education just as health-care providers currently receive supplemental payments for certain expensive procedures.

We cannot draw on our experience with vouchers to determine how such a system would work. No voucher program in the United States has been sufficiently extensive to approximate what would happen if vouchers replaced public schools. We can speculate that a voucher system would have many of the strengths and weaknesses of a public health system in which people are free to choose their HMO. On the positive side, there would be considerable competition among schools for profitable students. On the negative side, if they were free to select students, schools would turn away unprofitable students. If not, they would tailor their programs to attract the students who were most profitable.

Without knowing the details of the voucher system, it is impossible to know in advance precisely how this would work. One possibility is that all regular education students would received the same voucher while special education students would receive a voucher for a greater amount that depended on the precise nature of their special education status. If poor children were more expensive to educate, schools could follow policies designed to discourage poor students from applying – locating far from poverty neighborhoods, refusing to participate in the school lunch program, etc. On the other hand, if the voucher payment for poor students were sufficiently high relative to other students, then schools would try to attract poor students. More generally, within any group for which the voucher payment is constant, schools would try harder to attract the least expensive to educate.[34]

This concern is supported by research on managed care in the Medicare system in the United States. Under Medicare, health maintenance organizations receive a fixed payment per enrollee (with some modest adjustments for personal characteristics associated with the cost of treatment). Researchers[35] have found evidence that they choose the services they offer in order to attract healthy patients and deter patients with chronic illnesses from enrolling.

[34]For a discussion of the health-care case see Frank,Richard, and Jacob Glazer and Thomas McGuire, "Measuring Adverse Selection in Managed Health Care," *Journal of Health Economics*, 19 (Novemer 2000): 829-854.

[35]Cao, Zhun and Thomas McGuire, "Service-Level Selection in Medicare," *Journal of Health Economics*, forthcoming.

3.9. Evidence from New Zealand. Beginning in 1989 and culminating with additional reforms in 1991, New Zealand moved from a highly centralized education system under the control of the Department of Education to one with a system of locally managed schools competing for students.[36] The reforms can most easily be described as a system of universal vouchers and universal charter schools under the governance of parent boards. Schools were given wide latitude to run themselves and received a fixed payment from the central government for each student enrolled (with a supplement for students from disadvantaged backgrounds).

In contrast with most charter schools in the United States, oversubscribed schools were free to select the students they preferred from among their applicants. By 1995, almost half of the schools in the two largest metropolitan areas were over-subscribed. In the United States, most over-subscribed charter schools must select their students on the basis of a lottery. This difference should not be exaggerated. Charter schools in the United States have the potential to attract the types of students they want by choosing policies that make them attractive to those types of students. They can expel students for minor discipline violations or set high academic standards and flunk those who fail to meet them.

The evidence from New Zealand provides a mixed message. On the one hand, while there are no national exams that can be used to quantify changes in student performance, most teachers and principals believe that the program has improved education,[37] and most boards are pleased with their autonomy. On the other hand, there has been a substantial increase in social segregation across schools. Successful schools are able to select the brightest most motivated students and draw their students from advantaged families. A significant minority of schools would be classified as failing. Their boards are overwhelmed by the administrative responsibilities that come with their autonomy. They end up attracting only those students unable to obtain places in the oversubscribed schools. Their students are overwhelmingly drawn from disadvantaged populations. While, in principle, such schools should be allowed to fail, there is not a large population of education entrepreneurs eager to open schools that target disadvantaged students. Indeed, since most parents want their children educated with more advantaged, brighter and more motivated students, creating a school

[36]This section draws heavily on Fiske, Edward B. and Helen F. Ladd, *When Schools Compete*, Washington, DC: Brookings Institution Press, 2000.

[37]*ibid*, p243-4.

targeted at difficult to educate children is not a natural pathway to commercial success.

The New Zealand experience reinforces the view that competition among middle-class schools is an effective mechanism for generating high-quality education. It does not, however, support the view that charter schools and vouchers on a large scale are likely to address the lack of competition and choice for disadvantaged urban children.

4. Using Tests to Increase School and District Accountability

Both federal and many state laws now require schools and school districts to administer state tests. While, as discussed above, many of these state tests are high stakes for students, some are used primarily to evaluate schools.

The system used in Texas has served as a model for other states and for the federal legislation. The Texas Board of Education is mandated to rate the performance of schools and school districts according to a set of "academic excellence indicators," including TAAS or TAKS results, dropout rates and student attendance rates.[38] State law also prescribes that student performance data be disaggregated by ethnicity and socioeconomic status. The performance rating system holds that school performance is not acceptable if the performance of any subgroup is not acceptable. Based primarily on percentage of students passing each of the state tests, the more than 6,000 schools in Texas have been rated since 1994 as "exemplary," "recognized," "acceptable" or "unacceptable." Schools are eligible for cash awards for high ratings; and if they are rated as low performing twice in a row, they are subject to sanctions from the Texas Education Agency, including possible closure.

Under the "No Child Left Behind" (NCLB) act of 2001,[39] each state must establish standards in reading, math and science and must test children annually in grades 3 - 8 and once in high school to determine whether they are meeting those standards. Currently schools' performances are calculated separately for math and English/language arts and separately for each test by ethnicity and socioeconomic status. Schools are supposed either to meet a standard which increases each year or to make adequate progress towards that standard every year.

[38]Texas Education Agency, *Accountability Manual, 2004*, http://www.tea.state.tx.us/perfreport/account/2004/manual/index.html

[39]NCLB reauthorizes the Elementary and Secondary Education Act and is therefore also known as ESEA.

Schools must publish a "report card" on their performance. A school that fails to meet the standard or improvement requirement on any test for any subgroup fails to make adequate yearly progress. Schools that fail to make adequate yearly progress two years in a row will be provided with support and subject to sanctions. In particular, school districts are required to offer parents the choice of another school in the district if their current school is deemed to be failing to make adequate progress. Ultimately such schools can be closed down or their staffs replaced.

Even the simple act of publishing test results may have important consequences. Realtors use information about school performance when they show houses to clients. If housing prices and rents rise in districts with high scores, then homeowners will put pressure on school authorities to ensure that their schools do well on the tests. As such, publishing test results is a natural extension of the argument of the Tiebout model – if people do not know the quality of the schools in different neighborhoods, how can their choice of where to live respond to school quality.

There are, however, three cautionary tales with respect to the use of tests to ensure accountability. The first is one we have already discussed. Schools will respond to the test. If too much weight is placed on math relative to science, schools will ignore science and focus on math. If the test measures calculation and not calculus, schools may not teach calculus.

Angela Dills[40] has addressed this question indirectly by looking at the relation between housing prices and changes in student performance on the TAAS. She reasons that if parents value the improvements in education measured by the TAAS, then they should bid up housing prices in communities where performance on the TAAS improves. She finds no evidence of such a relation. Thus it appears that, at least in the eyes of parents, the benefits of improved performance on the state exams are offset by reduced performance on other dimensions.

The second concern is that, even along those dimensions that the test is designed to measure, tests measure year-to-year changes very imprecisely. Changes in school performance from year-to-year reflect at least three factors – real changes in the way the school prepares students for the test, random variation in the ability of the students from year-to-year and random variation in the preparation of students or other random factors affecting performance. In this last category,

[40]Dills, Angela K. "Do Parents Value Changes in Test Scores? High Stakes Testing in Texas," Clemson University, unpublished, Sept. 2002,

one school might have taken the test during a bad thunderstorm that distracted many students or the grade might have had a special project on some topic that was addressed one year but not the next.

Thomas Kane and Doug Staiger[41] show that randomness is an important element in gains especially for those schools experiencing the largest gains or largest losses. When states such as North Carolina give rewards to the 25 schools showing the greatest improvement, the overwhelming majority of those awards go to small schools. A school with a single class of twenty students is much more likely to have large year-to-year changes in student quality than a school with 500 students in a grade.

Figure 4 plots the change in average score on the Massachusetts 10th grade math class for schools that administered the exam in all three years 2000 to 2002. The vertical axis shows the change from 2001 to 2002 while the horizontal axis shows the change from 2000 to 2001. The size of the circle is proportional to the size of the 10th grade in 2001.

It is apparent from looking at the figure that the points are arrayed along a downwards sloping line – schools that improved a lot from 2001 to 2002 tend to be those schools that improved little the previous year. Relatively few schools stand out from the mass two years in a row in either direction. Moreover, most of those that do tend to be small.

This does not mean that we cannot learn anything from changes in test scores, just that we need to be careful about putting too much weight on annual changes. Changes that persist over a period of years are likely to be real rather than statistical flukes. However, many short term "successes" will quickly disappear.

In addition, among schools of the same size, more diverse schools will have greater difficulty making adequate yearly progress. A school with 250 students, all of the same race needs to meet the standards only for the group as a whole. A school with 250 students equally divided among blacks, whites, Hispanics, Asians and native-Americans would have to meet the standard for each group. Since each group is small, there is more randomness than for the group as a whole. And, of course, the school risks failing with any one of five groups rather than with just one group.

The third concern about putting considerable emphasis on high-stakes exams is the strong pressure on educators to cheat. There are

[41]Kane, Thomas J. and Douglas O. Staiger, "The Promise and Pitfalls of Using Imprecise School Accountability Measures," *Journal of Economic Perspectives*, 16 (Fall 2002): 91-114.

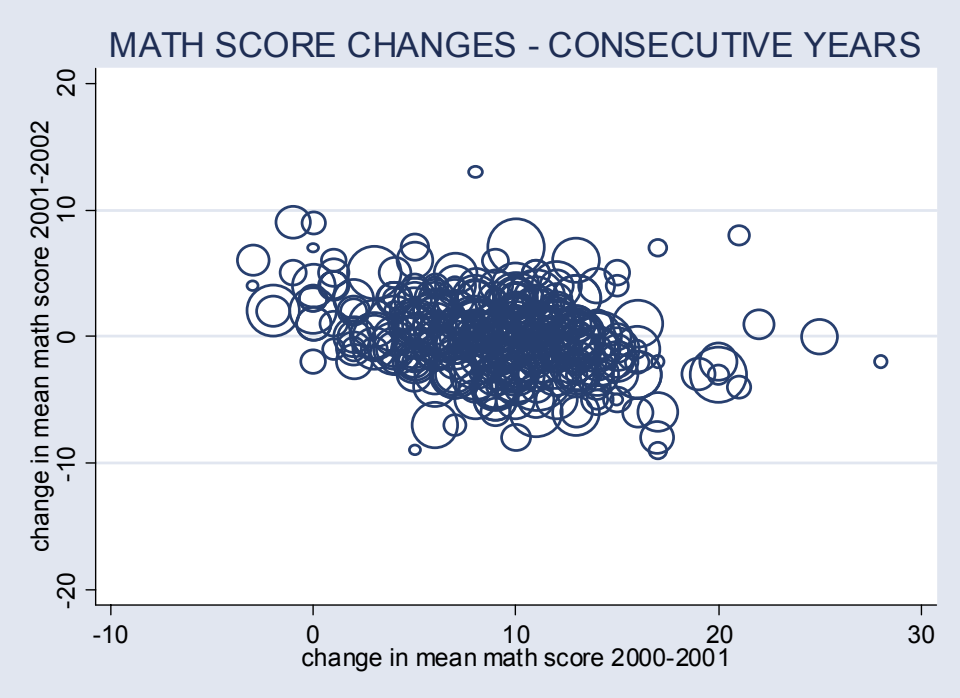

Figure 4:

high-profile newspaper stories of teachers giving students the rights answers or erasing answers and filling in the correct answers, but it is difficult to know how common such actions are.

Steve Levitt and Brian Jacob[42] used a clever technique to uncover improbable patterns in students answers. If a group of students got easy questions wrong and then got many hard answers right, it was unlikely that this occurred naturally. They estimated that such cheating occurred in about 4-5% of Chicago classrooms and that the incidence of cheating rose 75% when sanctions for low-performing schools were introduced. Moreover, the increase in cheating occurred in schools in the bottom third of the performance distribution. Their work turned up evidence that led to the disciplining of some teachers, but in many cases, results were unlikely but there was not sufficient evidence to prove that any individual teacher cheated.

[42]Jacob, Brian A. and Steven D. Levitt, "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating," *Quarterly Journal of Economics.* 118 (August 2003): 843-77.

5. Concluding Remarks

There is good reason to believe that raising the educational attainment and improving the quality of schools attended by poor children would improve their outcomes as adults. Although proving causality is difficult, higher school quality is associated with children completing more years of school. So any improvement in school quality is likely to increase educational attainment as well as to ensure that students learn more while in school.

There is some evidence that competition among schools raises school quality. In many areas the competition among suburban school districts and the importance of the schools for the community put pressure on political and educational leaders to ensure that the schools provide high quality education efficiently.

Many, perhaps most, urban schools are immune from these pressures. Education reform movements try to improve urban schools by substituting other mechanisms for the competition among districts found in suburban areas. There is little evidence that any of these mechanisms has been effective. Texas, often held up as an example of tremendous success, has primarily succeeded in improving performance on its own test with some much more limited spillover to performance on other tests. This has come at some cost in terms of an increased dropout rate, energy expended by administrators to look good on the assessment system (for example by recorded students as transferring rather than dropping out) and by reduced emphasis on aspects of education not covered by the assessment system.

Just as there are good and bad regular public schools, there are also good and bad charter schools. If anything, charter schools appear to be, on average, worse than public schools serving comparable populations, probably because they have less experienced leadership. There is little evidence that increased competition from charter schools has the desired effect on the regular public schools with which they compete. While vouchers undoubtedly help some of the students they serve, the evidence is weak that voucher recipients generally get large gains from the voucher.

Our limited experience with vouchers, charter schools, intra and inter-district school choice does not suggest that revamping our education system to create a system of universal vouchers (or charters) would solve the problem of inner-city education. The experience in New Zealand indicates the opposite.

This does not mean that such reforms have no role to play in improving education. Because of the multitask principal agent problem,

putting too much emphasis on state tests is problematic, but it is possible to use the tests while putting less weight on them. State tests could play the same role for students leaving school after high school that the SAT and ACT play for those applying to college. Charter schools and pilot schools within urban school system could be selected more rigorously both in terms of leadership and innovation and then carefully monitored and evaluated to determine their efficacy. The experience with CSD 4 in New York suggests that inspired leadership and innovation can have a large impact on schools even in high poverty neighborhoods.

6. Further Reading

Eric P. Bettinger, "The Effect of Charter Schools on Charter Students and Public Schools," unpublished paper, Case Western Reserve University, 2002.

Julian R. Betts and Robert M. Costrell, "Incentives and Equity under Standards-Based Reform" in Diane Ravitch, ed., *Brookings Papers on Education Policy: 2001*, Washington, DC: Brookings Institution, 2001.

Robert Bifulco and Helen F. Ladd, "The Impacts of Charter Schools on Student Achievement: Evidence from North Carolina," Paper presented at the Association of Public Policy Analysis and Management Annual Meeting, 2003.

Sandra E. Black, "Do Better Schools Matter? Parental Valuation of Elementary Education," *Quarterly Journal of Economics*. 114 (May 1999): 577-99.

Katharine L. Bradbury and Christopher J. Mayer and Karl E. Case, "Property Tax Limits, Local Fiscal Behavior, and Property Values: Evidence from Massachusetts under Proposition 2 1/2," *Journal of Public Economics*, 80 (May 2001): 287-311.

Richard Buddin and Ron Zimmer, "Academic Outcomes," in Zimmer, Ron et al eds., *Charter School Operations and Performance: Evidence from California,* Santa Monica, CA: Rand Corporation, 2003.

Peter W. Cookson, Jr. and Sonali M. Shroff, *School Choice and Urban School Reform*, Teachers College, Columbia University, December 1997, http://eric-web.tc.columbia.edu/monographs/uds110/

Angela K. Dills, "Do Parents Value Changes in Test Scores? High Stakes Testing in Texas," Clemson University, unpublished, Sept. 2002

Edward B. Fiske and Helen F. Ladd, *When Schools Compete*, Washington, DC: Brookings Institution Press, 2000.

David W. Grissmer et al, *Improving Student Achievement: What State NAEP Test Scores Tell Us*, Santa Monica, CA: Rand Corporation, 2000.

Eric A. Hanushek, *Making Schools Work: Improving Performance and Controlling Costs*, Washington, DC: Brookings Institution Press, 1994.

Caroline M. Hoxby, "Does Competition among Public Schools Benefit Students and Taxpayers?" *American Economic Review*,.90 (December 2000): 1209-38.

Eric A. Hanushek, John F. Kain, and Steven G. Rivkin, "The Impact of Charter Schools on Academic Achievement," unpublished, Hoover Institution, 2002.

William E. Hoyt, "Public School Choice, Property Values, and Investment in Public Education", University of Kentucky, unpublished, 1997.

Brian A. Jacob and Steven D. Levitt, "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating," *Quarterly Journal of Economics*. 118 (August 2003): 843-77.

Thomas J. Kane and Douglas O. Staiger, "The Promise and Pitfalls of Using Imprecise School Accountability Measures," *Journal of Economic Perspectives*, 16 (Fall 2002): 91-114.

Stephen P. Klein et al, "What Do Test Scores in Texas Tell Us?" *Rand Organization Issue Paper No. 202*, 2000.

Alan B. Krueger and Pei Zhu, "Another Look at the New York City School Voucher Experiment," *NBER Working Paper No. 9418*, 2002.

Paul E. Peterson and William G. Howell, "Efficiency, Bias, and Classification Schemes: Estimating Private-School Impacts on Test Scores in the New York City Voucher Experiment Latest Results from New York City School Voucher Research – African American in Private Schools Score Higher, *Program on Education Policy and Governance Working Paper PEPG/02-15 - 22,* 2002

Randall Reback, "Capitalization under School Choice Programs: Are the Winners Really the Losers?" National Center for the Study of Privatization in Education Teachers College, Columbia University, Occasional Paper No. 66, 2002.

Cecilia E. Rouse, "Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program," *Quarterly Journal of Economics*. 113 (May 1998): 553-602.

7. Questions for Discussion

1. What are the principal reasons that education may have a causal effect on earnings?

2. If education is primarily a sorting device and if high-stakes testing raises standards, who is helped and who is hurt by the tests.

3. If students learn marketable skills in school, will raising standards increase or decrease the dropout rate?

4. What is the multitask principal agent problem? How is it relevant to the testing debate?

5. Explain how competition among school districts generates efficient provision of education in the Tiebout model. What are the weaknesses of the model for representing the real world?

6. Why are the theoretical advantages and disadvantages of decentralized funding for education? Does empirical research support these predictions?

7. Discuss the theoretical and empirical effects of intradistrict and interdistrict school choice on school quality?

8. Do voucher programs improve educational outcomes for participants?

9. What is the evidence regarding the effects of charter schools on student performance?

10. What is meant by each of the following: intradistrict school choice, interdistrict school choice, vouchers, charter schools?

11. What does the experience of CSD 4 in New York City tell us about the effectiveness of intradistrict school choice programs?