Volume Title: The Economics of School Choice

Volume Author/Editor: Caroline M. Hoxby, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-35533-0

Volume URL: http://www.nber.org/books/hox03-1

Conference Date: February 22-24, 2001

Publication Date: January 2003

Title: School Choice and School Productivity. Could School Choice Be a Tide that Lifts All Boats?

Author: Caroline Minter Hoxby

URL: http://www.nber.org/chapters/c10091

8

School Choice and School Productivity
Could School Choice Be a Tide that Lifts All Boats?

Caroline M. Hoxby

A productive school produces high achievement in its pupils for each dollar it spends. Formally, a school's productivity is defined as achievement per dollar spent, controlling for incoming achievement differences of its students. In this chapter, I comprehensively review how school choice might affect productivity. I begin by describing the importance of school productivity, then explain the economic logic that suggests that choice will affect productivity, and finish by presenting much of the available evidence on school choice and school productivity. Readers are likely to be most intrigued by the final section of the paper, in which I examine the achievement and productivity effects of three important recent choice reforms: vouchers in Milwaukee, charter schools in Michigan, and charter schools in Arizona. However, readers are much less likely to find the evidence to be a "black box" if they read the earlier sections of the paper, which set up the relationship between choice, school conduct, student achievement, and productivity. I encourage impatient readers who jump to the final section to return to the earlier sections for answers to the questions that will naturally arise once they have seen the evidence.

8.1 Why the Productivity Consequences of School Choice Matter A Lot

Although a great deal of research has dealt indirectly with school productivity (most famously, the "does money matter?" debate), productivity has been neglected by research on school choice. School choice research has concentrated on *allocation* questions, which include the following: Who ex-

Caroline M. Hoxby is professor of economics at Harvard University and a research associate of the National Bureau of Economic Research.

ercises school choice? Who chooses which school? How does choice change the allocation of resources? How does reallocation of students change peer effects? The allocation questions are largely questions of redistribution. Although it is theoretically possible that school choice could improve achievement for *all* students through reallocation, such an outcome would require that, for *every* student, the benefits of going to a school that was a better match exceed the costs imposed upon him or her by school choice. The costs might include a worse peer group or a decline in resources.[1]

In general, then, allocation-oriented research presents a view of school choice that is rife with tensions about redistribution (which students gain, and which students lose?). One way to relieve these tensions is to devise allocation-related remedies, such as controls on how resources and peers shift when choice is introduced (see, e.g. Hoxby 2001). However, when advocates of school choice argue that every child would benefit from school choice, they are usually relying on the idea that school *productivity* would increase sufficiently to swamp any negative allocation effects that some students might experience. The basic logic is that choice would give schools greater incentives to be productive because less productive schools would lose students to more productive schools. That is, if a school could raise a student's achievement while spending the same amount as the current school, it would be expected to draw the student away from his or her current school. This process would shrink the less productive and expand the more productive school, until one of two things happened: the more productive replaced the less productive school or the less productive school raised its productivity and was thereby able to maintain its population of students. (This is the broad idea: later I discuss specific mechanisms through which choice might raise productivity.) In other words, a general increase in school productivity could be a rising tide that lifted all boats, and the gains and losses from reallocation might be nothing more than crests and valleys on the surface of the much higher water level.

Thus, the first reason that the productivity consequences of school choice matter is that they potentially determine whether choice will benefit all children. For the rising tide scenario to be a realistic probability and not just a possibility, however, one must ask what productivity schools could reasonably be expected to achieve. That is, what is the range of productivity over which choice *could* cause productivity to vary? Recent history suggests that school productivity could be much higher than it is now—60 to 70 percent

1. Strictly speaking, what is required is that (a) the current distribution of peers or teaching methods is inoptimal and (b) that school choice would cause people to redistribute themselves in such a way that a Pareto improvement in the distribution of peers or teaching methods would occur. It is reasonable to think that families might redistribute themselves so as to achieve better alignment between teaching methods and their children's learning styles. It is less reasonable to think that families, acting independently, could implement a Pareto improvement in peer effects.

higher. Consider the simplest productivity calculation, achievement per dollar. Such a calculation (which I later describe in detail) suggests that average public school productivity was about 65 percent higher in 1970–71 than in 1998–99. This means that, if choice were simply to restore school productivity to its 1970–71 level, then the *average* student in the United States would be scoring at an advanced level where fewer than 10 percent of students now score. This improvement in achievement would be so large that it would overwhelm any worst-case scenario suggested by allocation research on school choice.

8.1.1 How Much Higher Could School Productivity Plausibly Be?

How does one make such a calculation? We have one measure of student achievement in the United States that reflects the achievement of the entire population of students, is nationally representative, and is designed for comparison over a long period of time and across schools: the National Assessment of Educational Progress (NAEP). Other measures of achievement tend to fail at least one, usually a few, of these requirements.[2] If one simply calculates NAEP points per thousand real dollars spent per pupil, one generates the results shown in table 8.1.[3] (All money amounts in this chapter are adjusted into 1999 dollars using the Consumer Price Index, unless otherwise indicated.) They show that, between the 1970–71 and 1998–99 school years, productivity fell by between 54.9 percent (based on math tests for nine-year-olds) and 73.4 percent (based on reading tests for seventeen-year-olds). The bottom section of table 8.1 shows actual NAEP scores in its upper row and, in its lower row, what NAEP scores would be if schools returned to 1970–71 productivity (1972–73 productivity, in the case of math). For all of the tests, the average American student would have a score that fewer than 10 percent of American students currently attain. In fact, the *average* seventeen-year-old would have a score that fewer than 5 percent of American seventeen-year-olds currently attain. The mean American student would be classified by the NAEP as an advanced student.

One might wonder whether demographic changes in the United States account for the fall in school productivity, as measured by the simple calculation described above. Perhaps schools were not losing productivity; perhaps they were simply working with students from worse family backgrounds. There is no definitive way to address this issue, but a standard approach is to do the following:

2. The high school dropout rate, for instance, only reflects variation in the outcomes of low-achieving students. Students' self-selecting into the tests generates incurable biases when the Scholastic Aptitude Test (SAT) or ACT (formerly known as the American College Testing Program) is used for comparisons over time or across schools. The SAT and ACT also only reflect variation in the outcomes of high-achieving students. High school grades have been shown to be relative measures that cannot be compared successfully across schools.

3. The source for the table is U.S. Department of Education (2000).

Table 8.1 School Productivity in the United States, 1971–99

	Productivity[a] Based on the Test of:					
	Reading (9-year-olds)	Math (9-year-olds)	Reading (13-year-olds)	Math (13-year-olds)	Reading (17-year-olds)	Math (17-year-olds)
School year						
1970–71	46.0		56.6		63.3	
1972–73		45.5		55.3		63.2
1974–75	41.7		50.8		56.7	
1977–78		41.3		49.7		56.5
1979–80	41.3		49.6		54.8	
1981–82		40.8		50.1		55.7
1983–84	36.4		44.4		49.9	
1985–86		34.8		42.2		47.4
1987–88	30.5		37.1		41.8	
1989–90	28.9	31.8	35.5	37.3	40.1	42.2
1991–92	29.0	31.7	35.8	37.7	40.0	42.3
1993–94	28.7	31.4	35.1	37.2	39.2	41.6
1995–96	28.1	30.6	34.3	36.3	38.0	40.6
1998–99	26.8	29.4	32.8	35.0	36.5	39.0
Productivity decrease (%) since earliest year shown						
No adjustments	71.5	54.9	72.5	58.1	73.4	62.0
Adjusted for demographics	74.3	58.2	75.6	62.3	78.0	65.1
Adjusted for wages of females with advanced degrees	55.9	39.1	56.8	42.1	57.6	45.5
Actual mean NAEP score (1998–99)	212	232	259	276	288	308
Predicted NAEP score at productivity of earliest year shown (1998–99)	363	359	447	437	500	499

Source: U.S. Department of Education *Digest of Education Statistics 1999* (2000) and U.S. Department of Education *National Assessment of Educational Progress 1999 Long-Term Trend Summary Data Tables* (2000).

[a]NAEP points per thousand dollars of per-pupil spending.

- regress 1998–99 achievement on the characteristics of students who took the test in that year and thereby determine the effect of each characteristic (African American, Hispanic, single-parent family, family income, and so on);
- predict what achievement would have been in 1998–99 if the student population were the same as the 1970–71 student population—that is, substitute 1970–71 characteristics into the prediction equation with 1998–99 coefficients; and
- use predicted achievement to determine what 1998–99 productivity would have been if the student population had remained what it was in 1970–71.[4]

If one uses this method to hold student characteristics constant, then one finds that the decline in productivity from 1970 to 1999 is very slightly *larger* than the unadjusted estimates would suggest. See the row of table 8.1 in which productivity decrease is adjusted for demographics. For instance, consider the measured decrease in productivity based on the mathematics scores of seventeen-year-olds. It is a 62.0 percent decrease if student characteristics are not held constant, but it is a 65.1 percent decrease if student characteristics are held constant.[5] The decline in productivity is greater when one holds student characteristics constant mainly because a smaller share of students had high school graduate or college graduate parents in 1971–72 than in 1998–99. Such students tend to score better on the NAEP exam than students whose parents are high school dropouts. In addition, there were smaller shares of students in 1971–72 whose families had the incomes typical of families today. The shares of students who are African American and Hispanic have risen since 1971–72, and these students do tend to score worse on NAEP exam than non-Hispanic white students. However, the effect of changing racial composition is overwhelmed by the effect of changes in parents' education and income. Other changes in the composition of the student population, such as area of the country, have little effect on the adjustment.

If demographic changes do not account for the fall in school productivity, perhaps changes in career opportunities for women do. That is, over the

4. The calculation is

$$\rho = \frac{\mathbf{X}_{1972\text{–}72}\,\hat{\boldsymbol{\alpha}}_{1998\text{–}99}}{pps_{1998\text{–}99}},$$

where ρ is productivity, $\mathbf{X}_{1973}$ is the vector of characteristics of the 1971–72 student population, pps_{1998} is per-pupil spending in 1998–99, and $\hat{\boldsymbol{\alpha}}$ is the vector of estimated coefficients from the regression

$$NAEP_{1998\text{–}99} = \mathbf{X}_{1998\text{–}99}\,\boldsymbol{\alpha}_{1998\text{–}99} + \varepsilon.$$

$NAEP_{1998}$ is a 1998–99 NAEP score (in reading, math, or science) and $\mathbf{X}_{1998}$ is the vector of characteristics of the 1998–99 student population.

5. The source of data for the calculations described in U.S. Department of Education (1999).

1970–99 period, it may have cost schools an increasing amount to hire a female with a given level of skills because non-teaching opportunities for women were opening up. One can examine this hypothesis by inflating nominal spending using a wage index for females rather the CPI. In order to give this hypothesis as much explanatory power as possible, I used the wage index for females in the college-educated occupation that experienced the most wage growth: professional specialty occupations (lawyers, physicians, etc.).[6] Use of this index exaggerates the degree to which females' wages account for the measured decline in productivity for two reasons. First, women in professional specialty occupations have always had higher-quality educations and higher ability than the average American school teacher, and highly skilled and able workers have experienced rapid earnings growth relative to all other workers (including less skilled college graduates) since 1970. Second, teachers are not the only input that schools require. They also need office equipment, buildings, less skilled service workers (custodians, bus drivers, food preparers), and other inputs; the prices of such inputs have not risen nearly as fast as the wages of female professional specialty workers. As long as we recognize that inflating by female professional specialty workers' wages is likely to give us a smaller decrease in productivity than has really occurred, the calculation is informative. The row in table 8.1 in which productivity increase is adjusted to the wages of females with advanced degrees shows that the wage-adjusted decrease in productivity ranges from 39.1 to 57.6 percent, whereas the CPI-based decrease in productivity ranges from 54.9 to 73.4 percent. In other words, although the wage-adjusted productivity losses *are* smaller, they are still very substantial.

The facts suggest that school conduct, not changing student characteristics or female career opportunities, is the main source of the decline in productivity. Consequently, policies that improve school conduct could potentially generate very large increases in productivity.[7] Of course, it is not

6. Earnings of full-time, full-year females working in professional specialty occupations are taken from U.S. Department of Commerce (1976, 1983, 1995, 1999). The index is for women in a constant age range. However, experience is not held constant. Because women have been gaining more experience for every year they age (they have spent less time out of the labor force for family reasons), using an index that holds experience (not age) constant would produce results that look more like the results using the CPI. This is yet another reason why the productivity loss shown in the row of table 8.1 that is adjusted for wages of females with advanced degrees is understated.

7. One could criticize the constant-student productivity by saying that some student characteristics mean different things in 1998 from in 1971. For instance, coming from a single-parent family is more common in 1998 than in 1971, and thus it may be a different experience now from what it was in 1971. On the whole, however, such criticisms work in the wrong direction. Most of the student characteristics that are more common now than in the past are also less stigmatized—being a member of minority or being from a single-parent family, from the South or Southwest, and so on. The decline in productivity would be larger if one were to take account of the fact that having a single-parent household, say, is not as bad for achievement as it was in 1971.

enough to point out that school productivity could plausibly be much higher than it is. One must investigate whether choice actually induces schools to raise productivity. Such investigations—both how one conducts them and what they show—are the main content of this paper. Before I take up such matters, however, one more vital point about school productivity must be made.

8.1.2 How School Productivity Affects American Industry and Growth

For as long as we have been able measure the factor content of American net exports and the sources of American economic growth, they have been intensive in human capital. This was observed early on by Leontief (1956) and confirmed by a series of other researchers (Keesing 1966; Krueger 1968; Jorgenson 1984; and Jorgenson and Fraumeni 1989, 1992). In other words, the United States has a comparative advantage in producing goods and services that make intensive use of educated labor. This comparative advantage has existed because America has always had a relative abundance of educated labor. That is, the United States has always been able to produce education in its population relatively cheaply. America's "new economy" products (microprocessors, software, knowledge services) are some of the most human capital–intensive products in the world. However, we know from basic trade theory that the human capital–intensive economy is built on a foundation of American ability to produce education in its population relatively cheaply. Although it true that America can import some human capital (for instance, software engineers), imported human capital *cannot* be a source of comparative advantage in the middle to long run.[8] Thus, if Americans wish to continue enjoying a growing economy that is centered around human capital–intensive products, they cannot be indifferent about rapidly falling productivity in their schools. A school sector with falling productivity translates into America's having relatively costly human capital, which translates into a loss of comparative advantage in human capital–intensive goods.

In short, the effect of choice on school productivity is not interesting simply because it could overwhelm the allocation effects of choice on achievement; it is interesting because it also has broad implications for the macroeconomy, for trade, and for Americans' jobs.

8.2 How Productivity Fits into the School Choice Literature

The productivity implications of choice have been sadly neglected by the literature on school choice. This neglect has nothing to do with the impor-

8. Other countries can import human capital too, so imported human capital cannot be cheap relative to its cost in the rest of the world. Moreover, under a logical extension of current trends, countries that are currently net exporters of human capital would become the world's net exporters of human capital–intensive products.

tance of productivity (which is great, as has been discussed) and has everything to do with the roots of the theoretical literature. Models of school choice have grown out of models of local public goods provision, which have traditionally focused exclusively on allocation problems, such as who gets what local public good and how one person's local public good choice affects other people. This focus has been inherited by the school choice literature, and although allocation-focused models of choice are instructive, the intellectual history of the literature should not dictate neglect of productivity. Indeed, it is worth while to take a step back to look at some related research that demonstrates how important productivity effects can be when competition is introduced into a market.

Health care is an obvious and recent example. Legislation passed in the late 1980s and early 1990s allowed managed care organizations to compete.[9] The competition has affected the allocation of health care, but a wealth of research also documents the dramatic effects of competition on the productivity, which far exceeded what supporters of managed care had hoped. From 1990 to 2000, health care costs grew just one-half as quickly as in the previous decade, but key health indicators (such as life span) grew just as rapidly in the 1990s as in the 1980s. These facts suggest that productivity surged in the more competitive environment,[10] in part because competition induced providers to adopt efficiency-enhancing technology (such as computers that reduce paperwork) and to discourage conduct that created rents (such as doctors' referring patients to their friends without regard to cost). It should be noted that market competition was associated with productivity gains in nonprofit and public hospitals, as well as private hospitals.

Trucking and parcel services are other examples. Many researchers have documented how, following deregulation in the 1970s, competition in trucking grew dramatically. The result was a sharp increase in productivity, as documented by Rose (1987), Michel and Shaked (1987), Traynor and McCarthy (1991), and others. For the same money, a trucking customer could obtain faster, more specialized service after competition than before. In parcel services, the introduction of competition improved productivity not only because the private firms (United Parcel Services, Federal Express, DHL Worldwide Express, etc.) had higher productivity and productivity growth than the U.S. Postal Service did. The competition also induced the U.S. Postal Service to raise substantially its own productivity. Many commentators had doubted whether the U.S. Postal Service could rise to the occasion and compete, but it has maintained a large market share in parcel de-

9. Managed care organizations include health maintenance organizations, primary provider networks, and certain other physician and hospital networks.

10. See *Economic Report of the President* (2000). The literature on the productivity effects of managed care is voluminous, but Cutler and Sheiner (1998) may serve as a good introduction to it.

livery—largely by introducing services (like express mail and priority parcel post) that are competitive on price and quality with services offered by the private firms.

In fact, it is somewhat odd that school productivity should be so neglected in the school choice literature, because—although productivity was also neglected in the local public goods literature—there is increasing interest among economists in the productivity of not-for-profit, semipublic, and regulated enterprises. Economists are increasingly interested in giving market-like incentives to such enterprises in order to keep workers from rent-seeking despite the fact that they lack conventional profit-maximizing incentives. For instance, yardstick competition among not-for-profit providers of social services (awarding contracts to training programs on the basis of their performance relative to other sites) is increasingly used as a method of inducing productivity gains.

Finally, it is worth noting that one type of school-related research does implicitly contain substantial evidence on productivity, although productivity is rarely mentioned and productivity calculations are never made. I refer to research that compares students' outcomes in public and private schools and that attempts to eliminate selection bias. (Selection bias is the potential bias caused by the fact that students who self-select into private schools might be unobservably different from students who remain in public schools.) The body of research on this topic is well established and even includes some recent research comparing students who are assigned *by lottery* a private school voucher or no voucher (so that they remain in the public schools). Peterson et al. (chap. 4 in this volume) illustrate the best strain of such research. The consensus in public versus private achievement research appears to be that private schools produce statistically significantly better achievement, at least among minority children and children from lower- to lower-middle-income households.

This body of research could be reformulated as a comparison of public and private school productivity, because there is always an attempt to compare achievement and hold constant the quality of student inputs. Unfortunately, other inputs, especially spending, are *not* constant between private and public schools, and this body of research is often silent about this fact (and almost never controls for it). In particular, the typical private school in the United States spends only about 60 percent as much per pupil as the typical public school, but private school spending is also much more variable than is public school spending, so that minimum private school spending is lower than minimum public school spending and maximum private school spending is higher than maximum public school spending. Thus, even if researchers were to find that public and private schools produced identical achievement, it would probably be true that private schools were considerably more productive (because they spend less on average). However, if one is to get an accurate comparison of public and private school productivity, one really ought to make a productivity calculation for each school (thereby taking account of differences in the distribution of spending) and compare these calculations for students with the same backgrounds.

8.3 Why Should Choice Affect School Productivity?

Why, logically, should choice be expected to affect productivity? That is, what mechanism guarantees that low-productivity producers will be driven out by high-productivity producers? Ultimately, this is a question about (a) what schooling producers maximize and (b) what the production function for schooling is like. In fact, there are several answers to this question, and the answer is different for different types of schooling producers: for-profit firms (like Edison Schools), not-for-profit private schools, charter schools, and regular public schools. In this section, I describe the mechanism by which choice might affect productivity for each of these types of schools. Across all the cases, I do maintain one assumption: For any given cost to them, parents will choose the school that produces the schooling that they value most. For convenience, I will hereafter call what parents value in schools "school quality," but I do not assert that parents' notion of school quality necessarily matches that of the reader. In other work (Hoxby 1999a) I have presented empirical evidence that suggests that parents tend to prefer schools that have better academic achievement, emphasize academic standards, and promote a relatively structured (disciplined) school atmosphere.

8.3.1 A For-Profit School Producer That Takes Up Charter School Contracts

Let us start with a very straightforward case: a for-profit firm that opens a charter school. Such a case is fairly typical of Edison Schools and might become a common model if charter school programs were more widely enacted. The fee that the school could charge would be set by law, and parents would not be able to "top up" the fee. Also assume that the school must accept charter school applicants at random (a typical charter school restriction) and is risk neutral. In other words, a plan to include or exclude students cannot be part of the school's profit maximization strategy.

Then, the school would solve a problem such as

$$\max_{q,l,k} \pi = px(q) - c(q, l, w, k, r)x(q). \tag{1}$$

This problem simply says that the school maximizes the difference between revenues (the fixed fee p times the number of students who enroll x) and costs (per-pupil costs c times the number of students who enroll x). The school chooses the quality q that it offers, the staff or labor l that it hires, and the other inputs k that it employs (textbooks, equipment, and so on). The school accepts the going wage rate for staff w and the going price for other school inputs r. Per-pupil costs c are assumed to be increasing in quality, staff hired, and other inputs purchased. I have assumed that per-pupil costs are the same regardless of the school's scale. This is a good starting assumption, but it is probably not true. I relax it below.

Given that we have said that parents choose the school that offers the highest quality for a given price, it is clear that enrollment x is increasing in quality q. Specifically, the school enrolls all of the public school students in an area if it offers quality that is higher than that of any other area charter school or regular public school (which would be equally free to parents). That is,

$$x(q_{j^*}) = \sum_{t=1}^{N} 1 \text{ if } q_{j^*} > q_{j \neq j^*} \text{ for all } j, \tag{2}$$
$$x(q_{j^*}) = 0 \text{ if } q_{j^*} < q_{j \neq j^*} \text{ for any } j.$$

Assume that the school shares equally in enrollment if it offers exactly the same quality as another school.

In these circumstances, the best that the school can do is maximize quality subject to the constraint that its per-pupil cost must not exceed the charter school fee. Put another way, the school *must* maximize its productivity for a given cost, or another school will enroll all of the students in the area. Unproductive schools will be driven from the market. Note that the firm earns just enough profit on each student to repay its shareholders a market rate of return for the use of their capital, so the best it can do is maximize the number of students on whom it earns this slim profit. It does this by offering the highest possible quality that the charter school fee can sustain.

Managers of for-profit schooling firms believe that there are economies of scale in schooling because a firm can pay lower prices for its inputs if it pools purchasing, curricular research and development, and information processing across multiple schools.[11] If there are economies of scale, then large firms may be able to earn economic profits (profits that exceed the profits necessary to pay the cost of capital) in local markets where they compete with other schools that, for one reason or another, remain too small to take advantage of economies of scale.

8.3.2 A For-Profit School That Takes Up Vouchers

The case of a for-profit school producer that takes voucher students is quite similar to the case just examined, except that parents are assumed to be allowed to top up a voucher with extra tuition payments from their own funds. Otherwise, assume that the case is the same: The school must accept voucher applicants on a random basis conditional upon the applicants'

11. Pooling may occur even if the schools are in different areas and offer somewhat different levels of quality.

being willing to pay the school's fees with a combination of the voucher and extra tuition payments. Because the school can now set its fees, its problem is slightly more complicated:

$$\max_{p,q,l,k}\pi = px(p, q) - c(x, q, l, w, k, r)x(p, q). \tag{3}$$

That is, when a school sets its fees, it takes into account that a higher fee means, on the one hand, greater revenue per student who enrolls but, on the other hand, lower enrollment (because a higher fee discourages enrollment for any given level of quality offered by the school). It can easily be seen that, for any given fee p, the school must maximize the quality it produces subject to the constraint that costs are less than or equal to p. In other words, the school must still maximize productivity if it is not to lose all its enrollment to another school that offers higher quality for the same price.

Note that, in the equation above, I have allowed for economies of scale. Thus, when a school considers aiming for a "better" but smaller niche of parents, who are willing to be charged higher fees for better quality, it must take into account the loss of economies of scale (and the consequent increase in its costs).

8.3.3 A Nonprofit School That Takes Up Charter School Contracts

The for-profit case is a nice place to begin because the firm's incentives to maximize productivity are obvious. The vast majority of school producers that take up charter school contracts or voucher students are, however, not-for-profits. At first glance, it might seem difficult to say what not-for-profits maximize, but in fact relatively simple modifications of the for-profit case capture not-for-profit behavior. The key difference between a nonprofit and a for-profit organization is the distribution of surplus.[12] A for-profit school distributes profits to its owners (private owners or shareholders). Thus, in the problem above, it was reasonable to assume straightforward profit maximization because owners benefit directly from profits. If a not-for-profit school has surplus (a difference between revenues and costs), it cannot pay them in a straightforward way to anyone. It can, however, use surplus in a variety of ways that make surplus a valuable thing to have. Surplus can be used to make working conditions pleasant for the school's staff (staff lunches, smaller classes, more classroom supplies, and so on) even if these conditions do not contribute to productivity. Surplus also allows a school to pursue social goals that its staff value: experiments with teaching methods, development of new curricula, a diverse student body, the exposing of students to nature. There are a few things to note about such distributions of surplus. First, they are nearly always inefficient compared to distribution

12. This point has been discussed by numerous researchers. See Glaeser and Shleifer (forthcoming) for a recent model of nonprofit entrepreneurs and for a review of the literature.

of cash (which is fungible). That is, some of the surplus is lost in the process of being transformed into goods or services that the staff values. As a result, the school staff faces weaker incentives than they would face if they could be given cash incentives. Second, although it is relatively simple to distribute a nonprofit school's surplus to its staff in the forms mentioned, it is difficult to distribute it (legally) to a single owner or even a concentrated subset of the staff. Thus, a school has less incentive to expand simply to increase the absolute size of the surplus: The surplus will increase as it expands, but so will the number of staff over whom the surplus must be divided. This is unlike the for-profit situation where owners have an incentive to expand their schooling production so long as they can earn some positive surplus on each additional enrollee.

One can incorporate these features of the distribution of surplus into a nonprofit charter school's maximization problem:

$$\text{(4)} \qquad \max_{q,l,k} \frac{(\alpha \cdot \pi)}{l} = \frac{\alpha \cdot [px(q) - c(q,l,w,k,r)x(q)]}{l}$$

where enrollment is given by

$$\text{(5)} \qquad x(q_{j^*}) = \sum_{i=1}^{N} 1 \text{ if } q_{j^*} \text{ for all } j,$$

$$x(q_{j^*}) = 0 \text{ if } q_{j^*} < q_{j \neq j^*} \text{ for any } j,$$

just as before.

This problem simply says that a staff member at a nonprofit charter school wants the school to maximize $(\alpha\pi)/l$, where π is total surplus (what the for-profit school would call profit), α is a factor that is less than 1 (the share of surplus that remains after it has been transformed into goods for the staff), and l is the number of staff. Under this maximization problem, the school's incentives to expand enrollment are weaker (than those of a for-profit school), but its incentives to maximize productivity are strong. The school will still be driven out by competitors if it does not produce the maximum quality q attainable given the constraint that its costs must not exceed the fixed charter school fee.

Two comments about the nonprofit school's maximization problem are in order. First, if there are economies of scale, the school will have stronger incentives to expand enrollment than suggested just above. Second, one's measure of the productivity of a nonprofit school may slightly understate its true productivity if the school earns surplus and buys staff rewards with it that appear to be inputs (although they really make no contribution to outcomes that parents value). The understatement will be slight because competition among nonprofit schools will drive the surplus toward zero (even as each seeks to maximize its surplus).

8.3.4 A Nonprofit School That Takes Up Vouchers

The case of a nonprofit school that takes voucher students is just like the case of a for-profit school that takes voucher students except that its surplus can be distributed only in the indirect way described above. That is, the voucher school's maximization problem is

$$(6) \qquad \max_{p,q,l,k} \frac{(\alpha \cdot \pi)}{l} = \frac{\alpha \cdot [px(q) - c(x, q, l, w, k, r)x(p, q)]}{l}$$

The nonprofit voucher school must maximize productivity if it is not to lose its enrollment to a similar school that offers higher quality for the same fee p. The only complication is that the school needs to choose its fee and quality simultaneously, and the only caveat is that the school has weaker incentives to expand enrollment than a for-profit voucher school.

8.3.5 A Summary for Fee-Based Schools (For-Profit and Nonprofit)

In all of the cases above, the school's revenues are derived from student fees. It is this fee basis that is crucial because it means that parents' choices determine whether a school is viable or not. If a school's students are enticed away by a competing school that charges the same fees, the school naturally has to increase its productivity (either by raising its quality for the given fee or lowering the fee it charges for its quality). As shown, the for-profit or nonprofit basis of the school is somewhat less crucial. A for-profit schooling firm will have stronger incentives to enter new markets and gain new enrollment, but both nonprofits and for-profits have incentives to maximize productivity.

People often wonder whether there will be an elastic supply of charter or voucher schools. This is an important question, especially for nonprofits, which do not have clear incentives to expand when they hope to earn only a slim surplus on additional students. If there are economies of scale, then the charter school fee or voucher that makes a school viable with a small number of students should guarantee that it is more than viable with higher enrollment. Thus, economies of scale suggest that both for-profits and nonprofits should have elastic supply once they are in business. On the other hand, there are some factors that might function like *diseconomies* of scale. For instance, a charismatic principal might become uninspiring if he or she managed a large school and therefore had little direct contact with students.

Buildings are often discussed as a possible factor that would limit the elasticity of supply of charter or voucher schools. This, however, would seem to be a short-run phenomenon that mainly plagues the start-up of new charter or voucher programs. The total number of students to be taught does not increase simply because a new school has entered, so the introduction of charter or voucher school competition does not require much of a net increase in school building. As enrollment shifts from less productive to more productive schools, buildings should be sold by the shrinking or exiting schools and purchased by the expanding or entering schools. In fact, there *is* an active market for school buildings and similar institutional buildings. If small fractions of school buildings could be sold easily, competition would require *no* net increase in school buildings. Schools are, however, somewhat indivisible: Although parts of school buildings are often sold or leased to separate schools, only certain fractions of a building will generally make a viable school. (For instance, most schools require an entrance area, a set of bathrooms, and so on.) Realistically, then, competition requires a small increase in the total stock of school buildings, simply to allow more flexibility as parents' ability to choose makes enrollment more variable.

In any case, it is clear that some factors (economies of scale) suggest that school supply will be very elastic, whereas other factors (which function like diseconomies of scale) suggest that school supply will be less elastic. The elastic response of charter and voucher schools is, thus, an empirical question and will depend on features of the reform, such as funds for the refurbishment of buildings.

8.3.6 Competition and the Productivity of Regular Public Schools

Does competition give *regular* public schools incentives to be productive? We have seen that fee-based schools face straightforward incentives, but what about regular public schools that are funded mainly by taxes?

If a regular public school faces competition from a charter or voucher school, and the charter school fee or voucher comes directly from its budget, then the regular public school *is* fee-based at the margin and will have marginal incentives to be productive. Whether these marginal incentives work well or not depends on the size of the fee or voucher. Some vouchers or fees are so small relative to regular per-pupil spending that they give public schools perverse incentives to drive students away. That is, a voucher or charter school fee that is small relative to per-pupil spending (or that is not financed from the public school's revenues) *raises* per-pupil spending nonnegligibly for each student who is driven away from the public schools. Public school staff may be able to enjoy greater surplus if they drive students away than if they try to attract them. Such perverse scenarios can be easily avoided by setting a sufficiently high voucher or charter school fee.

What if, however, a regular public school does not face competition from a charter or voucher school? If it is not fee-based at the margin, does it have any incentives to be productive? The answer is yes if the public school is financed by local property taxes and faces a high degree of traditional choice among public school districts.

Traditional choice among public school districts is what occurs when parents choose a school district by choosing a residence. This traditional form of choice is by far the most pervasive and important form of choice in American elementary and secondary schooling today. In order that this form of choice give schools incentives to be productive, it is essential that parents choose among *districts* that are fiscally independent. The mechanism that I am about to describe does not work with *intra*district choice.

Conventionally, public school districts in the United States have revenues that depend largely on local property taxes. If parents in a metropolitan area can choose among a large number of districts, they will tend to favor districts that produce higher achievement for a given local property tax liability or, equivalently, have lower local tax liability for a given level of achievement. That is, parents will tend to favor districts with high productivity. If a school district's productivity falls, it will be avoided by parents who happen to be moving. The resulting decrease in the demand for its houses will drive down the district's property prices. The falling property tax base will, in turn, drive down the school's budget, which depends on property tax revenues. The administrator will be encouraged to raise productivity, either by maintaining achievement in the face of a falling budget or by raising achievement sufficiently to make the district attractive to home buyers again.

Notice that, although only a fraction of households are moving at any given time, their observations of achievement and tax liabilities are "universalized" through the housing market so that *every* family's house price changes in such a way as to give schools incentives to be productive. Notice also that this productivity-inducing mechanism is sustainable over the long term because it depends on decentralized choices.[13] This is in contrast to centralized reward systems—for example, financial or other merit awards for successful school districts that are distributed by the state. Centralized rewards tend to be unsustainable because state governments cannot, ex post, credibly adhere to systems that reduce the amount of money going to failing school districts.

8.4 Finding Evidence on How Choice Affects School Productivity

In the next two sections of this paper, I show evidence on how choice affects school productivity. The next section focuses on traditional forms of choice (parents choosing among independent school districts and parents choosing private schools). Section 8.6 focuses on recent choice reforms: vouchers and charter schools. There are, however, some problems that arise in any analysis of how choice affects productivity, and the purpose of *this* section is to explain them.

8.4.1 The Endogenous Availability of Choice Options

One problem that plagues analysis is the fact that choice options do not arise randomly, but are frequently a response to school conduct. In partic-

13. The mechanism described is the subject of Hoxby (1999b), where it is described in much more detail.

ular, when people are dissatisfied with a particular school's conduct, they try to create alternative schools for themselves or maintain their access to existing alternative schools.

It is easy to see this phenomenon with respect to the creation (or maintenance) of private schools, charter schools, and voucher programs. In an area where the public schools are bad, parents are frustrated and are willing to make some effort or devote some money to obtaining alternative schooling. A collapse in the quality of local public schools (as sometimes occurs when an administrator leaves or school finance laws change) tends to send families scurrying toward local private schools. The result is an area in which private schooling is common because the public schools are bad. A recent illustration of this phenomenon is the substantial increase in private schooling that followed California's school finance equalization (Downes and Schoeman 1998).

Recent voucher and charter initiatives also illustrate this phenomenon. It is no accident that Washington, D.C. has both a voucher program and a rapidly growing population of charter schools. The Washington, D.C. district has historically had low productivity: Its per-pupil spending is in the 99th (highest) percentile for the United States, yet its average student scores between the 10th and 20th percentiles on the NAEP. Reports of malfeasance in the D.C. public schools, including the theft of school supplies and payrolls padded with nonworkers, are common.[14]

Although the mechanism is less obvious, choice options existing *because* the public schools are bad is a problem that also plagues traditional choice among public school districts. It turns out that voters resist district consolidation in areas where one or more districts (usually the largest central city district) has bad productivity. In areas where all the districts have good productivity, voters elect to consolidate them in order to enjoy economies of scale. In districts with bad productivity, subareas are keen to secede and form another district, whereas, in districts with good productivity, no such secessions occur. The end result of such phenomena is that areas with many districts often contain one or more districts with bad productivity.

Endogenous school choice in areas with bad public schools generates bias if a researcher naively estimates the effect of choice on productivity. Because schools with poor productivity induce the creation of choice, it can appear as though choice causes low productivity (instead of the other way around). Researchers can avoid this bias only by (a) comparing the same school district before and after a choice reform if panel data are available or (b) finding a source of variation in the availability of choice that is *not* correlated with the underlying causes of bad school productivity. The first so-

14. The sources are U.S. Department of Education, *Digest of Education Statistics* (2000) and U.S. Department of Education, *National Assessment of Educational Progress Long-Term Trend Summary Data Tables* (2000).

lution typically generates differences-in-differences strategies, in which schools that are "treated" with choice reforms are compared, before and after the reform, to similar control schools (which did not experience the reform). The second solution typically generates instrumental variables strategies, two of which are illustrated in section 8.5.

8.4.2 Unobserved Differences in Student Inputs That Appear to Be Differences in Productivity

Some families provide many learning opportunities and resources for their children at home; other families provide few. Children also differ in motivation and innate ability. When measuring a school's productivity, one should fully account for differences in student inputs so that one avoids describing a mediocre school as highly productive simply because it has such good student inputs that achievement is high even if it adds very little learning (beyond what its students learn at home and pick up for themselves). It is not possible, however, to measure all student inputs. In particular, motivation and innate ability are usually not observed and cannot be controlled for.

For finding the effect of choice on productivity, there are three ways that researchers can deal with this problem. Suppose a researcher wants to compare productivity across schools that face strong choice-based incentives (such as voucher or charter schools) and schools that face weak choice-based incentives (such as a large public school district that dominates a metropolitan area). Then, the researcher must ensure that a random mechanism (such as a lottery) that is *not* correlated with unobserved motivation or ability assigns students to schools. If such a mechanism is at work, schools will have an equal allocation of unobserved motivation or ability, and the difference in achievement per dollar spent will accurately reflect true differences in productivity. This approach is illustrated by Peterson et al. (chap. 4 in this volume).

An alternative is for a researcher to compare the achievement of *all* students from an environment in which there is little or no choice to that of *all* students from an environment in which there is a lot of choice. As long as the students cannot choose the environment to which they belong, this method generates good estimates. One example is comparing all students in a metropolitan area with little choice to all students in a metropolitan area with a great deal of choice. (Families are assumed to move among metropolitan areas for reasons *other* than the availability of choice.)

A final alternative is for a researcher to examine the achievement of students who are unlikely to benefit from choice unless it benefits *all* students. An example will illustrate this method. Suppose that a researcher wishes to compare school productivity before and after a private school becomes available, and the researcher sees that the private school draws students who were previously high achievers in the public schools. The researcher can compare measured productivity at the public schools before and after the private school's introduction, knowing that public schools' measured productivity is likely to rise only if the availability of private school choice benefits *all* students—that is, if the researcher may be reasonably confident that a measured increase in public school productivity is not generated by unobserved motivation and ability rising at the public school.

8.4.3 Measuring Productivity

Productivity is achievement per dollar spent in a school, and measuring productivity raises a few measurement issues, mostly related to measuring achievement. It goes almost without saying that one should avoid using measures of achievement, such as grades, that have different meanings in different schools and times. One should also avoid using scores on standardized tests that are taken by only a small, self-selected share of students, such as the Scholastic Aptitude Test (SAT1) or American College Test (ACT). Use of such tests generates self-selection bias that is impossible to solve without the use of other standardized tests that are given to the entire population of students. If one has such a populationwide standardized test, however, one should use it instead of the SAT1 or ACT.

Supposing that one has a standardized test administered to the entire population of students, there remains the question of whether to measure productivity with reading scores, math scores, science scores, elementary school scores, secondary scores, and so on. These are all valid measures of productivity, and the researcher is best off presenting several (especially math and reading). It is perfectly normal to find that a school has better productivity in some subjects or grades than in others. One may use scale scores, national percentile scores, or any other score designed by the test-maker to be comparable across schools and time.[15]

Measuring per-pupil spending presents few problems as long as the same definition is used for all schools. One may use either current spending or (preferably) total spending with smoothed capital expenditures.

8.5 The Effect of Traditional Forms of School Choice on Productivity

Parents' ability to choose among public school districts (through residential decisions) and to choose private schools are such established features of American education that they are taken for granted. Through these mechanisms, American parents have traditionally exercised some choice

15. Wages and income later in life are additional measures of achievement that are useful complements to measures based on standardized tests. They are useful because they are meaningful to people in a concrete way and because they are measured in dollars, as is the denominator of productivity. There are, however, several problems with using wages and income to measure achievement, including a paucity of data linked to schools, questionable validity for women, and the impossibility of analyzing a reform until at least twenty years after its occurrence. I do not present wage-based measures of productivity here, but see Hoxby (2000b) for some wage-based estimates.

over their children's schooling. These traditional forms of choice are useful for predicting the effects of choice on productivity, especially because the availability of traditional choice mechanisms varies greatly across metropolitan areas in the United States. Some metropolitan areas contain many independent school districts and a large number of affordable private schools. Other metropolitan areas are completely monopolized by one school district or have almost no private schooling.

In previous work, I have drawn upon traditional forms of choice to generate evidence about how choice affects productivity. I review this evidence here. In addition, I explain how traditional forms of choice generate important evidence on productivity that is otherwise unobtainable and illustrate empirical strategies for determining the effects of traditional forms of choice. For detail on the empirical work described here, see Hoxby (2000a, b). Rather than providing such detail here, I reserve space for evidence on the productivity effects of recent choice reforms (section 8.6).

8.5.1 Traditional Interdistrict Choice

The first traditional form of choice occurs when parents choose among independent public school districts by choosing a residence. The degree to which parents can exercise this form of choice depends on the number, size, and housing patterns of districts in the area of the parents' jobs. There are some metropolitan areas in the United States that have many small school districts with reasonably comparable characteristics. Boston, for instance, has seventy school districts within a thirty-minute commute of the downtown area and many more in the metropolitan area. Miami, on the other hand, has only one school district (Dade County), which covers the entire metropolitan area. Most metropolitan areas are, of course, somewhere between these two extremes. A typical metropolitan area has an amount of choice that corresponds to having four equal-sized school districts (or a greater number of unequally sized districts).

People with jobs in rural areas typically have only one or two school districts among which to choose. To avoid a much-choice/little-choice comparison that mainly reflects urban-rural differences in school productivity, it is useful to focus on metropolitan areas when analyzing traditional interdistrict choice.

It is essential that parents choose among districts that are fiscally and legally independent if this traditional form of choice is to be useful guide to the productivity effects of choice. This is because the mechanism previously described, by which parents' housing choices translate into incentives for a school to be productive, does not operate if, say, a district relies entirely on state revenue or is otherwise held harmless from repercussions associated with an inability to attract parents. *Intra*district choice among schools does not provide useful evidence about productivity effects because the schools in a district are fiscally dependent on one another, by definition.

How does one measure the degree of traditional interdistrict choice in a metropolitan area? A particularly good index of interdistrict choice is the probability that, in a random encounter, two students in the metropolitan area would be enrolled in different school districts. If there were only one district, as in Miami, this probability would be equal to zero. If there were many districts, as in Boston, this probability would be very close to one (greater than 0.95). We can calculate this choice index, C_m, using the following equation:

$$C_m = 1 - \sum_{j=1}^{J} s_{jm}^2, \tag{7}$$

where s_{jm}^2 is the square of district j's share of enrollment in metropolitan area m. Table 8.2 lists the names and choice indexes of metropolitan areas

Table 8.2

Metropolitan Areas with the Most Choice among Public School Districts		Metropolitan Areas with Very Little Choice among Public School Districts	
Metropolitan Area	Choice Index	Metropolitan Area	Choice Index
Albany, N.Y.	0.97	Honolulu, Hawaii	0
Bergen-Pasaic, N.J.	0.97	Miami, Fla.	0
Boston, Mass.	0.97	Las Vegas, Nev.	0
Pittsburgh, Pa.	0.96	Fort Lauderdale, Fla.	0
Riverside-San Bernardino, Calif.	0.96	Daytona Beach, Fla.	0
		Fort Myers, Fla.	0
Monmouth-Ocean, N.J.	0.96	Albuquerque, N.Mex.	0
Minneapolis, Minn.	0.96	Hagerstown, Md.	0
Atlantic City, N.J.	0.95	Jacksonville, N.C.	0
San Francisco, Calif.	0.95	Sarasota, Fla.	0
St. Louis, Mo.	0.95	Odessa, Tex.	0
Binghamton, N.Y.	0.94	Cheyenne, Wyo.	0
York, Pa.	0.94	Lakeland/Winter Haven, Fla.	0
Scranton, Pa.	0.94		
Johnstown, Pa.	0.94	Reno, Nev.	0
San Jose, Calif.	0.94	Boca Raton, Fla.	0
Dayton, Ohio	0.94	Wilmington, N.C.	0
Allentown, Pa.	0.94	Ocala, Fla.	0
Anaheim-Santa Ana, Calif.	0.94	Melbourne/Palm Bay, Fla.	0
Seattle, Wash.	0.94	Panama City, Fla.	0
Rochester, N.Y.	0.94	Bradenton, Fla.	0
Phoenix, Ariz.	0.94	Portland, Oreg.	0.07
Youngstown, Ohio	0.94	Midland, Tex.	0.11

Source: U.S. Department of Education (1994b).

Notes: Hawaii is one school district, so the school district is larger than the metropolitan area of Honolulu. California has school districts that have almost no fiscal independence, so it is somewhat deceptive to describe metropolitan areas like Riverside-San Bernardino, San Francisco, San Jose, and Anaheim-Santa Ana as having significant choice among school districts.

in the United States that have very high or low degrees of interdistrict choice. It is interesting to note that metropolitan areas as disparate as Saint Louis and Seattle have comparably high degrees of interdistrict choice. Metropolitan areas as disparate as Las Vegas and Wilmington have zero interdistrict choice.

8.5.2 Traditional Choice of Private Schools

The second way in which parents have traditionally been able to exercise choice in the United States is by enrolling their children in private schools. Traditionally, private school tuition in America is not subsidized by public funds (as it is in Canada and many European countries), so parents can only afford private school if they can pay tuition and also pay taxes to support local public schools. Partly as a result, private schools enroll only 12 percent of American students.

In the United States, 85 percent of private school students attend a school with religious affiliation, but such schools include a variety of Christian and non-Christian schools and have tuition that ranges from a token amount to over $10,000. The remaining 15 percent of private school students attend schools with no religious affiliation; these include most of the independent, college-preparatory schools that charge tuition of $5,000 or more. The modal private school student in the United States attends a Catholic school that charges between $1,200 and $2,700.

A key feature of American private schools is that they typically subsidize tuition with revenues from donations or an endowment (or implicit revenues from an in-kind endowment such as buildings and land). The share of schooling cost that is covered by subsidies is larger in schools that serve low-income students, but even relatively expensive private schools charge subsidized tuition. For instance, Catholic elementary schools, on average, cover 50 percent of their costs with nontuition revenues.

The number of private school places (of a given quality) that are available at a given tuition varies greatly among metropolitan area in the United States.[16] For instance, in some metropolitan areas, 15 percent of the elementary student population is enrolled in private schools where tuition is about two-thirds of the schools' per-pupil expenditure. (Typical amounts would be tuition of $1,800 and expenditure of about $2,700). In other metropolitan areas, fewer than 1 percent of the elementary school population is enrolled in such schools, although places might be available in schools where tuition is higher because there are no tuition subsidies. In short, the

16. The quality of a private school can be measured in various ways, the simplest of which is merely the amount of money the private school spends on educating a student. Because private schools face strong incentives to be productive, their costs are a good guide to their quality. Private school expenditure sometimes understates the true cost of educating a private school student because, especially in schools with religious affiliation, labor is donated by volunteers and church buildings are used for educational purposes.

supply of private schooling varies among metropolitan areas, and, thus, so does the degree to which parents have choice between public and private schools.

It is reasonable to use the actual share of students who attend private school in a metropolitan area as a measure of private school availability *if* the measure is properly instrumented. The instruments must be variables that cause the nontuition revenue of private schools to vary but are otherwise unrelated to local public school achievement. That is, one wants to use only the variation in private school availability that is generated by factors that affect the *supply* of private schooling, not by factors that affect the *demand* for private schooling (such as the local public schools being bad). I describe the best available instruments below.

8.5.3 Why Evidence from the Traditional Forms of Choice is Necessary

Evidence from the traditional forms of choice is necessary because it can reveal the long-term, general equilibrium effects of choice. Evidence based on recent reforms cannot.

In the short term, administrators who are attempting to raise their school's productivity to respond to competition have only certain options. They can induce their staff to work harder; they can get rid of unproductive staff and programs; they can allocate resources away from non-achievement-oriented activities (building self-esteem) and toward achievement-oriented ones (math, reading, and so on). In the slightly longer term, they can renegotiate the teacher contract to make the school more efficient. If administrators actually pursue all of these options, they may be able to raise productivity substantially.

Nevertheless, choice can affect productivity through a variety of long-term, general equilibrium mechanisms that are not immediately available to an administrator. The financial pressures of choice may bid up the wages of teachers whose teaching raises achievement and attracts parents. Choice may thus draw people into teaching (or keep people in teaching) who would otherwise pursue other careers. Indeed, there is evidence that choice changes the entire structure of rewards in teaching and could thereby transform the profession. (It appears that schools under pressure from choice reward teachers more on the basis of merit and allow administrators more discretion in rewarding good teachers.)[17] The need to attract parents may force schools to issue more information about their achievement and may thus gradually make parents into better "consumers." Because parents' decisions are more meaningful when schools are financed by fees they control, choice may make schools more receptive to parent participation. The need to produce results that are competitive with those of other schools may force schools to recognize and abandon pedagogical techniques and curric-

17. See Hoxby (2002b) for more on this point.

ula that are unsuccessful in practice although philosophically appealing. Finally, in the long term, choice can affect the size and very existence of schools. Choice makes districts' enrollment expand and contract; it makes private schools enter and exit. In the short term, we mainly observe how the existing stock of schools changes its behavior.

Both traditional forms of choice potentially create the long-term, general equilibrium effects that interest us.

8.5.4 The Effect of Traditional Interdistrict Choice on School Productivity

We have a good measure of the degree of interdistrict choice in a metropolitan area: C_m, defined above. We are concerned, however, that the interdistrict choice available is endogenous to the conduct of local public schools; in particular, districts consolidate with productive districts but secede from unproductive districts. To obtain unbiased estimates, we need geographic or historical factors that increase a metropolitan area's tendency to contain many independent districts but that have no direct effect on contemporary public school conduct. As explained in Hoxby (2000b), streams and rivers provide good instruments because, early in American history, they were natural barriers that influenced the drawing of district boundaries. They increased students' travel time to school, causing school districts to be drawn smaller initially.[18] They probably have no direct effect on how schools conduct themselves now.

Formally, the set of instruments for C_m is a vector of variables that measure the number of larger and smaller streams in a metropolitan area. I estimate the effects of interdistrict choice using regressions in which the dependent variable is either achievement (the numerator of productivity) or per-pupil spending (the denominator of productivity).[19] The key independent variable is the choice index (instrumented). The key variation in the regression is at the metropolitan area level, but I am able to control for a wide range of background variables that might also influence schools or students. For instance, I control for the effect of household income, parents' educational attainment, family size, single-parent households, race, region, metropolitan area size, and the local population's income, racial composition, poverty, educational attainment, and urbanness. Because I have good measures of racial, ethnic, and income segregation by school and school

18. This typically took place about the time of Anglo-American settlement, which varies with the area of the country. Many of the original petitions for district boundaries cite streams as a reason for not extending the district lines further. Streams are by far the most common natural boundary for school districts. Note, however, that many of the streams that are preserved in boundaries are small and have never had industrial importance. Today, many of the boundary streams are of negligible importance in travel.

19. Per-pupil spending is the denominator in the measure of productivity. Because I compare productivity across metropolitan areas with widely varying costs of living, I adjust perpupil spending using the Bureau of Labor Statistics metropolitan cost-of-living indexes.

Table 8.3 Effect of Traditional Interdistrict Choice on Productivity of Public Schools

	8th-Grade Reading Score	10th-Grade Math Score	12th-Grade Reading Score
Effect on Achievement (numerator of productivity)			
An increase of 1 in the index of interdistrict choice	3.818**	3.061**	5.770**
(no choice to maximum choice)	(1.591)	(1.494)	(2.208)
Effect on Per-Pupil Spending (denominator of productivity)			
An increase of 1 in the index of interdistrict choice		–7.63%**	
(no choice to maximum choice)		(3.41)	

Sources: The main source for this table is Hoxby (2000b). Observations are metropolitan area students from the National Education Longitudinal Study. The number of observations in each column are 10,790 (from 211 metropolitan areas), 7,776 (from 211 metropolitan areas), and 6,119 (from 209 metropolitan areas). The number of observations varies due to the availability of the dependent variable. Other data sources are the U.S. Department of Education (1993, 1994b), U.S. Department of Commerce (1994), and U.S. Geological Survey (1994).

Notes: Test scores are measured in national percentile points. Per-pupil spending is measured in natural log points so that the effect of choice is recorded in percentage terms. The coefficients shown come from instrumental variables estimation of regressions in which the dependent variable is one of the achievement measures shown or per-pupil spending. The independent variables in the regression include the index of choice (instrumented by a vector of streams variables; see text), several family background variables (household income, gender, race, parents' education), several neighborhood variables (mean household income in district, income inequality in district, racial composition of district, racial and ethnic homogeneity of district, educational attainment of adults in district), and several characteristics of the metropolitan area (population, land area, mean household income, income inequality, racial composition, racial homogeneity, ethnic homogeneity, educational attainment of adults, homogeneity of educational attainment, region of the country). The regressions are weighted by school enrollment. Standard errors are in parentheses and use formulas (Moulton 1986) for data grouped by districts and metropolitan areas.

**Statistically significantly different from zero at the 95 percent level of confidence.

district, I can even control for segregation that may be affected by interdistrict choice.

The principal results of these regressions are shown in table 8.3, which displays only the coefficients of interest, not the coefficients on control variables. The estimates show that interdistrict choice has a positive, statistically significant effect on productivity.[20] We can see this by looking at the two components of productivity: achievement (the numerator of productivity), which is shown in the top panel of the table; and per-pupil spending (the denominator of productivity), which is shown in the bottom panel.

The top panel shows that a metropolitan area with maximum interdistrict choice (index approximately equal to 1) has eighth-grade reading scores that are 3.8 national percentile points higher, tenth-grade math scores that are 3.1 national percentile points higher, and twelfth-grade reading scores that are 5.8 national percentile points higher. All of these effects

20. I consistently use the words "statistically significant" to mean "statistically significantly different from zero (for a two-sided test) with at least 90 percent confidence."

are statistically significant with at least 95 percent confidence. The bottom panel of table 8.3 shows that this better achievement is attained with *lower* per-pupil spending. Per-pupil spending is 7.6 percent lower in metropolitan areas where interdistrict choice is at its maximum level (choice index equal to 1), as opposed to its minimum level (choice index equal to 0). The combination of the top and bottom panels is striking: Schools can simultaneously have significantly higher achievement and significantly lower spending only if their productivity is substantially higher.

8.5.5 The Effect of Traditional Private School Choice on School Productivity

Recall that availability of private schooling varies among metropolitan areas in the United States. To estimate the effects of varying private school competition for public schools, we need factors that affect the supply of private schooling but have no direct effect on achievement. Such factors include historical differences in metropolitan areas' religious composition because religious groups left endowments that today generate differences in the amount of nontuition revenue enjoyed by private schools. A private school presented by history with a generous endowment can provide a given quality of schooling at a lower tuition (and can thus be more competitive with public schools) than a private school with little or no endowment.

Formally, the set of instruments for the share of enrollment in private schools is a vector of variables that measure the population densities of nine major religious denominations in 1950. As long as I control for *current* religious composition of metropolitan areas (which might affect the demand for private schooling), these historical religious population densities should mainly affect the supply of schooling and should have little or no direct effect on the achievement of public school students.[21] I estimate the effects of private school choice using regressions in which the dependent variable is either achievement (the numerator of productivity) or per-pupil spending (the denominator of productivity). The key independent variable is the percentage of metropolitan-area students in private schools (instrumented). I control for the same background variables that I used for interdistrict choice (see above).

The key estimates from these regressions are shown in table 8.4, which displays only the coefficients of interest, not the coefficients on control variables. The table shows that private school choice has a positive, statistically significant effect on *public* schools' productivity. For instance, compare two metropolitan areas, one with a moderately high degree of private school supply (about 17 percent of students in private schools) and the other with a moderately low degree of private school supply (about 7 percent of students in private schools). The difference between moderately high and low

21. See Hoxby (2000a) for further comment on this point.

Table 8.4 Effect of Traditional Private School Choice on Productivity of Public Schools

	8th-Grade Reading Score	8th-Grade Math Score	12th-Grade Reading Score	12th-Grade Math Score
Effect on Achievement (numerator of productivity)				
An increase of 1 in the index of interdistrict choice	0.271**	0.249**	0.342**	0.371**
(no choice to maximum choice)	(0.090)	(0.090)	(0.172)	(0.171)
Effect on Per-Pupil Spending (denominator of productivity)				
An increase of 1 in the index of interdistrict choice		0.85%		
(no choice to maximum choice)		(0.68)		

Sources: The main source for this table is Hoxby (2000a). Observations are metropolitan area students from the National Education Longitudinal Study. Other data sources are U.S. Department of Education (1993, 1994b) and U.S. Department of Commerce (1994).

Notes: Test scores are measured in national percentile points. The coefficients shown come from instrumental variables estimation of regressions in which the dependent variable is one of the achievement measures shown. The independent variables in the regression include the percentage of metropolitan area students enrolled in private schools (instrumented by a vector of religious composition variables from 1950; see text), several family background variables (household income, gender, race, parents' education), several neighborhood variables (mean household income in district, income inequality in district, racial composition of district, racial and ethnic homogeneity of district, educational attainment of adults in district), and several characteristics of the metropolitan area (population, land area, mean household income, income inequality, racial composition, racial homogeneity, ethnic homogeneity, educational attainment of adults, homogeneity of educational attainment, region of the country). The regressions are weighted by school enrollment. Standard errors are in parentheses and use formulas (Moulton 1986) for data grouped by districts and metropolitan areas.

**Statistically significantly different from zero at the 95 percent level of confidence.

private school choice is, thus, a 10 percentage point difference in the share of students in private schools. This means that we can interpret the coefficient shown in the top panel of table 8.4 as follows. A *public* school in the metropolitan area with moderately high private school choice (as opposed to moderately low private school choice) has eighth-grade reading scores that are 2.7 national percentile points higher, eighth-grade math scores that are 2.5 national percentile points higher, twelfth-grade reading scores that are 3.4 national percentile points higher, and twelfth-grade math scores that are 3.7 national percentile points higher.

Of course, in order to see whether these effects on achievement are generated by higher productivity or just higher spending, we need to examine the effect of private school choice on per-pupil spending in the public schools. This result is shown in the bottom panel of table 8.4. Compared to public schools in metropolitan areas with moderately low private school choice, public schools in areas with moderately high private school choice have per-pupil spending that is 0.53 percent (approximately half of 1 percent) higher. Not only is this change very small, but it is not statistically significantly different from zero. In other words, traditional private school choice has no effect on public school spending. This is probably because of offsetting effects. Increased availability of private school choice draws some students away from the public schools, raising per-pupil spending through the reduction in the number of pupils served but lowering per-pupil spending through the reduction in voters who will support higher public school spending.

In summary, the effect of private school choice on productivity is substantial and occurs purely through an effect on achievement: Per-pupil spending in the public schools does not change, but their achievement is higher.

8.5.6 Discussion of the Effects of Traditional Forms of School Choice

Are the effects of traditional choice on productivity large or small? One way to answer this question is to ask how much higher American school productivity would be if every school were to experience a high level of interdistrict choice and private school choice, as opposed to zero interdistrict choice and moderately low private school choice. There would be a 28 percent improvement in American school productivity, based on the estimates described above. 28 percent is close to half of the decline in American school productivity since 1970.

One should keep in mind, however, that both traditional forms of choice provide rather weak incentives compared to choice reforms like vouchers and charter schools. Moreover, many poor families cannot exercise either of the traditional forms of choice: A family can only choose among districts if it can afford to live in a variety of areas, and it can only exercise traditional private school choice if it can pay tuition. Thus, even if every metropolitan area in the United States had the maximum degree of the traditional forms of choice, poor families would probably be left with relatively unproductive schools.

8.6 The Effect of Recent Choice Reforms on School Productivity

As mentioned above, recent choice reforms can only partially answer our questions about how competition affects productivity. The recent vintage of most reforms means that we are unlikely to witness major changes in the supply of schools. Also, short-term reactions to choice can differ from long-term reactions. For instance, consider a regular public school that has had low productivity for years and that has become the target of voucher or charter school competition. Under pressure, the school might make dramatic productivity gains in the short run. The principal might quickly eliminate unsuccessful instructional programs or personnel and reallocate resources toward core instructional programs in reading, language, math, history, and science. The rate of productivity increase might, however, slow after the first few years as good policy changes become less obvious. On the other hand, even a school that is raising its productivity might appear to have productivity losses in the short run if it faces adjustment costs when it makes changes. For instance, a school that puts an academic monitoring system in place may face short-run costs for computers and training.

Can we learn much, then, from recent choice reforms? The answer is yes if we follow a few principles. First, although it is interesting to examine the productivity of the choice schools themselves (as Peterson et al. implicitly do in chap. 4 in this volume), it is even more important to study the productivity reactions of regular public schools that are newly facing competition. This is because the productivity reactions of regular public schools are in much more doubt than the productivity of choice schools. An unproductive choice school is unlikely to enter and even less likely to survive, but critics of school choice doubt whether regular public schools even have the knowledge or tools to raise their productivity. Second, we should focus on the productivity reactions of regular public schools that face nonnegligible incentives due to a choice reform. This immediately limits our investigation to a few choice reforms that meet the following requirements: (a) There is a realistic possibility that at least 5 percent of regular public enrollment could go to choice schools; (b) the regular public schools lose at least some money (not necessarily the entire per-pupil cost) when a student goes to a choice school, and (c) the reform has been in place for a few years. Three reforms that satisfy these basic requirements are school vouchers in Milwaukee, charter schools in Michigan, and charter schools in Arizona. I describe each of these reforms below in the course of examining the reaction to it. Apart from these three reforms, most choice reforms fail to meet at least one of these requirements. In fact, choice reforms are typically characterized by constraints on enrollment (for instance, no more than 1 percent of local students can attend choice schools) or perverse financial incentives (for instance, the local district loses no money when it loses a student to a choice school).[22]

8.6.1 The Effect of Vouchers on Achievement in Milwaukee Public Schools

Vouchers for poor students in Milwaukee were enacted in 1990 and were first used in the 1990–91 school year. Currently, a family is eligible for a voucher if its income is at or below 175 percent of the federal poverty level (at or below $17,463 for a family of four).[23] For the 1999–2000 school year, the voucher amount was $5,106 per student or the private school's cost per

22. See Rees (2000) for a thorough review of current school choice reforms. In most cases where I have not used materials directly obtained from the relevant state's department of education, I have relied upon Rees for a description of reforms.

23. As a rule, any child who is eligible for free or reduced-price lunch is also eligible for a voucher. The actual cutoff for reduced-price lunch is 185 percent of the federal poverty level, but the difference between 175 percent (the cutoff for the vouchers) and 185 percent is not rigorously enforced (and would be difficult to enforce).

student, whichever was less. For every student who leaves the Milwaukee public schools with a voucher, the Milwaukee public schools lose state aid equal to half the voucher amount (up to $2,553 per voucher student in 1999–2000). Milwaukee's per-pupil spending in 1999–2000 was $8,752 per pupil, so the district was losing 29 percent of the per-pupil revenue associated with a voucher student. Currently, the vouchers may be used at secular and nonsecular private schools.[24]

The voucher program had a difficult start. Although approximately 67,000 students were initially eligible for vouchers, participation was initially limited to only 1 percent of Milwaukee enrollment. In 1993, the limit was raised to 1.5 percent and, in 1998, to 15 percent of enrollment. The 1998 changes followed a prolonged legal dispute in which most voucher students had to use privately donated, not publicly funded, vouchers. For instance, in 1997–98, only 1,500 students (about 1.4 percent of Milwaukee students) were able to use publicly funded vouchers. Also, until 1998, the future of the program was very much in doubt.[25] Overall, although the voucher program began in 1990 and might have been expected to have had a small impact on the Milwaukee public schools beginning with the 1990–91 school year, the program generated very little potential competition until the 1998–99 school year. However, because the program was already somewhat established and familiar to Milwaukee residents by 1998, one would expect a quicker response to the program than for a completely new program. In short, it is plausible to look for a productivity impact, if any, over the few most recent school years. The 1996–97 school year effectively predates serious competition.

Not all schools in Milwaukee experienced the same increase in competition as the result of the voucher program. The greater a school's share of poor children, the greater the potential competition, because the greater was the potential loss of students (after 1998). Some Milwaukee schools had as few as 25 percent of their students eligible for vouchers, whereas other Milwaukee schools had as many as 96 percent eligible. Also, because private elementary schools cost significantly less than private high schools, more than 90 percent of vouchers were used by students in grades one through seven in 1999–2000. Thus, only elementary schools in Milwaukee faced significant potential competition.

These facts about the voucher program suggest that the following type of evaluation is most appropriate for examining the productivity response of

24. The information on the Milwaukee program and Wisconsin schools is obtained from several publications of the Wisconsin Department of Public Instruction (2000a–e).

25. The future of the program is still somewhat in doubt, for two reasons. First, state supreme courts' opinions conflict on the question of whether it is constitutional to have vouchers that can be used at schools with religious affiliation. Therefore, it is likely that the U.S. Supreme Court will eventually rule on such vouchers. Second, the Wisconsin legislature has threatened to fund the vouchers at such a low level that they are unusable.

Milwaukee public schools. First, one should focus on the productivity of Milwaukee schools in grades one through seven. Second, schools' productivity should be compared from 1996–97 (before significant competition) to 1999–2000 (after significant competition). Third, schools in Milwaukee can be separated into those that were "more treated" by competition because a large number of students were eligible and those that were "less treated." More-treated schools are likely to have responded more strongly to the program. We can think of the less-treated schools in Milwaukee as a partial control group, but *all* schools in Milwaukee were eligible for nonnegligible treatment. Therefore, it is desirable to have a control group of schools from Wisconsin that were truly unaffected by the voucher program. It turns out that it is not easy to find such schools in Wisconsin because Milwaukee's schools are much poorer and have much larger shares of black and Hispanic students than most other schools in the state. I chose the most similar schools available for the evaluation, but it is likely that the results will understate the productivity effects of school competition. We expect understatement because schools that have fewer poor and minority students typically enjoy greater productivity and higher productivity *growth* than schools with more poor and minority students. Thus, the control schools, which are richer than the treated schools, would probably have higher productivity growth (all else being equal) than the treated group of schools. Also, the less-treated schools in Milwaukee would probably have higher productivity growth (all else being equal) than the more-treated schools.[26]

Because my evaluation compares *treated and control* schools *before and after* 1998, it is what is sometimes called a difference-in-differences evaluation. It has a fairly obvious analog in scientific experiments.

Table 8.5 shows some demographic indicators for the three groups of elementary schools: most treated (Milwaukee schools where at least two-thirds of students were eligible for vouchers), somewhat treated (Milwaukee schools where less than two-thirds of students were eligible for vouchers), and untreated comparison schools. Note that 30 percent was the minimum share of students eligible for vouchers among the somewhat treated Milwaukee elementary schools. There are thirty-two most-treated and sixty-six somewhat-treated elementary schools. All of the Milwaukee elementary schools have enrollment of about seventy-one to seventy-two students in a grade.

In the most-treated schools, an average of 81.3 percent of students were

26. It is fairly obvious that better-off schools will have better productivity if one does not control for demographic differences among students. It is less obvious that better-off schools will also have better productivity growth, but in fact they do. For instance, prior to 1996, Wisconsin elementary students took statewide tests in reading (only). In the prevoucher period, productivity growth was negative (based on these reading tests) in the schools that were later to become most-treated and somewhat treated. In contrast, productivity growth (based on reading tests) was positive in the schools that form the untreated comparison group.

Table 8.5 Demographics of Wisconsin's Most-Treated, Somewhat Treated, and Untreated Comparison Schools

	% of Students Eligible for Free/ Reduced-Price Lunch	% of Students Who Are Black	% of Students Who Are Hispanic
Most-treated schools	81.3	65.4	2.9
Somewhat treated schools	44.5	49.1	13.7
Untreated comparison schools	30.4	30.3	3.0

Sources: Wisconsin Department of Public Instruction (2000a–e) and U.S. Department of Education (1994b).

Notes: "Most-treated" schools were Milwaukee elementary schools where at least two-thirds of students are eligible for free or reduced price lunches (and thus eligible for vouchers). There are thirty-two most-treated elementary schools, each of which has an average fourth grade enrollment of seventy-two students. "Somewhat treated" schools were Milwaukee elementary schools where fewer than two-thirds of students are eligible for free or reduced price lunch (and thus eligible for vouchers). In all of these schools, at least 30 percent of students are eligible for free lunch. There are sixty-six somewhat treated elementary schools, each of which has an average fourth-grade enrollment of seventy-one students. The untreated comparison schools are all the Wisconsin elementary schools that (a) are urban, (b) have at least 25 percent of their students eligible for free lunch, and (c) have at least 15 percent of their students being black. There are twelve untreated comparison elementary schools, each of which has an average fourth-grade enrollment of fifty-one students.

eligible for free or reduced-price lunches (and thus eligible for vouchers), 65.4 percent of students were black, and 2.9 percent of students were Hispanic. In the somewhat treated schools, an average of 44.5 percent of students were eligible for vouchers, 49.1 percent of students were black, and 13.7 percent of students were Hispanic.[27]

I included a Wisconsin elementary in the untreated comparison group if it (a) was not in Milwaukee, (b) was urban, (c) had at least 25 percent of its students eligible for free or reduced-price lunch, and (d) had black students composing at least 15 percent of its students. There were only twelve schools in Wisconsin that met these criteria. It was not possible to choose a group of untreated schools that were more closely matched to Milwaukee schools. In the untreated comparison schools, average enrollment in a grade was fifty-one students, 30.4 percent of students were eligible for free or reduced-price lunch (and, thus, would have been eligible for vouchers had they lived in Milwaukee), 30.3 percent of the students were black, and 3.0 percent of students were Hispanic.

Students in Wisconsin take statewide examinations in grades four, eight, and ten. Because I am necessarily focusing on the productivity reactions of elementary schools, I measure productivity by dividing a school's fourth-

27. Note that all of these demographic numbers reflect what the schools looked like in 1990, before the voucher program was enacted. This is the correct method for choosing treated and control schools. One does not want to measure the extent of treatment using measures of student composition that potentially reflect how students reacted to the voucher program.

Table 8.6 **Productivity Time Trends in Wisconsin Most-Treated, Somewhat Treated, and Untreated Comparison Schools, from Regressions with School Fixed Effects**

	Productivity Calculation is Based on Exam in:				
Annual Change in Productivity by School Type	Math	Science	Social Studies	Language	Reading
Most-treated schools	0.732[a]	0.889[ac]	0.475[bc]	0.248[ad]	–0.035[bd]
	(0.071)	(0.072)	(0.070)	(0.066)	(0.066)
Somewhat treated schools	0.527	0.729[a]	0.327	0.123[b]	–0.141
	(0.056)	(0.057)	(0.055)	(0.052)	(0.052)
Untreated comparison schools	0.342	0.255	0.188	–0.081	–0.235
	(0.172)	(0.176)	(0.170)	(0.160)	(0.162)

Sources: Wisconsin Department of Public Instruction (2000a–e) and U.S. Department of Education (1994b).

Notes: Productivity is measured in national percentile points per thousand dollars of per-pupil spending, where per-pupil spending is measured in 1999 dollars. The deflator used is the Consumer Price Index. Each regression includes a fixed effect for each school, a time trend for most-treated schools, a time trend for somewhat treated schools, and a time trend for untreated comparison schools. The observations are school-level averages based on fourth-graders' scores, and the regressions are therefore weighted by the schools' fourth-grade enrollment.

[a]Time trend is statistically significantly different from the time trend for untreated comparison schools at the 95 percent level of confidence.

[b]Time trend is statistically significantly different from the time trend for untreated comparison schools at the 85 to 95 percent level of confidence.

[c]Time trend is statistically significantly different from the time trend for somewhat treated schools at the 95 percent level of confidence.

[d]Time trend is statistically significantly different from the time trend for somewhat treated schools at the 85 to 95 percent level of confidence.

grade score (expressed in national percentile points) by its per-pupil spending in thousands of real (1999) dollars. Achievement is measured on five tests: mathematics, science, social studies, language, and reading. It is worth noting that, during the period in question, Wisconsin enacted a controversial new reading curriculum that emphasized whole-language methods, as opposed to phonics.

Table 8.6 shows productivity growth rates in most-treated, somewhat treated, and untreated comparison schools in Wisconsin between 1996–97 and 1999–2000. The statistics in the table are based on regressions in which the dependent variable is productivity and the independent variables are an indicator for each school, a time trend for most treated schools, and time trend for somewhat treated schools, and a time trend for untreated comparison schools. This regression incorporates the best differences-in-differences method, given the application, because it allows each school to have its own starting point for productivity. Intuitively, the regression is based on the idea that productivity growth rates might look like the following figure.

Figure 8.1 shows what productivity might look like in three schools, one

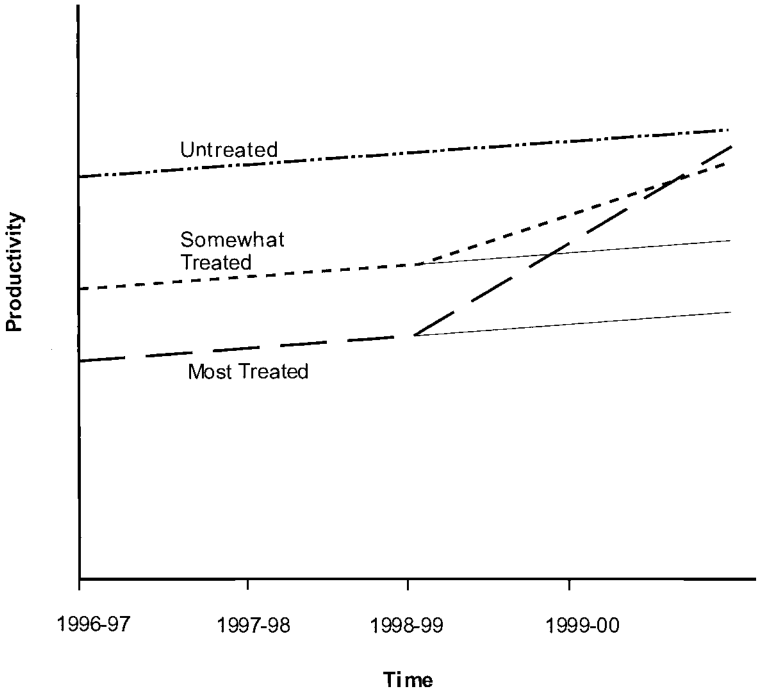

Fig. 8.1 How schools might change their productivity when facing competition

of which is most treated, one of which is somewhat treated, and one of which is untreated. It would be fairly typical to find that the most-treated schools had the lowest initial productivity if we did not correct for differences in student demographics, for the simple reason that poorer students tend to have lower achievement and the most-treated schools have more poor students. Thus, the figure shows the most-treated school having the lowest initial productivity, somewhat treated schools having medium initial productivity, and untreated schools having the highest initial productivity. If competition has little or no effect on productivity, then the time trends for productivity might all be stable, as indicated by the solid lines. On the other hand, if competition makes schools raise their productivity, then the time trends might look the dashed lines, in which the most-treated schools raise their productivity the most, somewhat treated schools raise their productivity somewhat, and untreated schools raise their productivity the least (or perhaps not at all).[28]

28. Actually, we expect the untreated schools to have higher *initial* productivity growth because richer schools tend to have better productivity growth, all else being equal. This tendency (richer schools, higher productivity growth) will make the difference-in-difference estimates understate the effect of competition on productivity.

Formally, the regression equation can be written as follows:

$$(8)\quad \frac{\text{Ach}_{it}}{\text{PPExp}_{it}} = \alpha_1 I_1 + \ldots + \alpha_N I_N + \beta^{\text{most treated}} I_i^{\text{most treated}} \text{time}_t + \beta^{\text{somewhat treated}} I_t^{\text{somewhat treated}} \text{time}_t + \beta^{\text{untreated}} I_i^{\text{untreated}} \text{time}_t + \varepsilon_{it},$$

where Ach_{it} is a national percentile rank score for students at school *i* in year *t*, PPExp_{it} is the per-pupil expenditure at school *i* in year *t*, I_1 through I_N are indicator variables for schools, α_1 through α_N are initial productivity levels at individual schools, $I^{\text{most treated}}$ is an indicator variable for the school being most treated, $I^{\text{somewhat treated}}$ is an indicator variable for the school being somewhat treated, $I^{\text{untreated}}$ is an indicator variable for the school being untreated, and time_t is the school year. The coefficients $\beta^{\text{most treated}}$, $\beta^{\text{somewhat treated}}$, and $\beta^{\text{untreated}}$ pick up the different productivity growth rates for most-treated, somewhat treated, and untreated schools, respectively.

The left-hand column of table 8.6 shows that, based on mathematics achievement, productivity grew annually by about 0.7 national percentile points per thousand dollars between 1996–97 and 1999–2000 in the most-treated schools. It grew more slowly in somewhat treated schools (about 0.5 national percentile points per thousand dollars) and yet more slowly in untreated schools (about 0.3 national percentile points per thousand dollars). Productivity growth based on science, social studies, and language (grammar) is shown in the next three columns, all of which show patterns that are similar to the mathematics-based pattern. In all these columns, productivity growth in the most-treated schools is statistically significantly different from that in the untreated schools, with a 95 percent confidence level. Reading-based measures of productivity are falling in all the schools over the period in question, perhaps because of whole language methods. However, reading-based productivity is falling least quickly in schools that were most treated to voucher school competition.

Table 8.7 shows statistics that are very similar to those in table 8.6. They are easier to interpret for those unfamiliar with regression analysis, but they are less ideal because each school does not have its own initial level of productivity. For instance, examine the top panel, which shows productivity calculations based on the mathematics exam. In 1996–97, the most-treated schools earned 4.18 national percentile points for every thousand dollars of per-pupil spending. In the same year, the somewhat treated and untreated schools earned 4.08 and 5.65 national percentile points (respectively) for every thousand dollars. Over the next few years, however, productivity growth was the highest in the most-treated schools, second highest in the somewhat treated schools, and lowest in the untreated schools (see right-hand column). In fact, by 1999–2000, productivity in the most-treated schools was closer to that of the untreated schools than it was to that of the somewhat treated schools! The productivity growth rates shown in table 8.7

Table 8.7 **Productivity in Wisconsin's Most-Treated, Somewhat Treated, and Untreated Comparison Schools**

	1996–1997	1999–2000	Annual Change
Productivity calculation is based on math exam			
Most-treated schools	4.18	6.09	0.64
Somewhat treated schools	4.08	5.50	0.47
Untreated comparison schools	5.65	6.65	0.33
Productivity calculation is based on science exam			
Most-treated schools	3.87	6.04	0.72
Somewhat treated schools	3.91	5.67	0.59
Untreated comparison schools	6.33	6.92	0.20
Productivity calculation is based on social studies exam			
Most-treated schools	5.05	6.19	0.41
Somewhat treated schools	5.26	5.80	0.18
Untreated comparison schools	6.90	7.21	0.10
Productivity calculation is based on language exam			
Most-treated schools	5.07	5.64	0.19
Somewhat treated schools	5.07	5.28	0.07
Untreated comparison schools	6.04	5.85	–0.06
Productivity calculation is based on reading exam			
Most-treated schools	5.35	5.31	–0.01
Somewhat treated schools	5.46	4.98	–0.16
Untreated comparison schools	6.68	6.04	–0.21

Sources: Wisconsin Department of Public Instruction (2000a–e) and U.S. Department of Education (1994b).

Notes: Productivity is measured in national percentile points per thousand dollars of per-pupil spending, where per-pupil spending is measured in 1999 dollars. The deflator used is the Consumer Price Index. Statistics are based on weighted averages over schools in the relevant group, where each school is weighted by its enrollment.

are dramatic for the most-treated schools. The basic pattern (highest productivity growth in the most-treated schools) is repeated in the other panels of the table, for the science, social studies, language, and reading examinations.

Tables 8.8 and 8.9 are very much like tables 8.6 and 8.7, except that they show achievement growth instead of productivity growth. That is, they leave out the changes in productivity that come about as a result of changes in per-pupil spending. An examination of them shows that achievement growth displays patterns like that of productivity growth, which suggests that the improvements in productivity in the most-treated and somewhat treated schools occurred because achievement was rising in those schools, not because achievement was holding steady while per-pupil spending fell. (Indeed, use of the vouchers causes per-pupil spending to rise in the Milwaukee public schools, so if achievement were to hold steady, productivity would fall if schools did not respond to competition by raising it.)

Look, for example, at table 8.8. It shows that math scores rose by about 7

Table 8.8 Achievement Growth in Wisconsin's Most-Treated, Somewhat Treated, and Untreated Comparison Schools, from Regressions with School Fixed Effects

Annual Change in Test Scores by School Type	Math	Science	Social Studies	Language	Reading
Most-treated schools	7.06[ac]	8.39[ac]	4.97[bc]	2.98[ad]	0.57[bd]
	(0.61)	(0.62)	(0.59)	(0.56)	(0.57)
Somewhat treated schools	5.27	6.99[a]	3.68	1.88[b]	–0.37
	(0.48)	(0.49)	(0.46)	(0.44)	(0.45)
Untreated comparison schools	3.71	2.96	2.40	–0.10	–1.42
	(1.48)	(1.50)	(1.43)	(1.37)	(1.38)

Sources: Wisconsin Department of Public Instruction (2000a–e) and U.S. Department of Education (1994b).

Notes: Test scores are measured in national percentile points. Each regression includes a fixed effect for each school, a time trend for most-treated schools, a time trend for somewhat treated schools, and a time trend for untreated comparison schools. The observations are school-level averages for fourth-graders, and the regressions are therefore weighted by the schools' fourth-grade enrollment.

[a]Time trend is statistically significantly different from the time trend for untreated comparison schools at the 95 percent level of confidence.

[b]Time trend is statistically significantly different from the time trend for untreated comparison schools at the 85 to 95 percent level of confidence.

[c]Time trend is statistically significantly different from the time trend for somewhat treated schools at the 95 percent level of confidence.

[d]Time trend is statistically significantly different from the time trend for somewhat treated schools at the 85 to 95 percent level of confidence.

percentile points per year in the most-treated schools, by about 5 percentile points per year in somewhat treated schools, and by about 4 percentile points in untreated schools. Alternatively, examine table 8.9. It shows that social studies scores in the most-treated schools rose by 4.2 percentile points per year, whereas social studies scores in untreated schools rose by only 1.5 percentile points per year.

Overall, an evaluation of Milwaukee suggests that public schools have a strong, positive productivity response to competition from vouchers. The schools that faced the most potential competition from vouchers had the best productivity response. In fact, the schools that were most treated to competition had dramatic productivity improvements. On the one hand, such bursts of productivity growth may slow down after a few more years of competition. On the other hand, the productivity effects of competition may be understated because the control group of schools was a slightly unfair comparison group, with fewer poor and minority students.

8.6.2 The Effect of Charter Schools on Achievement in Michigan Public Schools

In 1994, Michigan enacted a charter school law as part of a series of changes in its method of financing schools. Michigan charter schools receive a per-pupil fee that is essentially the same as the state's foundation level of per-pupil spending (the state's minimum level of per-pupil spending, given

Table 8.9 Fourth-Grade Test Scores in Wisconsin's Most-Treated, Somewhat Treated, and Untreated Comparison Schools

	1996–97	1999–2000	Annual Change
Math			
Most-treated schools	34.5	53.3	6.3
Somewhat treated schools	33.7	48.2	4.8
Untreated comparison schools	50.0	60.6	3.5
Science			
Most-treated schools	31.9	52.8	7.0
Somewhat treated schools	32.3	49.7	5.8
Untreated comparison schools	56.0	62.9	2.3
Social studies			
Most-treated schools	41.6	54.2	4.2
Somewhat treated schools	43.4	50.7	2.4
Untreated comparison schools	61.0	65.6	1.5
Language			
Most-treated schools	41.8	49.4	2.5
Somewhat treated schools	41.8	46.2	1.5
Untreated comparison schools	53.4	53.2	–0.1
Reading			
Most-treated schools	44.2	46.5	0.8
Somewhat treated schools	45.1	43.6	–0.5
Untreated comparison schools	59.0	55.0	–1.3

Sources: Wisconsin Department of Public Instruction (2000a–e) and U.S. Department of Education (1994b).

Notes: Test scores are measured in national percentile points. Statistics are based on weighted averages over schools in the relevant group, where each school is weighted by its enrollment.

the characteristics of the school's student population). For instance, in 1999–2000, the average charter school student in Michigan had $6,600 spent on his education, whereas the average regular public school student had about $7,440 spent on his education. Detroit public schools spent $8,325 per pupil, and the average charter school student in Detroit had about $6,590 spent on his education. A district that loses a student to a charter school loses approximately the foundation level of per-pupil revenue. Charter competition tends to be most substantial in the elementary grades because the charter fees more adequately cover costs for the lower grades. By the 1999–2000 school year, approximately 3.5 percent of all nonprivate elementary students in Michigan were enrolled in charter schools. The corresponding number for secondary students was 0.7 percent. Charter schools can receive their charters from statewide organizations, such as universities, so they can compete with local public schools, unlike charter schools in many other states that have their charters granted and renewed by their local district.[29]

29. The information on Michigan charter schools and all the data on Michigan schools are taken from publications of the Michigan Department of Education (2000a–d).

A difference-in-differences strategy, analogous to the strategy used on Milwaukee, is appropriate for evaluating the effect of charter school competition on Michigan public schools. There are two additional issues, however, that did not arise with Milwaukee. First, it was easy to define ex ante the treatment and control schools in Wisconsin: No school outside of Milwaukee received any voucher treatment, and the scale of treatment within Milwaukee schools varied with students' poverty, a variable that we observe. In Michigan, "treatment" and "control" and "before" and "after" must be defined on a district-by-district basis, so that a district is being "treated" and is in the "after" period once it is forced to recognize that it is losing a critical share of students to charter schools. Of course, we do not know what this critical share might be, but it is useful to know that the mean year-to-year change in a Michigan school's enrollment *prior to 1994* was 5.1 percent. Therefore, a small drawing-away of enrollment by a local charter school would be hard to differentiate from normal year-to-year variation in enrollment. However, a persistent drawing-away of enrollment of more than 5 percent, say, would be likely to be noticed and attributed to charter schools. I initially looked for a critical level of 6 percent, and, because it worked well, I kept it. A critical level of 7 or 8 percent works very similarly.[30]

The left-hand side of table 8.10 lists the Michigan districts in which charter schools account for at least 6 percent of total enrollment inside the district's boundaries.[31] There are 597 districts in Michigan and only 34 listed in the table, so a nonnegligible charter school presence is still the exception and not the rule. Michigan's large city districts are well represented among the districts that face charter school competition: Detroit, Lansing, and Kalamazoo all have at least 6 percent of enrollment in charter schools.

Second, the Michigan districts that had to face competition from charter schools were not selected randomly or according to a simple rule. Instead, charter schools probably formed as a response to local circumstances. In some cases, charter schools may have formed where parents were unusually active and concerned about education (good circumstances for public school productivity and achievement). In other cases, charter schools may have formed where parents and teachers were frustrated because the district

30. Results for a critical level of 7 or 8 percent are available from the author. If one chooses a critical level much higher than 8 percent, the results depend unduly on just a few districts, simply because only a few districts ever face more than an 8 percent drawing-away of their students. Descriptive statistics for the Michigan data set are also available from the author.

31. Note that the charter schools' share of local enrollment is based, in table 8.10, on the assumption that students attend charter schools in the district in which they reside. Because students who are in particularly unappealing districts are disproportionately likely to attend a charter school outside their district if they do attend a charter school, the statistics on which the table is based slightly understate the enrollment losses of bad districts. It is possible to construct estimates of the share of a district's students who attend charter schools, but such estimates are somewhat noisy and (in any case) generate results that are qualitatively similar to the results shown in tables 8.11 and 8.12. The alternative set of results may be found in the working paper version of this paper, available from the author.

Table 8.10 Michigan School Districts and Arizona Municipalities Where at Least 6 Percent of Enrollment Entered Charter Schools

Michigan School Districts		Arizona Municipalities	
Alba	Inkster-Edison	Benson	Keams Canyon
Bark River-Harris	Jackson[b]	Bisbee	Kingman[b]
Big Rapids	Kalamazoo[a]	Camp Verde	Mayer
Boyne Falls	Kenowa Hills	Cave Creek	Page
Buena Vista	Kentwood[b]	Chinle	Phoenix[a]
Caledonia	Lansing[a]	Chino Valley	Pima
Charlevoix	Mount Pleasant	Clarkdale	Prescott
Coldwater	Oak Park	Concho	Queen Creek
Detroit[a]	Onekama	Coolidge	Safford
Elk Rapids	Pentwater	Cottonwood	Saint Johns
Flat Rock	Petoskey	Enrenberg	Scottsdale[a]
Forest Hills[b]	Sault Sainte Marie	Flagstaff[b]	Sedona
Godwin Heights	Southfield[b]	Fountain Hills	Show Low
Grand Blanc[b]	Wane-Westland[a]	Gilbert[b]	Sierra Vista
Hartland	Westwood	Globe	Tempe[a]
Hillsdale	Wyoming[b] Avondale[b]	Golden Valley	Tuba City
Holland[b]		Green Valley	Vail
Huron		Higley	Winslow

Sources: Michigan Department of Education (2000a–d) and Arizona Department of Education (2000a–d).

Notes: The share of students who live in a district and attend charter schools is difficult to calculate because students can attend charter schools located outside of their districts (Michigan) or municipality (Arizona). These statistics are calculated under the assumption that students attend a charter school located in their district (Michigan) or municipality (Arizona).

[a]Very large city district (enrollment in one grade typically exceeds 1,000).

[b]Large city district (enrollment in one grade is typically between 500 and 1,000).

was run poorly (bad circumstances for public school productivity and achievement). Thus, it is important that the difference-in-differences strategy look *within* a school—that is, how a *given* school changes when it is faced with new competition. I present differences-in-differences results that control for school fixed effects, which pick up all the unobserved characteristics of a school that are stable over the several-year period that I analyze.

The difference-in-differences strategy might not be convincing, however, if the districts that were eventually forced to complete with charter schools had preexisting productivity *trends* that were different from other public schools in Michigan. Different preexisting trends would not be unlikely because charter schools chose where to locate: A charter school would expect to find little demand for its services in a district that was improving rapidly on its own. In cases where different trends are possible, a more sophisticated, detrended differences-in-differences strategy is appropriate. Therefore, I also present estimates of how schools' productivity *trends* changed when they began to face charter competition.

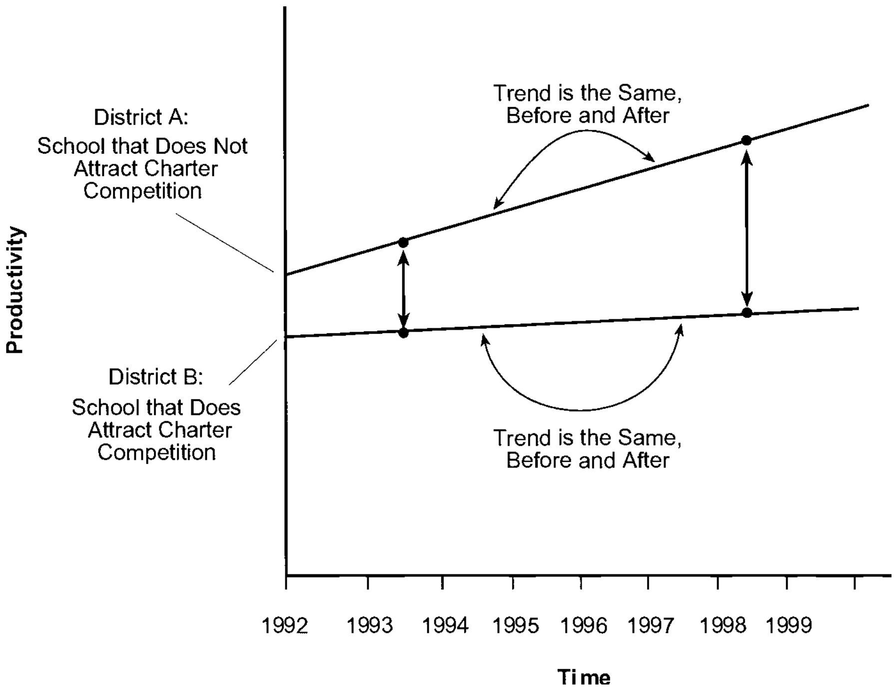

Fig. 8.2 The fact that there is no reaction to competition can be discerned, even with different preexisting trends

Figures 8.2 through 8.4 illustrate why a detrended differences-in-differences strategy can be a useful complement to a typical differences-in-differences strategy. In figure 8.2, the top line represents the productivity of district A, which initially enjoys strong positive productivity growth. The bottom line represents the productivity of district B, which initially has very low productivity growth. Suppose that charter schools are deterred from entering district A because it is already improving rapidly and parents are pleased with the current course of events. Suppose that charter schools do enter district B, however, and are able to claim a critical share of local parents (who were not pleased with the course that the public school was on) by 1996. Finally, suppose that the district B does not respond to the charter school competition: It remains on its initial path after 1996. A simple differences-in-differences strategy would compare the change in district A's *level* of productivity to the change in district B's *level* of productivity. (Notice the indications on the figure of possible "before" and "after" points that could be used for comparison). In such a comparison, charter school competition would seem to have a negative effect (although it truly has no effect), simply because charter schools enter where districts' productivity trends are already worse. On the other hand, if we compared the change in district A's

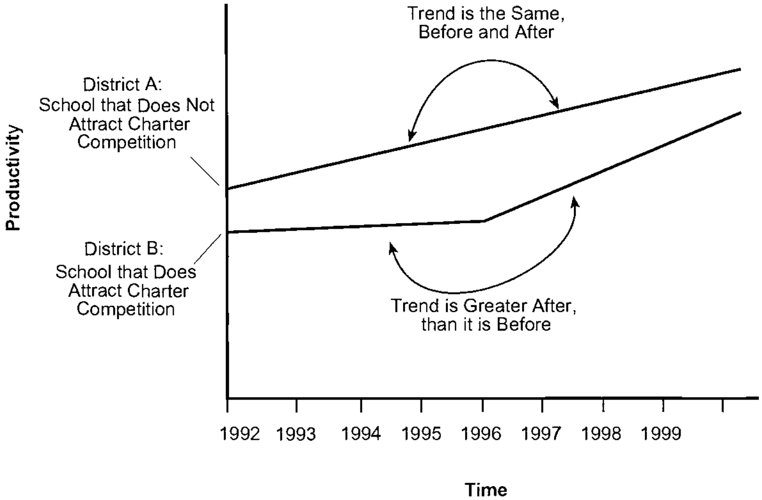

Fig. 8.3 A positive reaction to competition can be discerned, even with different preexisting trends

trend in productivity to the change in district B's *trend* in productivity, we would correctly see that district B did not respond to competition.

Figures 8.3 and 8.4 illustrate situations in which district B responds positively when it begins to face charter school competition (figure 8.3) and responds negatively when it begins to face competition (figure 8.4). Observe that the difference between district A's change in trend and district B's change in trend is an accurate indicator of the response to charter school competition. In short, the advantage of detrended difference-in-differences is that it generates consistent estimates even when schools that eventually face charter competition have different preexisting trends from schools that never face competition. The disadvantage of detrended difference-in-differences is that it demands a lot of information from the data because each school's preexisting *trend* in achievement (not just its level of achievement) must be identified. Because it is so demanding statistically, detrended difference-in-differences will not generate statistically significant estimates of effects that are small. Thus, we can foresee that the estimated effects for higher grades (which are likely to be small because charter competition affected them relatively little) are likely to be hard to identify using detrended difference-in-differences.

To summarize, it is important that difference-in-differences strategies control for *each* school's initial conditions (levels or trends). We need to control for schools' unobservable characteristics, especially characteristics that might attract charter competition. Difference-in-differences strategies also control for what was happening to Michigan schools in general over

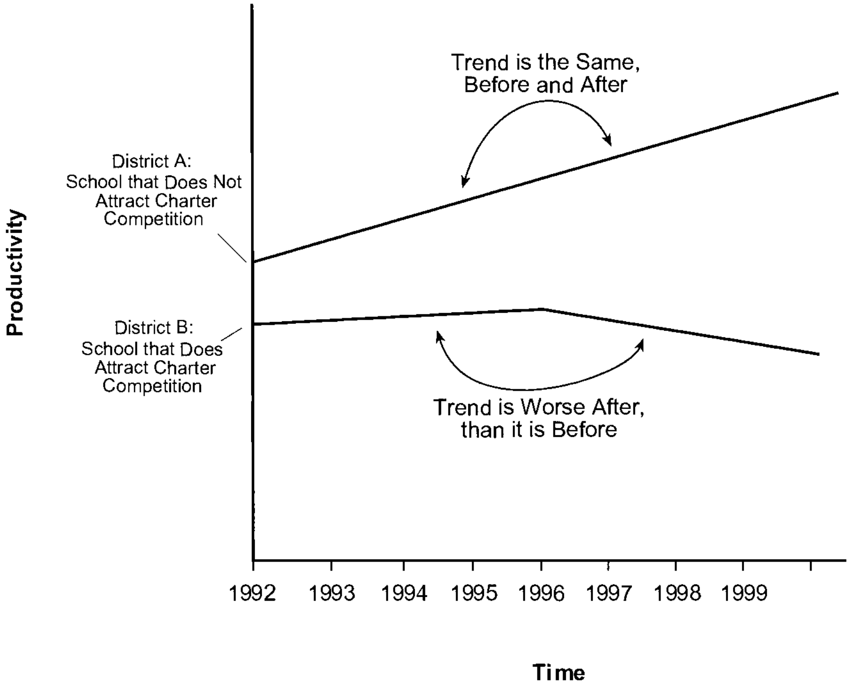

Fig. 8.4 A negative reaction to competition can be discerned, even with different preexisting trends

the period. This is important as well because Michigan enacted a major school finance reform in 1994 that affected all schools in the state. The strategies will identify changes that occurred in schools facing competition, *above and beyond* whatever occurred in other schools in the state (which were presumably responding to the finance reform).

I use regression to carry out both the simple difference-in-differences analysis and the detrended difference-in-differences analysis. The top panel of table 8.11 presents the estimated effect of charter school competition on productivity, using the simple differences-in-differences analysis. The bottom panel presents the estimated effect on productivity, using the detrended analysis. Formally, the regression used in the top panel is

$$(9) \qquad \frac{\text{Ach}_{ijt}}{\text{PPExp}_{ijt}} = I_{ij}^{\text{school}}\delta^{\text{school fixed effects}} + I_t^{\text{year}}\delta^{\text{year fixed effects}} + \delta^{\text{critical charter competition}} I_{jt}^{\text{charter}\geq 6\%} + \varepsilon_{jt} + \varepsilon_{ijt}$$

where Ach_{it} is the scale score for school i in district j in year t, PPExp_{it} is the per-pupil expenditure for the same school, $\mathbf{I}^{\text{school}}$ is a vector of school indi-

Table 8.11 Effects of Charter School Competition on Michigan Public Schools' Productivity

	4th-Grade Reading Exam	4th-Grade Math Exam	7th-Grade Reading Exam	7th-Grade Math Exam
Dependent Variable (productivity based on exam)				
Difference-in-differences (levels)				
Change in productivity level after district is faced with charter school competition	1.60**	1.37**	1.87**	1.53**
	(0.45)	(0.39)	(0.86)	(0.73)
Dependent Variable (change in productivity based on exam)				
Detrended difference-in-differences				
Change in productivity trend after district is faced with charter school competition	0.31*	0.27*	0.15	0.06
	(0.17)	(0.14)	(0.46)	(0.54)

Source: Michigan Department of Education (2000a–d).

Notes: Standard errors in parentheses. Regressions include school fixed effects and year fixed effects. Charter schools represent at least 6 percent of enrollment in district. The table is based on regressions of school-level data from 1992–93 to 1999–2000. In the top panel, the dependent variable is a school's productivity—specifically, a school's scale scores divided by its per-pupil spending in thousands of 1999 dollars. In the bottom panel, the dependent variable is the trend (annual change) in a school's productivity (or this year's productivity minus last year's). The regression includes school indicator variables to pick up characteristics of schools that are constant over the period (location, neighborhood, organization) and year indicator variables that allow for statewide changes from year to year in the test itself or in the pressure to perform on the test. The inflator for per-pupil spending is the Consumer Price Index. The numerator for productivity is the school's scale score on the Michigan Assessment of Educational Progress (MEAP) tests, which are administered to fourth- and seventh-graders. See the text for details on the tests.

**Change in productivity is statistically significantly different from zero with 95 percent confidence.

*Change in productivity is statistically significantly different from zero with 90 percent confidence.

cator variables, $\boldsymbol{\delta}^{\text{school fixed effects}}$ is the vector of school fixed effects, $\mathbf{I}^{\text{year}}$ is a vector of year indicator variables, $\boldsymbol{\delta}^{\text{year}}$ is the vector of year fixed effects, $I^{\text{charter}\geq 6\%}$ is an indicator variable for the district's having at least 6 percent of enrollment in charter schools, and $\delta^{\text{critical charter competition}}$ picks up the effect of facing a critical level of charter competition. Note that the year fixed effects pick up changes over time in the test or in the pressure to perform well on the test. The school fixed effects pick up unobserved characteristics of each school that are stable.

The regression used in the bottom panel of table 8.11 is identical, except for the dependent variable, which is the *difference* between this year's and last year's productivity:

$$(10)\quad \frac{\text{ACH}_{ijt}}{\text{PPExp}_{ijt}} - \frac{\text{Ach}_{ij,t-1}}{\text{PPExp}_{ij,t-1}} = I_{ij}^{\text{school}}\delta^{\text{school fixed effects}} + I_{t}^{\text{year}}\delta^{\text{year fixed effects}} + \delta^{\text{critical charter competition}} I_{jt}^{\text{charter}\geq 6\%} + \varepsilon_{jt} + \varepsilon_{ijt}$$

The estimates in the top panel of table 8.11 indicate that Michigan public schools raised their productivity in response to competition from charter schools. Productivity rose by 1.60 (scale points per thousand dollars spent)

Table 8.12 Effects of Charter School Competition on Michigan Public Schools' Achievement

	4th-Grade Reading Exam	4th-Grade Math Exam	7th-Grade Reading Exam	7th-Grade Math Exam
Dependent Variable (productivity based on exam)				
Difference-in-differences (levels)				
Change in achievement level after district is faced with charter school competition	1.21** (0.65)	1.11* (0.62)	1.37** (0.60)	0.96* (0.48)
Dependent Variable (change in achievement based on exam)				
Detrended difference-in-differences				
Change in achievement trend after district is faced with charter school competition	2.40* (1.37)	2.50** (1.04)	0.25 (0.58)	0.77 (0.69)

Source: Michigan Department of Education (2000a–d).

Notes: Standard errors in parentheses. Regressions include school fixed effects and year fixed effects. Charter schools represent at least 6 percent of enrollment in district. The table is based on regressions of school-level data from 1992–93 to 1999–2000. In the top panel, the dependent variable is a school's achievement—specifically, a school's scale scores on the Michigan Assessment of Educational Progress (MEAP) tests, which are administered to fourth- and seventh-graders. In the bottom panel, the dependent variable is the trend (annual change) in a school's scale scores (this year's achievement minus last year's). The regression includes school indicator variables to pick up characteristics of schools that are constant over the period (location, neighborhood, organization) and year indicator variables that allow for statewide changes from year to year in the test itself or in the pressure to perform on the test. From 1992 to 2000, the means and standard deviation of schools' average scores (weighted by the number of test-takers) were mean of 611, standard deviation of 19 on fourth-grade reading; mean of 528, standard deviation of 16 on fourth-grade math; mean of 600, standard deviation of 17 on fourth-grade reading; mean of 521, standard deviation of 14 on fourth-grade math.

**Change in achievement is statistically significantly different from zero with 95 percent confidence.

*Change in achievement is statistically significantly different from zero with 90 percent confidence.

based on the fourth-grade reading exam, by 1.37 based on the fourth-grade mathematics exam, by 1.87 based on the seventh-grade reading exam, and by 1.53 based on the seventh-grade mathematics exam. All of these estimates are statistically significantly different from zero with a high level of confidence.

Moreover, the bottom panel of table 8.11 shows that charter school competition made Michigan public schools improve their productivity *relative to their own initial trends.* Productivity trends based on fourth-grade tests improve to a degree that is statistically significant. Not surprisingly, given the greater impact of charter competition on lower grades, the seventh-grade results are statistically insignificant.

It is difficult to interpret productivity improvements until we know whether they arise as a result of improvement in achievement or a fall in per-pupil spending, or both phenomena occurring simultaneously. Therefore, table 8.12 examines the effects of charter competition on achievement. Its structure is identical to that of table 8.11, except that it shows results for achievement instead of productivity. That is, it leaves out the changes in productivity that come about as a result of changes in per-pupil spending. Table 8.12 shows that the effect of charter school competition on achievement looks much like the effect on productivity, which suggests that Michigan's regular public schools raised their productivity mainly by raising their achievement for a given level of per-pupil spending, rather than by maintaining a steady level of achievement and cutting their per-pupil spending. For instance, the top panel of table 8.12 shows that fourth-grade reading and mathematics scores were, respectively, 1.21 and 1.11 scale points higher in schools that faced charter competition *after* they began to face competition. Seventh-grade reading and mathematics scores were, respectively, 1.37 and 0.96 scale points higher. Recall that these improvements in scores are not only relative to the schools' own initial performance (the first difference), but also relative to the gains made over the same by schools that did not face charter competition (the difference-in-differences).

The bottom panel of table 8.12, which shows detrended difference-in-differences results, shows how the schools facing charter competition accomplished these achievement gains. For instance, examine the fourth-grade reading and mathematics coefficients, which are statistically significant. (The seventh-grade detrended difference-in-differences coefficients are, as in table 8.11, statistically insignificant.) Schools that faced charter competition raised their annual improvement in achievement by 2.40 scale points a year in fourth-grade reading and 2.50 scale points in mathematics. Recall that this is a change relative to their previous rate of change in achievement, which was actually about 0.4 scale points *lower* on average than that of schools that were never faced with charter competition. In fact, the results give us a picture much like that shown in figure 8.4: The achievement trend of schools that eventually face charter competition is initially lower than that of other schools, but, once charter competition commences, schools that face competition have a higher rate of growth.

The change in achievement for schools subjected to charter competition is statistically significant and positive, but it is not unrealistically large, particularly when one considers that such schools were making up for years of slower achievement growth. Even with mathematics and reading achievement growth that is about 2.5 scale points per year better than that of other schools, a district like Detroit would take approximately two decades to catch up with the achievement of one of its affluent suburbs, like Grosse Point. (Of course, it is possible that, as Detroit caught up, a suburb like Grosse Point would feel competitive pressure to increase its own rate of achievement growth. This would lengthen the catch-up period but further raise Michigan students' scores.)

Overall, the picture that one draws from Michigan is the following. Public schools that were subjected to charter competition raised their productivity and achievement in response, not only exceeding their own previous performance but also improving relative to other Michigan schools not subjected to charter competition. The improvements in productivity and achievement appear to occur once charter competition reaches a critical level that coincides with the enrollment at which charter schools' taking students would be easily discernible (not to be confused with regular fluctuations in enrollment). The increase in productivity and achievement is larger and more precisely estimated in fourth grade, probably because elementary schools faced more competition from charter schools than middle schools did.

8.6.3 The Effect of Charter Schools on Achievement in Arizona Public Schools

Like Michigan, Arizona enacted a charter school law in 1994. Arizona's charter school law is widely regarded as the most favorable to charter schools in the United States, as it allows charter schools to have considerable fiscal and legal autonomy. There are also few constraints on the growth of charter schools in Arizona. As a result, 5.3 percent of Arizona's nonprivate enrollment was in charter schools in 1999–2000. This percentage is the highest of any American state.

In Arizona, state-sponsored charter schools get a fee equal to the state's share of revenue (45 percent of total revenue for a regular public school). District-sponsored charter schools get a fee equal to local per-pupil revenue but are less able to compete with the regular public schools because they must seek renewal of their charters from the very districts with which they compete.

My evaluation of Arizona follows the same strategy as I employ for Michigan, so I will merely highlight a few differences between the Michigan and Arizona situations here. In Arizona, a municipality may contain multiple districts: for instance, a few elementary districts, a middle school district, and a high school district. A local charter school may therefore be competing with multiple districts. Therefore, I associate regular public schools and charter schools with a municipality, not a district. All Arizona fourth- and seventh-graders were required to take the Iowa Test of Basic Skills (ITBS) through 1995–96 and have been required to take the Stanford 9 test since then. The shift in the test does not pose problems for the analysis because both tests offer national percentile rank scores (which have a 0.97 correlation at the school level). Moreover, all of the schools switched tests in the same year, so it is simple to establish each school's prereform trend and postreform trend allowing for a statewide shift in the intercept.[32] I use national percentile rank scores at the school level for the school years from 1992–93 to 1999–2000. I again use 6 percent of enrollment as the critical level at which charter schools are held to be a nonnegligible competitive

32. More precisely, a separate statewide shift is estimated for each percentile rank. The information on Arizona charter schools and all the data on Arizona schools are taken from publications of the Arizona Department of Education (2000a–d).

threat. I use the same critical level as I use for Michigan in order that the two states' results be as comparable as possible. However, a variety of critical levels between 6 percent and 11 percent produce similar results for Arizona.[33]

The right-hand panel of table 8.10 lists the Arizona municipalities that had at least 6 percent of local enrollment in charter schools. Municipalities of all sizes are represented. The list includes some of Arizona's largest cities (Phoenix, Tempe, Scottsdale), some medium-sized cities (Avondale, Flagstaff, Gilbert, Kingman), and thirty smaller municipalities.

As in Michigan, it is important that the Arizona difference-in-differences strategies control for each school's initial conditions. We need to control for schools' unobservable characteristics, particularly because some of those characteristics may actually attract charter competition. Also, it is important that the difference-in-differences strategy generate estimates that control for what was happening to Arizona schools in general over the period. Although Arizona did not experience a major school finance reform, it did have a very activist state department of education that enacted numerous programs (including a school report card program so that parents would be better informed about performance).

Table 8.13 has the same structure as table 8.11: The effect of charter competition on productivity is estimated using difference-in-differences in the top panel and using detrended difference-in-differences in the bottom panel. The equations estimated are (9) and (10), shown above.

The estimates in the top panel of table 8.13 suggest that Arizona public schools raised their productivity in response to competition from charter schools. Productivity rose by 0.55 (national percentile points per thousand dollars spent) based on the fourth-grade reading exam, by 0.70 based on the fourth-grade mathematics exam, by 0.38 based on the seventh-grade reading exam, and by 0.53 based on the seventh-grade mathematics exam. All of these estimates are statistically significantly different from zero with a high level of confidence.

The bottom panel of table 8.13 shows that charter school competition made Arizona public schools improve their productivity *relative to their own initial trends.* Productivity trends based on fourth-grade tests improve to a degree that is statistically significant. The seventh-grade results are statistically insignificant, but this is not surprisingly because charter competition had a greater impact on lower grades.

As noted previously, interpreting productivity gains is hard until we look at one of the components of productivity separately. Table 8.14 shows the effect of charter competition on achievement—that is, it leaves out the changes in productivity that come about as a result of changes in per-pupil

33. These results and descriptive statistics for the Arizona data set are available from the author. Choosing a level much higher than 11 percent makes the results depend unduly on just a few districts, simply because only a few districts ever face more than an 11 percent drawing away of their students.

Table 8.13 Effects of Charter School Competition on Arizona Public Schools' Productivity

	4th-Grade Reading Exam	4th-Grade Math Exam	7th-Grade Reading Exam	7th-Grade Math Exam
Dependent Variable (productivity based on exam)				
Difference-in-differences (levels)				
Change in productivity level after district is faced with charter school competition	0.55** (0.16)	0.70** (0.19)	0.38* (0.21)	0.53** (0.17)
Dependent Variable (change in productivity based on exam)				
Detrended difference-in-differences				
Change in productivity trend after district is faced with charter school competition	0.31* (0.17)	0.28** (0.13)	0.33 (0.22)	0.35 (0.26)

Source: Arizona Department of Education (2000a–d).

Notes: Standard errors in parentheses. Regressions include school fixed effects and year fixed effects. Charter schools represent at least 6 percent of enrollment in district. The table is based on regressions of school-level data from 1992–93 to 1999–2000. In the top panel, the dependent variable is a school's productivity—specifically, a school's national percentile rank (NPR) score divided by its per-pupil spending in thousands of 1999 dollars. In the bottom panel, the dependent variable is the trend (annual change) in a school's productivity (this year's productivity minus last year's). The regression includes school indicator variables to pick up characteristics of schools that are constant over the period (location, neighborhood, organization) and year indicator variables that allow for statewide changes from year to year in the test itself or in the pressure to perform on the test. The inflator for per-pupil spending is the Consumer Price Index. The numerator for productivity is the school's national percentile rank on a nationally normed standardized test (the Iowa Test of Basic Skills or the Stanford 9). See the text for details on the tests.

**Change in productivity is statistically significantly different from zero with 95 percent confidence.

*Change in productivity is statistically significantly different from zero with 90 percent confidence.

spending. The table has the same structure as the previous two tables. It shows that the effect of charter school competition on achievement looks much like the effect on productivity, which suggests that Arizona's regular public schools raised their productivity mainly by raising their achievement for a given level of per-pupil spending, rather than by maintaining a steady level of achievement and cutting their per-pupil spending. The top panel of table 8.14 shows that fourth-grade reading and mathematics scores were, respectively, 2.31 and 2.68 national percentile points higher in schools that faced charter competition *after* they began to face competition. Seventh-grade mathematics scores were 1.59 national percentile points higher. (The estimate for seventh-grade reading is statistically insignificant.) These are important gains, especially when one recalls that these gains are relative not only to the schools' own initial performance (the first difference), but also to the gains made over the same period by schools that did not face charter competition (the difference-in-differences).

The bottom panel of table 8.14 shows the detrended difference-in-differences results, which suggest that schools facing charter competition raised achievement relative to their own previous trends. Such schools

Table 8.14 Effects of Charter School Competition on Arizona Public Schools' Achievement

	4th-Grade Reading Exam	4th-Grade Math Exam	7th-Grade Reading Exam	7th-Grade Math Exam
Dependent Variable (achievement based on exam)				
Difference-in-differences (levels)				
Change in achievement level after district is faced with charter school competition	2.31**	2.68**	1.11	1.59*
	(0.69)	(0.79)	(0.95)	(0.89)
Dependent Variable (change in achievement based on exam)				
Detrended difference-in-differences				
Change in achievement trend after district is faced with charter school competition	1.40*	1.39*	1.48	1.29
	(0.79)	(0.81)	(1.13)	(1.10)

Source: Arizona Department of Education (2000a–d).

Notes: Standard errors in parentheses. Regressions include school fixed effects and year fixed effects. Charter schools represent at least 6 percent of enrollment in district. The table is based on regressions of school level data from 1992–93 to 1999–2000. In the top panel, the dependent variable is a school's achievement—specifically, a school's national percentile rank (NPR) score on a nationally normed standardized test (the Iowa Test of Basic Skills or the Stanford 9). See the text for details on the tests. In the bottom panel, the dependent variable is the trend (annual change) in a school's achievement (this year's achievement minus last year's). The regression includes school indicator variables to pick up characteristics of schools that are constant over the period (location, neighborhood, organization) and year indicator variables that allow for statewide changes from year to year in the test itself or in the pressure to perform on the test.

**Change in achievement is statistically significantly different from zero with 95 percent confidence.

*Change in achievement is statistically significantly different from zero with 90 percent confidence.

raised their annual improvement in achievement by 1.40 national percentile points a year in fourth-grade reading and 1.39 national percentile points in mathematics. Recall that this is a change relative to their previous rate of change in achievement, which was actually about 0.6 national percentile points *lower* on average than that of schools that were never faced with charter competition. Again, the results give us a picture much like that shown in figure 8.4: Schools that eventually face charter competition start with lower rate of growth in achievement but begin to catch up with higher growth rates once charter competition commences.

The improvements in achievement among schools subjected to charter competition are significant, but not unrealistically large. Even if its scores rise about 1.4 national percentile points more each year than do the scores of schools that do not attract competition, the typical Phoenix area school that is now competing with charter schools will take ten years to catch up with top-performing Phoenix area schools.

Overall, the evaluation of Arizona suggests conclusions that are broadly similar to those one draws from the Michigan evaluation. Charter competition focused on public schools that initially had below-average achievement and productivity growth, but charter competition induced public schools to improve their productivity and achievement. The improvements are relative to the schools' own past performance and also relative to gains made, over the same period, by schools that were not subjected to charter competition.

8.6.4 Discussion of the Effects of Recent School Choice Reform

Are the productivity effects of the Milwaukee vouchers, Michigan charter schools, and Arizona charter schools sufficient to make us think that choice could remedy the American school productivity problem? All three forms of choice did boost productivity. If all schools in the United States were to enjoy productivity growth rates like those in Milwaukee's most-treated schools, American schools could return to their 1970–71 productivity levels in under a decade. Of course, we should be cautious about extrapolating from the short voucher and charter school experiences described in this section. On the one hand, the bursts of productivity growth seen in Milwaukee may settle down to a lower level of growth. On the other hand, many of the long-term, general equilibrium effects of choice are not yet in operation.

In order to get a sense of the magnitude of the productivity effects, without having to extrapolate so much, consider the following alternative question. Is it likely that the productivity effects of Milwaukee's voucher program (the "rising tide") are likely to overwhelm the allocation effects for students who experience the worst possible allocation changes in Milwaukee? We can get a sense of the students who are available in Milwaukee to be reallocated if we examine the very high scoring (top decile) and very low scoring (bottom decile) elementary schools in the city. Such schools score about 32 national percentile points apart on the math exam. Thus, a Milwaukee student's worst-case scenario would be to experience a fall of about 32 national percentile points in his or her peer group. Moreover, let us make the extreme assumption that the student is *very* strongly influenced by his or her peers so that the student's scores fall by 32 points. This scenario is truly pessimistic! It is not strictly impossible, but it is so pessimistic that it is barely plausible. Nevertheless, if the student enjoys the achievement growth rates that Milwaukee students in the most-treated schools are enjoying now, he or she will "grow out of" the bad allocation effects within 4.5 years. That is, the student will be better off for having experienced vouchers within five years of the voucher program's affecting his or her schooling.

Many commentators on school choice are obsessed with the possibility that choice schools will "cream-skim" from the public schools. Thus, it seems odd even to raise the possibility of reverse cream-skimming. Nevertheless, given that public school students are positively affected by choice, one might worry that the effects are due to reverse cream-skimming. It is, however, easy to show that the effects of choice on public school students cannot be largely the result of reverse cream-skimming: There are simply too few students changing schools to affect average test scores to the degree they were actually affected. For instance, between 1996–97 and 1999–2000, the Milwaukee public schools lost no more than 498 fourth-graders to voucher schools. (The actual number is smaller because 498 is the total increase in vouchers for fourth-graders, and some of the vouchers went to students who had previously been attending private schools, not the Milwaukee public schools). Witte, Sterr, and Thorn (1995) inform us that disappointed voucher applicants (applicants who lost the lottery and therefore remained in the Milwaukee public schools) scored 5.6 points lower in reading and 10.2 points lower in math than the average Milwaukee student. They also show that voucher applicants performed at about the same level as other low-income Milwaukee students who were eligible for the vouchers. If we assume that the departing voucher students were like the disappointed applicants, then their departure would raise fourth-grade scores in Milwaukee public schools by at most 0.4 points in reading and 0.8 points in math between 1996–97 and 1999–2000. These gains would imply an annual improvement of 0.14 points in reading and 0.26 points in math. Compare such improvements to 1.3 points in reading and 1.8 points in math, which are the actual annual gains of Milwaukee public school students, above and beyond the gains recorded by the control students in non-Milwaukee schools. (The just-quoted numbers can be derived from table 8.9, once you know that there were 2,376 students in schools facing more competition and 4,554 in schools facing less competition.) In short, the change in Milwaukee scores that could plausibly be caused by reverse cream-skimming is an order of magnitude too small to account for the actual change in Milwaukee scores.

For Michigan and Arizona, there are no scores available for disappointed charter applicants, but I have compared the demographics of charter school students and regular public school students in these states elsewhere (Hoxby 2000a). For instance, in Michigan's ten largest districts, some charter schools enroll a higher share of black students, some charter schools enroll a smaller share of black students, and some charter schools enroll a virtually identical share of black students as the regular public schools do. In the ten next largest districts, there is a similar lack of pattern. In Arizona, charter schools' shares of Hispanic students typically differ by only few percent from those of their municipalities. Moreover, there is no consistent pattern to the differences that do exist. In short, demographic data suggest that cream-skimming and reverse cream-skimming are not important phenomena in Michigan and Arizona.

8.7 Conclusions

In this paper, I have presented evidence that suggests that we should care deeply about the productivity effects of school choice, not only because they potentially relieve tensions generated by the allocation effects of choice but also because American schools are in a productivity crisis. Policies that boost American schools' productivity are sorely needed, if only to return the schools to their 1970 productivity levels.

I have also explained how schools that face choice-driven incentives can be induced to raise their productivity. I presented models of for-profit choice schools, nonprofit choice schools, and even regular public schools that just face interdistrict choice.

In section 8.5 of the paper, I show evidence that traditional forms of choice raise school productivity. I present results for traditional forms of choice because they can have long-term, general equilibrium effects of productivity, such as may arise when schools enter or exit or when a different reward system draws better individuals into teaching. If all schools in the United States experienced high levels of the traditional forms of choice, school productivity might be as much as 28 percent higher than it is today.

In section 8.6 of the paper, I present evidence on three recent choice reforms: vouchers in Milwaukee, charter schools in Michigan, and charter schools in Arizona. In each case, I find that regular public schools boosted their productivity when exposed to competition. In fact, the regular public schools responded to competitive threats that were surprisingly small. In each case, the regular public schools increased productivity by raising achievement, not by lowering spending while maintaining achievement. This achievement-oriented response may, of course, be related to the nature of the actual reforms. One can summarize the productivity effects of a reform like Milwaukee's voucher program by noting that a student would have better achievement in five years under the voucher program *even* if his peer group plunged by the maximum amount possible in Milwaukee and his achievement fell one-for-one with that of his peer group.

Of course, one must be cautious about extrapolating unduly from recent reforms or traditional forms of school of choice that only partially mimic choice reforms. Nevertheless, it seems safe to conclude that analyses that ignore the productivity effects of choice are likely to be misleading. Improvements in productivity may be *the* key effects of choice.

References

Arizona Department of Education. 2000a. *Arizona educational directory.* Phoenix, Ariz.: Arizona Department of Education. Electronic files.

———. 2000b. *Arizona pupil achievement testing, statewide report.* 1988–95 editions. Phoenix, Ariz.: Arizona Department of Education.

———. 2000c. *Average daily membership and average daily attendance reports,* 1988–2000 editions. Phoenix, Ariz.: Arizona Department of Education.

———. 2000d. *School report card program,* 1996–2000 editions. Phoenix, Ariz.: Arizona Department of Education. Electronic PDF files and spreadsheet files.

Cutler, David M., and Louise Sheiner. 1998. Managed care and the growth of medical expenditures. In *Frontiers in health policy research,* vol. 1, ed. Alan M. Garber, 77–115. Cambridge: MIT Press.
Downes, Thomas, and David Schoeman. 1998. School finance reform and private school enrollment: Evidence from California. *Journal of Urban Economics* 43: 418–43.
Economic report of the President. 2000. Washington, D.C.: United States Government Printing Office.
Glaeser, Edward, and Andrei Shleifer. Forthcoming. Not-for-profit entrepreneurs. *Journal of Public Economics.*
Hoxby, Caroline M. 1999a. The effects of school choice on curriculum and atmosphere. In *Earning and learning: How schools matter,* ed. Susan Mayer and Paul Peterson, 281–316. Washington, D.C.: Brookings Institution.
———. 1999b. The productivity of schools and other local public goods producers. *Journal of Public Economics* 74 (1): 1–30.
———. 2000a. Do private schools provide competition for public schools? NBER Working Paper no. 4978 (revised). Cambridge, Mass.: National Bureau of Economic Research, August.
———. 2000b. Does competition among public schools benefit students and taxpayers? *American Economic Review* 90 (5): 1209–38.
———. 2001. Ideal vouchers. Harvard University, Department of Economics. Manuscript, January.
———. 2002a. How school choice affects public school students' achievement. In *Choice with equity,* ed. Paul Hill, 141–78. Stanford: Hoover Institution Press.
———. 2002b. Would school choice change the teaching profession? *Journal of Human Resources* 37 (4): 846–91.
Jorgenson, Dale W. 1984. The contribution of education to U.S. economic growth, 1948–73. In *Education and economic productivity*, ed. Edwin Dean, 95–162. Cambridge, Mass.: Harper and Row, Ballinger.
Jorgenson, Dale W., and Barbara M. Fraumeni. 1989. The accumulation of human and nonhuman capital, 1948–84. In *The measurement of saving, investment, and wealth,* ed. Robert E. Lipsey and Helen Stone Tice, 227–82. Chicago: University of Chicago Press.
———. 1992. Investment in education and U.S. economic growth. *Scandinavian Journal of Economics* 94 (supplement): S51–S70.
Keesing, Donald B. 1966. Labor skills and comparative advantage in international economics: Progress and transfer of technical knowledge. *American Economic Review* 56 (March): 249–58.
Krueger, Anne O. 1968. Factor endowments and per capita income differences among countries. *Economic Journal* 78 (September): 641–59.
Leontief, Wassily. 1956. Factor proportions and the structure of American trade: Further theoretical and empirical analysis. *Review of Economics and Statistics* 38 (November): 386–407.
Michel, Allen, and Israel Shaked. 1987. Trucking deregulation and motor-carrier performance: The shareholders' perspective. *Financial Review* 22 (2): 295–311.
Michigan Department of Education. 2000a. *Directory of Michigan public school academies.* Lansing, Mich.: Michigan Department of Education.
———. 2000b. *K-12 database.* Lansing, Mich.: Michigan Department of Education. Electronic file.
———. 2000c. *Michigan educational assessment program and high school test results.* Lansing, Mich.: Michigan Department of Education. Electronic file.
———. 2000d. *School code master.* Lansing, Mich.: Michigan Department of Education. Electronic file.

Moulton, Brent R. 1986. Random group effects and the precision of regression estimates. *Journal of Econometrics* 32 (3): 385–97.
Rees, Nina Shokraii. 2000. *School choice 2000: What's happening in the states.* Washington, D.C.: Heritage Foundation.
Rose, Nancy L. 1987. Labor rent sharing and regulation: Evidence from the trucking industry. *Journal of Political Economy* 95 (6): 1146–78.
Traynor, Thomas L., and Patrick S. McCarthy. 1991. Trucking deregulation and highway safety: The effect of the 1980 Motor Carrier Act. *Journal of Regulatory Economics* 3 (4): 339–48.
U.S. Department of Commerce, Bureau of the Census. 1976. *Statistical abstract of the United States.* Washington, D.C.: Government Printing Office.
———. 1983. *Statistical abstract of the United States.* Washington, D.C.: Government Printing Office.
———. 1994. *County and city data book.* 1993 edition. Washington, D.C.: Bureau of the Census. Electronic file.
———. 1995. *Statistical abstract of the United States.* Washington, D.C.: Government Printing Office.
———. 1999. *Statistical abstract of the United States.* Washington, D.C.: Government Printing Office.
U.S. Department of Education, National Center for Education Statistics. 1980–2000. *Digest of education statistics.* 1979–99 annual editions. Washington, D.C.: U.S. Government Printing Office.
———. 1993. *Common core of data.* 1990 edition. Washington, D.C.: National Center for Education Statistics. Computer file.
———. 1994a. *National education longitudinal study, 1988: Third follow-up.* Washington, D.C.: National Center for Education Statistics. Restricted access computer file.
———. 1994b. *School district data book: 1990 census school district special tabulation.* Washington, D.C.: National Center for Education Statistics. Computer file.
———. 1995. *Common core of data: School years 1987–88 through 1992–93.* Washington, D.C.: National Center for Education Statistics. Computer file.
———. 1999. *National assessment of educational progress restricted use database.* 1990–99 editions. Washington, D.C.: National Center for Education Statistics. Electronic files.
———. 2000. *National assessment of educational progress 1999 long-term trend summary data tables.* Washington, D.C.: National Center for Education Statistics. Electronic files.
U.S. Geological Survey. 1994. *Geographic names information system.* Washington, D.C.: U.S. Geological Survey. Electronic file.
Wisconsin Department of Public Instruction. 2000a. *Knowledge and concepts examinations: Test results.* Madison, Wis.: Wisconsin Department of Public Instruction. Electronic file.
———. 2000b. *Milwaukee parental school choice program: Facts and figures.* Madison, Wis.: Wisconsin Department of Public Instruction.
———. 2000c. *Reading comprehension test results.* Madison, Wis.: Wisconsin Department of Public Instruction. Electronic file.
———. 2000d. *School finance data.* Madison, Wis.: Wisconsin Department of Public Instruction. Electronic file.
———. 2000e. *School performance report.* Madison, Wis.: Wisconsin Department of Public Instruction. Electronic file.
Witte, John F., Troy D. Sterr, and Christopher A. Thorn. 1995. Fifth-year report: Milwaukee parental choice program. Madison, Wis.: Department of Political Science and the Robert La Follette Institute of Public Affairs.