

chapter 34

Schooling I: The Performance of Schools and Colleges

A school can be defined as an institution specializing in the production of training, as distinct from a firm that offers training in conjunction with the production of goods.

—Becker (1962)¹

It often seems that every other day we hear either a member of the media or some disgruntled politician lamenting the large class sizes, poorly equipped schools, declining student performance, and both the underpaid and often underqualified teachers that seemingly characterize the U.S. educational system. The knee-jerk response is that more should be spent on public education in general and primary education in particular. Yet this raises a difficult question: what is the “right” amount to spend? Presumably, this question does have an answer but “zero” seems to be the wrong one as does, “the entire GDP of the U.S. economy.”

Fortunately, the methods developed in Chapter 5, which dealt with the economics of human capital, have equipped us with powerful tools that we can use to address this issue in a sensible manner (see, in particular, the methods developed on page 174). To recap, the basic issue involves measuring the rate of return to educational investments and comparing it with the rate of return on other investments (often measured by the market rate of interest). If the former rate exceeds the latter, then this signals greater levels of investment in education are desirable. Similar, but opposite, remarks apply if the latter exceeds the former. The efficient level of investment in education occurs when its rate of return equals the market rate of interest.

While this is all well and fine on paper, it does raise another thorny issue. Just how does one set about measuring the rate of return on educational investments? One approach is to relate dollar expenditures on schooling to a metric of student performance (as measured by, say, their SAT scores or by their future earnings). In fact, precisely such an

LEARNING OBJECTIVES

By reading this chapter, you should be able to:

- Recognize the broad trends in the U.S. educational system over the past 50 years or so and understand the principal findings of the influential Coleman report.
- Understand the principles that underlie an educational production function.
- Explain the principal econometric obstacles that hinder attempts to obtain a faithful measure of the true rate of return to educational investments.
- Explain the principles underlying the econometric methods used to deal with the problem of selectivity bias.

undertaking was attempted over 40 years ago, leading to the publication of the influential *Coleman Report*, whose main properties are described in Definition 34.1.

Since its publication, the Coleman Report has spawned volumes of research that have placed its central findings under the econometric microscope. Broadly speaking, the emergent literature may be classified into one of two distinct categories. One branch has reconfirmed the central tenets of the report: schooling inputs have (at best) modest effects on student performance (see, for example, Betts 1995). The second branch, exemplified by Johnson and Stafford (1973) and Card and Krueger (1992a), has uncovered a link between school inputs and students' subsequent *earnings*, as distinct from their *test scores*.

DEFINITION 34.1: The Coleman Report

The Coleman Report (more formally, the *Equality of Educational Opportunity* report) was mandated by Congress as part of the 1964 Civil Rights Act (see Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld, and York (1966)). Its goal was to study inequality in the provision of public education. In doing so, it provided one of the first systematic attempts to collect school data around the country. In particular, it measured the test scores and socioeconomic characteristics of the students together with the school inputs. The primary—and indeed shocking—conclusion of the report was that school inputs had a nugatory impact on student achievement relative to the students' family background and the demographic makeup of the student body.

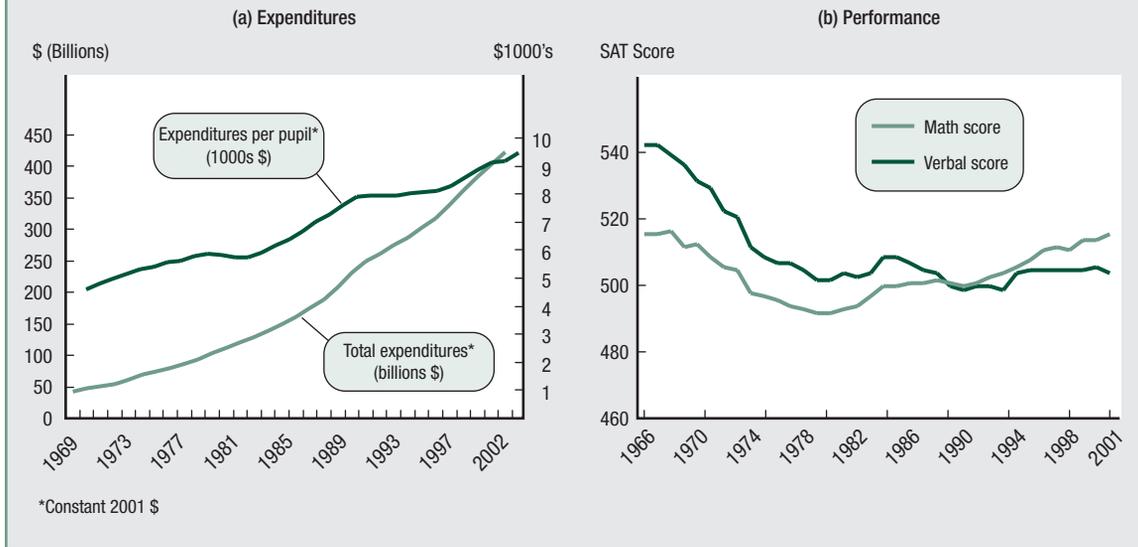
The two issues that are arguably the most fundamental concerns in the economics of education are obtaining a reliable estimate of the rate of return to educational investments and determining both the optimal means of financing and governing the educational system. This chapter focuses on the first of these issues, and Chapter 35 focuses on the second.

34.1 The Evidence

Our Nation is at risk. Our once unchallenged preeminence in commerce, industry, science, and technological innovation is being overtaken by competitors throughout the world. . . . We report to the American people that while we can take justifiable pride in what our schools and colleges have historically accomplished and contributed to the United States and the well-being of its people, the educational foundations of our society are presently being eroded by a rising tide of mediocrity that threatens our very future as a Nation and a people. What was unimaginable a generation ago has begun to occur—others are matching and surpassing our educational attainments.²

Primary and Secondary Schools

As we saw in Chapter 5, total annual expenditures on education in the United States are enormous. Figure 34.1a displays total expenditures on elementary and secondary schools over the period 1969–2000 (measured in billions of constant 2001 dollars).³

FIGURE 34.1 Expenditures and Performance of Primary and Secondary Schools: 1969–2002

Sources: *Digest of Education Statistics*, (2003), U.S. Department of Education. Fig. (a) Tables 29 and 166, and Fig. (b) Table 134. Data available at <http://nces.ed.gov/quicktables> (accessed July 1, 2010).

Part of the increase in expenditures obviously reflects an increase in the sheer number of students attending school. However, *real* expenditures per student have increased dramatically over the period as well. In 2000 expenditures per student were almost twice those of comparable expenditures in 1969.

Expenditures on education, measured as a fraction of U.S. GDP, witnessed striking changes over the past 50 years. For example, the fraction of GDP spent on primary and secondary education increased from 2.3% of GDP in 1949 to 4.3% of GDP in 2001.⁴ Nevertheless, by 1969 expenditure patterns, measured as a fraction of GDP, had all but settled down. In 1969, all educational expenditures, including those on higher education, accounted for 6.9% of GDP. In 2001, this figure had only risen to 7.1% of GDP.

What is the aggregate result of this explosion in educational expenditures? The results are, to say the least, not particularly comforting. Figure 34.1*b* depicts average SAT scores by year between 1969 and 2000 for mathematics and verbal reasoning. Despite the enormous increase in expenditures per student, verbal test scores are lower today than they were in 1969!

Higher Education

As described in Economic Application 34.1, over the past 100 years higher education in the United States has undergone a profound transformation. Perhaps the most notable feature being the astronomical increase in U.S. college and

university enrollments.⁵ Thus, 238,000 students were enrolled in U.S. colleges and universities in 1899. This figure had increased to 1.1 million students by 1930, 8.6 million students by 1970, and in excess of 15.4 million students by 2004.

Expenditures on higher education are enormous and continue to grow rapidly. Measured in constant 1999 dollars, total expenditures on degree granting institutions more than doubled over the last three decades, increasing from \$109 billion in 1970 to \$231 billion in 1999. Finally, the academic labor market also underwent striking changes. In 1970, U.S. colleges employed some 474,000 full- and part-time instructional faculty. Today the figure is over 1.02 million. In 1970, the ratio of full- to part-time faculty was about 3.6:1. By 1999, this figure had decreased to 1.35:1, reflecting the strong growth of nontenured part-time instructional faculty.⁶

ECONOMIC APPLICATION 34.1

The Evolution of Higher Education in the United States

The principal features of today's institutions of higher education arguably stemmed from post-World War II developments, such as the GI Bill and the increase in federal funding for higher education. As Goldin and Katz (1999) observe, these features are

[t]he large average size of its institutions, the coexistence of small liberal arts colleges and large research universities; the substantial share of enrollment in the public sector; a viable and long-lived private sector; professional schools that are typically embedded within universities; and varying levels of per capita funds provided by the states.⁷

Goldin and Katz (1999) argue that the modern form of the American higher education industry took place during the period between 1890 and 1940.

One of the primary impetuses behind the reorganization were the “knowledge shocks,” which took place in the late 19th century, leading to *specialization* in knowledge production. In universities and colleges, this fragmentation was mirrored in the emergence of separate departments of physics, chemistry, economics and so on. ■

34.2 Educational Production

The data on the schooling sector suggest a number of puzzles. The most important one . . . is that the constantly rising costs and ‘quality’ of the inputs of schools appear to be unmatched by improvement in the performance of students. It appears from the aggregate data that there is at best an ambiguous relationship between and at worst a negative relationship between student performance and the inputs supplied by schools.

—Hanushek (1986)⁸

A *production function* is a relationship that determines the feasible *output* that can be produced from a given set of inputs.⁹ An *educational production function* is no different. Broadly speaking, it determines the relationship between the various inputs used in the educational process and the resulting output measured in terms of student performance. It is the primary workhorse in the economics of education and can be used to model the outcome of the schooling process at every level from kindergarten through to graduate school. Unfortunately, however, the precise form of the educational production function is unknown and must be estimated, often using extremely imperfect data.

At its most general level, the educational production function, $F(\cdot)$, takes the form:

$$h_i = F(Q_i, H_i, X_i) \quad (34.1)$$

Where, i indexes some person in a group of N different persons, and the meanings of the other symbols are as follows:

- h_i is list (i.e., a *vector*) describing student i 's performance (e.g., her SAT score, GPA, together with measures of her subsequent earnings possibilities in myriad occupations).
- Q_i is a vector describing the quality of the school environment as it pertains to student i . For example, Q_i might include the student–teacher ratio, the quality and motivation of teachers, expenditures on anything and everything from classrooms to computers and even the abilities of the student's peers.
- H_i is a vector of characteristics specific to the student. The vector H_i might include the student's innate ability, her effort, and her motivation.
- X_i is a vector of givens that exert a systematic effect on the schooling relationship. The vector X_i might include the student's sex, race, parental expenditures on books and on other educational materials, and her parent's educational levels.

From a policy perspective, Q_i is the most pertinent vector because it contains elements that, ultimately, are under the control of the federal, state, or local government. For instance, communities must make choices regarding, say, average class sizes and the level of teacher pay—both of which belong to Q_i . Each of these vectors typically will include some variables that can be directly observed by the econometrician (e.g., the student's SAT or GPA scores) and others that cannot (e.g., the student's personal drive to succeed). Most important, as stressed by, for example, Hanushek (2002) the inputs Q_i , H_i , X_i should be *cumulative* reflecting the student's educational experiences up to the point in time at which her performance is measured. Thus Cuthbert's performance in geometry depends not only on the ability of this year's math teacher but also on the abilities of previous math teachers.¹⁰

With a little imagination, it is possible to make each of the lists just described long enough to fray even the nerves of the most sangfroid policy maker, educator,

or economist. However, given its near \$500 billion importance, it is a field in which abject failure is simply not an option.

Fortunately, following the publication of the Coleman Report, an impressive and sometimes ingenious body of literature has emerged that has sought to cut through the dense thicket of difficulties and to estimate the educational production function described by Equation 34.1.¹¹ Decisive progress has been made that now allows us to say something about the return to educational investments—and (more to the point) something sensible at that.

Estimating the Returns to Investments in Schools

Schools differ dramatically in “quality,” but not because of rudimentary factors that many researchers (and policy makers) have looked to for explanation of these differences. For example, differences in quality do not seem to reflect variations in expenditures, class sizes or other commonly measured attributes of school teachers. Instead, they appear to reflect differences in teacher “skills” that defy detailed description, but that can possibly be directly observed.

—Hanushek (1986)¹²

In applied work, it often seems that, as a general rule, the more important an unknown parameter is for public policy purposes, the more obstacles that stand in the way of obtaining a faithful estimate of its value. Unfortunately, measuring the rate of return to educational investments is no exception. Ignoring, for the moment, the problem of obtaining satisfactory data, which in itself is often a major obstacle, there are three primary difficulties. First, the outcome of the educational process is *multidimensional*. Thus, at one extreme, a student might be taught to just excel at tests, while at the other he might be provided with a fundamental grasp of substantive material. Second, many distinct inputs are used in the educational process. These include items like the quality of the library, the student–teacher ratio, the number of teaching assistants, and the funds that are available for extracurricular activities. Finally, (broadly speaking) there is the problem of *omitted variables*. More specifically, there are crucial variables, such as students’ innate abilities, that economists cannot directly measure or are not included in their data sets.

Together, the first and second problems imply that in reality there are *many distinct rates of return* to education. For example, one rate of return could reflect improvements in SAT scores that result from (say) smaller class sizes. Another could measure the expected increase in each student’s future earnings that results from increased expenditures that make computers more accessible to students in the classroom. However, data limitations imply that the applied economist may possess information only on a few broad aggregates such as (say) total expenditures on education, rather than the desired detailed investment levels broken down into specific categories. The third problem, omitted variables, permeates all of the applied work in this area. Here the key concern is that one or more variables

unobserved by the econometrician (and so omitted from the analysis) are correlated with educational investments *and* student performance. As we shall see, this creates formidable difficulties in accurately measuring their true rate of return.

The Primary Estimation Problem: Data. The first step in this empirical research program is the identification of a few salient variables that are believed to describe the h_i , Q_i , H_i , and X_i vectors mentioned earlier. For instance, h_i is often measured by the student's SAT score; Q_i may be captured by school expenditures and the average class size; H_i is sometimes identified with the student's innate ability. Part of this innate ability can be measured by the econometrician (for instance, the student's IQ). However, as described shortly, the more recent literature addresses the fact that H_i may also include unmeasured aspects of innate ability. Finally, the vector X_i includes many controls, such as the student's race, sex, and family background factors (likewise the educational attainment of the student's parents).

Educational Output. The production technology (Equation 34.1) maps from myriad inputs to student outcomes, h . It is therefore clearly essential to have a satisfactory outcome of student performance, h . In the empirical literature h is often just measured by standardized test score results, such as students' SAT scores. Other measures include graduation rates or each student's subsequent labor-market performance. The problem with using test scores is that one suspects they fail to fully capture the essence of educational output.¹³

A student's subsequent labor-market performance is an alternative and sometimes attractive measure of the outcome of the schooling process; after all, it tells us how much the market values a given skill set. However, one difficulty with this measure is that if an individual's earnings are observed (say) 20 years after he graduated from high school, then it is necessary to sort out the effects of his labor-market experience from the quality of his schooling. Another problem is that because of either unmeasured ability or family background differences, earnings data alone may fail to provide us with an accurate measure of the value added from schooling. For example, a subsequent high wage may simply reflect the fact that good students are sent to good schools, so the causation runs from ability to earnings and school choices.

Educational Inputs. Hanushek (1986) notes,

Family inputs tend to be measured by sociodemographic characteristics of families, such as parental education, income, and family size. Peer inputs, when included, are typically aggregate summaries of the sociodemographic characteristics of other students at the school. School inputs include measures of the teachers (education level, experience, sex, race, and so forth), of the school organization (class size, facilities, administrative expenditures, and so forth), and of district community factors (for example, average expenditure levels).¹⁴

Perhaps the core problem confronted in measuring inputs is ensuring that

The inputs should be related to the students being analyzed; and the educational process should be viewed as cumulative—past inputs have some lasting effect, although their value in explaining output may decline over time.¹⁵

The snag is that, as just noted, many data sets contain information only on current student performance, as measured by test results and contemporaneous inputs (e.g., the current average class size). Yet, as a practical matter, current performance depends on a suitably weighted stream of educational inputs, both past and present.

Omitted Variable Bias

Omitted variable bias is a problem that pervades applied work in this area. Over recent years, however, economists have taken important steps to overcome the problem. In fact, this endeavor is perhaps one of the great success stories of applied labor economics. In this section, we lay out the reasons why omitted variable problems are pervasive and the estimation hazards that arise if they are not addressed.

With this goal in mind, let h represent student quality measured by the student's SAT score, and let q represent educational investments. For the sake of argument, suppose that q is adequately captured by the *teacher–student ratio*.

In the best of all possible worlds, a simple statistical model of the following form:

$$\ln h = \sigma q + \text{controls} + \varepsilon \quad (34.2)$$

would adequately describe the relationship between the explanandum (the thing to be explained), h , and the explanan (the thing that does the explaining), q . Here $\ln h$ is the (natural) logarithm of h , $\sigma \geq 0$ is the key parameter of interest—the rate of return to investments in education, *controls* refers to a laundry list of all variables that exert a systematic influence on h (race, sex, family background, and so on), and ε is a random disturbance term of the sort encountered in Appendix A.¹⁶ If this were all there were to the story, then standard statistical methods (such as OLS) could be used to infer the unknown parameter σ .

The snag is that, as a practical matter, it is likely that some of the controls will not be observed by the econometrician and, worse still, that they will be systematically correlated with h and q . For example, suppose that the dependent variable, $\ln h$, depends on the student's innate ability, $a \in A$, and the teacher–student ratio q according to:¹⁷

$$\ln h = \sigma q + \alpha \cdot a \quad (34.3)$$

where $\sigma \geq 0$ measures the *causal* effect of school investments on a student's performance, and α measures the effect of his innate ability. If both q and $a \in A$ can

Tip!

Notice that an increase in the teacher–student ratio, q , represents a positive investment in education. In particular, a larger value of q corresponds to a smaller class size.

be measured, then standard methods can be employed to infer the unknown parameters σ and α .¹⁸ The trouble is that measuring innate productive ability, $a \in A$, and distinguishing it from h is formidably difficult in practice. Thus did Dougal ace his math test because he's inherently good at mathematics (high a) or because he learned a lot at school (high q)? Suppose that $a \in A$ is not observed by the econometrician. What then? The answer is that, in general, standard statistical methods will lead to biased estimates of σ . To see this, consider the following two hypotheses:

- [H1] Remedial training Suppose that both low-ability and disadvantaged (*low* $a \in A$) students are assigned to small remedial classes to help them catch up.
- [H2] Rewarding Excellence Suppose that schools attempt to nurture high-ability students by assigning them to small classes.

Hypotheses **H1** and **H2** seem eminently plausible and, in fact, depending on the circumstances, innocuous enough. Yet they are potentially devastating for the statistical inference problem because the econometrician may lack information on which one (if any of them) is used by the schools in his or her data set.

To see the damage that is potentially done by this lack of information, suppose that under **H1** the teacher–student ratio, q , and student ability, $a \in A$, are related by:

$$q = \phi - a \quad (34.4)$$

where $\phi > 0$, is an unknown parameter. This simple formulation says that q declines as $a \in A$ rises, indicating that low-ability students are assigned to small classes. Using this equation to substitute out $a \in A$ as:

$$a = \phi - q \quad (34.5)$$

yields:

$$\ln h = \sigma q + \alpha(\phi - q) \quad (34.6)$$

and, after collecting terms,

$$\ln h = (\sigma - \alpha)q + \alpha\phi \quad (34.7)$$

By assumption, the econometrician can see only h and q . Notice, however, that the *estimated* parameter equals $\hat{\sigma} \equiv \sigma - \alpha$ which is smaller than the true parameter σ .¹⁹ From this exercise, we conclude that:

MAJOR RESULT 34.1

Negative Bias

If schools systematically invest more resources in low-ability students, according to **H1** then the estimated return to educational investments is negatively biased.

It is quite conceivable that the true values of σ and α could be such that $\hat{\sigma} \equiv \sigma - \alpha < 0$, implying the *estimated returns* to educational investments are negative—even if the true returns, σ , are positive!²⁰

In contrast, suppose that **H2** is the valid description of the school's behavior: high-ability students are assigned to small classes. In fact, suppose that the school's actions are described by:

$$q = \phi + a \quad (34.8)$$

which says high-ability students are more likely to be assigned to small classes. Solving for $a \in A$ and substituting gives:

$$\ln h = (\sigma + \alpha)q + \alpha\phi \quad (34.9)$$

In this case, the estimated parameter $\hat{\sigma} \equiv \sigma - \alpha$ exceeds the true parameter σ . Here, even if the true value of σ is zero, we might falsely conclude it is positive.

MAJOR RESULT 34.2

Positive Bias

If schools systematically devote more resources to high-ability students, according to **H2**, then the estimated returns to educational investments is positively biased.

Consequently, depending on the unobserved mechanism whereby schools assign students (with unobserved abilities) to different size classes, the estimated parameter $\hat{\sigma}$ may lie below (**H1**) or above (**H2**) the true parameter σ .

Another important source of bias stems from omitted family background factors. For example, parents may differ according to the value they place on their children's education. It is easy to imagine concerned parents providing some education for their children at home and moving close to what they perceive to be high-quality schools. The unfortunate econometrician may see only final student outcomes and public expenditure on schools (family inputs are often notoriously difficult to measure). The unguarded conclusion might then be that investments in education have an extremely high return, even though, in reality, they only encourage attendance by already highly talented (or motivated) students.

As a general proposition, absent additional data, from a statistical perspective the problem of omitted variable bias is fatal. The estimates we obtain from our econometric analysis only tell us (in the case of negative bias), that the rate of return to educational investments is at least $\hat{\sigma} \equiv \sigma - \alpha$. The key to making further progress is acquiring more information. This additional information in conjunction with the appropriate statistical technique can often be used to obtain a more reliable estimate of the true unknown parameter, σ . In section 34.3, we briefly describe some of the recent body of work that has attempted to overcome the problems that result from omitted variables.

The Theory of Education

The previous arguments indicated that there are statistical grounds (under **H1**) for suspecting that the measured returns to schooling investments may be biased downward toward zero. This could help explain the lack of compelling evidence linking schooling investments and student performance noted in, for example, the Coleman Report. In this section, we show that there are theoretical grounds for suspecting that measured public educational investments may have a small—even zero—measured rate of return. Let's begin by describing in overview a recent model by Lazear (2001).

Educational Production. The core insight of Lazear (2001) is recognizing the *public good* dimension of education. If students behaved like passive “dummies,” then one teacher could teach 500 students as easily as he or she could teach 10. The trouble is that, in practice, teaching is subject to congestion effects.

More specifically, as the class size increases, this leads to a greater possibility of disruptive classroom behavior by one or more students. At one extreme, we are all familiar with cases like those in which Betsy, for example, uses her ruler to slap the back of Dougal's unsuspecting pate for sport. However, other less nefarious—but equally disruptive—actions include those in which one student asks a question that everyone else in the class already knows the answer to. From the perspective of these other students, the time the teacher takes to answer the question is simply wasted.

The model proposed by Lazear (2001) captures the effects of classroom disruption. He shows that the optimal class size is greater for better-behaved students. His model predicts, congruent with the evidence, that the level of educational attainment may bear little relation to the observed class size.

A More Detailed Discussion. Model 34.1 is adapted from Lazear (2001).

MODEL 34.1

Lazear's Educational Model

- Consider a school that has Z students and m classes. Let $n \equiv Z/m$ denote the student–teacher ratio.
- The value of human capital created over an **uninterrupted** period of teaching is V . The cost of the teacher and the classroom (measured in terms of wages and rent) is W .
- The probability that during any given period of time a given student **disrupts** the class is $1 - p$.
- The school is a private school that seeks to maximize its profits, π , by choosing the student–teacher ratio, n , and by charging tuition.

- (e) There is free entry of private schools into the market. As schools enter, the value of W is bid up to the point at which every school's profits are zero.

Remark: Under competitive conditions, private schools bid for teachers and pupils. For simplicity, we focus only on the case in which they bid for teachers.

The probability that a given student does *not* disrupt the class during the period is p . This implies that in a class with n students the probability that the class suffers from no disruptions at all is p^n . It is helpful to think of p^n as capturing the average amount of teaching that takes place during the class. For example, if $p = 0.98$ (so students misbehave only 2% of the time), then, in a small class of $n = 25$ students, $p^n = 0.6$, indicating that 40% of the class time is wasted. The value of human capital generated in a class that suffers no interruptions whatsoever is $\$V$. It follows that, given p and n , the amount of human capital generated during the typical class period is just $\$V \times p^n$. For simplicity, assume that Vp^n equals the tuition the school charges each student for admittance.²¹ The school's profits equal the revenues it receives from tuition minus its costs:

$$\pi = Z\{Vp^n - W/n\} \quad (34.10)$$

The optimal class size, n^* , is characterized by the following condition:

$$-V(\ln p)p^{n^*} = W/(n^*)^2 \quad (34.11)$$

The left-hand side is the marginal cost of a larger class size (measured in terms of more disruptive behavior), and the right-hand side is the marginal benefit of a larger class size (measured in terms of reduced average costs per student).²² In equilibrium, the free-entry condition ensures that schools earn zero profits:

$$\pi = Z\{Vp^n - W/n\} = 0 \quad (34.12)$$

Using this condition yields:

$$n(p)^* = -1/\ln p \quad (34.13)$$

where $n(p)^*$ is the optimal class size, written in terms of each student's disruptive-ness, p . Notice that n^* increases with p , indicating that better behaved (high p) students are optimally assigned to larger classes, and, as an immediate corollary, poorly behaved students are assigned to smaller ones. What is more, conditional on the optimal class size, $n(p)^*$, the value of human capital generated per student, denoted $Q(p)$, is given by:

$$Q(p) \equiv Vp^{n(p)^*} = V \exp[-1] \quad (34.14)$$

which is independent of p !

This result is important. It says that the class size optimally adjusts to maintain a constant level of educational output per student by assigning the most disruptive students to the smallest classes. The outcome of this optimizing process is

that (consistent with the evidence) the model predicts there will be no systematic relationship between measured student performance and class size.

Although the model focuses on private profit-maximizing schools, similar results are obtained if it is extended to examine non-profit-maximizing public schools. Lazear extends his model to encompass interesting issues such as differences in teacher ability and the importance of discipline at school.

Crowding Out. In much of the schooling literature, children are often conceived of as little automatons whose correspondingly little heads are filled with either more or less information at the whim of educators. Yet children make choices. In fact, if investment in education makes learning easier for students, then they may optimally respond by studying less!

34.3 Primary and Secondary Schooling

As we saw earlier, the value of U.S. public investments in education are enormous.²³ To make informed choices regarding the proper level of public financing, it is essential that policy makers have an accurate measure of the *rate of return to educational investments*. This information is critical for determining whether, at the margin, additional scarce tax dollars should be invested in education or else used for some other purposes.

Moreover, almost as a by-product, an accurate measure of the rate of return allows communities to evaluate the performance of their schools. This information is obviously essential for identifying those schools whose performance is subpar and formulating appropriate policies to deal with those that are.

Writing in 2002, Hanushek notes that, since the publication of the Coleman Report in 1966, some 276 estimated educational production functions appeared in academic journals. His major findings are reproduced, for convenience, in Table 34.1.

Schooling Input	Number of Estimates	Significant		Insignificant
		Positive	Negative	
Teacher–student ratio	276	14%	14%	72%
Teacher education	170	9%	5%	86%
Teacher experience	206	29%	5%	66%
Expenditures per pupil	163	27%	7%	66%
School facilities	91	9%	5%	86%

Source: Hanushek, 2002, p. 2076

The basic theme that emerges from sifting through the evidence is that the effectiveness of schools appears to bear little relationship to those factors that appear to be important on a priori grounds (in particular, notice the striking results presented in the last column of the table). Hanushek (1986) further adds

The results are startlingly consistent in finding no strong evidence that teacher–student ratios, teacher education, or teacher experience have an expected positive effect on student achievement. According to the available evidence, one cannot be confident that hiring more educated teachers or having smaller classes will improve student performance.²⁴

In view of this evidence, what factors explain the wide differences observed in student outcomes? Hanushek (1986) concludes

[F]amily background is clearly very important in explaining differences in achievement. Virtually, regardless of how measured, more educated and more wealthy parents have children who perform better on average.²⁵

In recent years, Hanushek's (1986) findings have come under close scrutiny. For instance, Hedges, Laine, and Greenwald (1994) perform a meta-analysis of the published results on educational production functions reported in Hanushek (1986). (A *meta-analysis* is a statistical technique that can be used to make inferences by combining information from many different empirical studies such as those just reported.) They conclude,

The data are more consistent with a pattern that includes at least some positive relation between dollars spent on education and output, than with a pattern of no effects or negative effects.²⁶

34.4 Measuring the Returns to School Quality

Many of the early studies surveyed by Hanushek (1986) find little relation between school quality and test score outcomes. However, one of the most significant changes that has taken place in recent applied work is the eschewal of measuring educational outcomes via standardized test score results and, instead, measuring them through labor-market outcomes, such as future earnings.²⁷ Card and Krueger (1994) remark,

To economic analysis, earnings are a natural focus of study because they reflect the market valuation of skills acquired in school. If better schools impart more or better knowledge, this should be reflected in the higher earnings of students. . . . Many economists also question whether standardized test scores are a reliable indicator of student performance. For example, there is evidence that teachers can coach students to perform well on standardized tests, without any lasting effect on their knowledge.²⁸

Most interesting, despite the new emphasis on using earnings data as a measure of performance, some studies now have found an increased effectiveness in the use of test scores as predictors of earnings for recent male cohorts!²⁹ One reason for this closer connection between the two measures is the increased systematic use of test scores by employers in their hiring decisions, implying that students have more to gain from performing well in tests. Next, let's briefly discuss several interesting papers that have sought to measure the returns to school quality.

Card and Krueger

In a very influential study, Card and Krueger (1992a) (CK) use earnings data—as opposed to test scores—to measure the returns to school quality, using the cohorts of men born between 1920 and 1949. They use several statewide measures to measure average school quality, including the student–teacher ratio, average term length, and relative teacher pay. Their results indicate that the rates of return are greater in states with higher quality schools. For instance, decreasing the student–teacher ratio by 5 (students per teacher) is associated with a 0.4% increase in the rate of return to schooling.

Most interesting, they find little evidence, holding constant school quality, that either parental education or parental incomes affect the rate of return. They conclude by remarking:

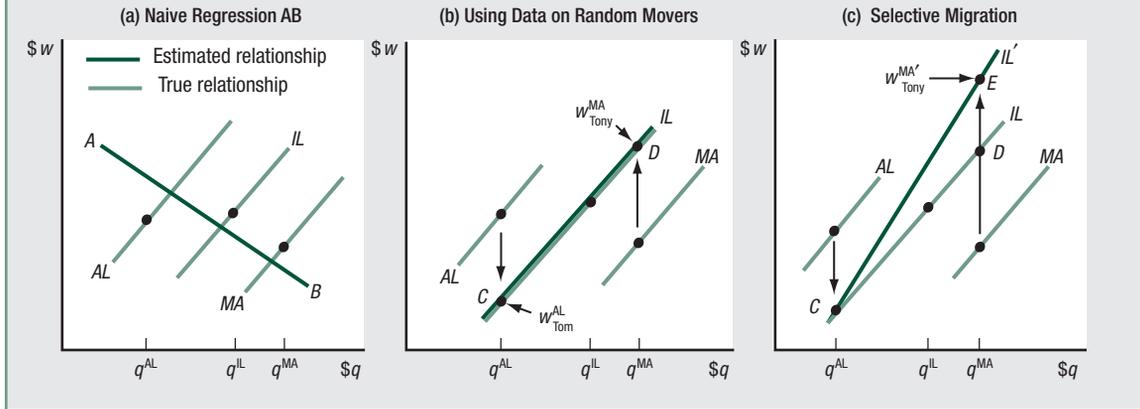
Our findings underscore the paradox . . . : school quality appears to have an important effect on labor market performance but is widely believed to have no impact on standardized achievement tests.³⁰

The Methodology. The empirical framework advanced by CK is interesting because it offers an elegant control for regional variations in labor-market performance. Specifically, their estimation strategy involves using information on workers who were educated in one state, and who subsequently moved and found work in another one. This strategy overcomes potential problems of obtaining spurious results that could arise from systematic differences in regional labor markets (however, as we shall see, it does raise some other problems). For instance, wages and education levels historically have been somewhat lower in the southern states than the northern ones.

To see the difficulties that this might generate, let q denote school quality (measured in, say, dollars per student), and let $w \equiv \ln W$ represent log earnings. Suppose that $w = \sigma q + \theta$, where θ measures state-specific productivity differences, and σ is the *true* (i.e., causal) rate of return to educational investments.

Figure 34.2a plots the (hypothetical) log wage, w ,–school quality, q , relationship for three different states: Alabama (AL), Illinois (IL), and Massachusetts (MA). The slope of each of the lines equals σ , which is the assumed causal rate of return to educational investments, q (in practice, σ might depend on the state

FIGURE 34.2 Estimating the Returns to Educational Investments



as well, but we will ignore this complication). Notice that the relative positions of the three lines differ because of labor-market differences inherent in each of the states. The three points that lie on the lines AL , IL , and MA represent the average wage and average education levels in each state. Notice, in Figure 34.2a, that a naive regression of w on q points to the negative relationship AB !

Figure 34.2b depicts the principles underlying the CK econometric correction. For the sake of argument, consider the line IL and suppose that it corresponds to the city of Chicago. By focusing on movers from the other states to the Chicago labor market, CK can control for unmeasured labor-market effects. To see why, suppose that, as shown in Figure 34.2b, Tom and Tony move from Alabama and Massachusetts, respectively, to Chicago, and that they both possess the average education level corresponding to their home state (q^{AL} and q^{MA}). After they move, their respective earnings are w_{Tom}^{AL} (point C) and w_{Tony}^{MA} (point D). Notice the slope of the estimated line CD equals the true rate of return σ . CK exploit this general idea using data on many workers who move across state lines.

Criticisms. Since the paper's publication, CK's results have attracted considerable attention; nevertheless, they have also been subject to criticism. Two issues are especially troublesome. First, a key assumption of the CK approach is that movers are similar to those who remain in the state in which they acquired their education. If, however, movers differ systematically from the average person in the state, then this will bias the results.

For instance, suppose that Tony moves from MA to Chicago only because his extremely high unmeasured (by the econometrician) ability is better rewarded there. As explained in Figure 34.2c, this sort of systematic behavior leads to an overestimate of the returns to educational investments. Because of Tony's extremely

high ability he earns $w_{Tony}^{MA'}$ on moving to Chicago (notice that his wage exceeds the wage the *average* person from MA would earn were he or she to move to Chicago—which is located at point D). The resulting estimated line, IL' is steeper than the *true* relationship, IL , which implies the estimated return to school quality is biased upward. Heckman, Layne Farrar, and Todd (1996) test for this possibility, discovering that selective migration is indeed a legitimate concern.

A second criticism is that CK use aggregate data to measure the quality of schooling, such as the average class size in the state of Alabama or Massachusetts. The snag is that there is considerable within state variation in school quality. Worse still, this variation may be correlated with the decision to move from one city to another. For example, perhaps there is a tendency for those who attended mediocre schools to remain in the state in which they completed their education. Betts (1995) tests this proposition using panel data that controls for individual differences in school quality.³¹ He concludes,

While there are significant differences between the labor market performance of students who attend different schools, these differences are not significantly related to standard measures of school quality.³² These results accord with the literature on school quality and test scores, as surveyed by Hanushek (1986).³³

The current state of play in the returns to schooling literature involves the attempt by economists to reconcile the differences between the results of studies that use individual and aggregate statewide controls for educational quality. In general, studies that include individual controls find little or no effect of school quality on earnings, but those that use aggregate data uncover positive effects.

For example, Betts (1996) examines whether the difference between the two sets of results stems from the fact that many of the studies that include individual controls look at younger workers with little labor-market experience and those that use aggregate data look at older workers. His general conclusions are negative; there is no age dependence. In contrast, Loeb and Bound (1996) control for the possibility that returns to school investments may have declined through time (CK look at male cohorts between 1920 and 1940). They find substantial effects of school quality on earnings.³⁴

Duflo (2001) attempts to measure the effects of schooling investments on outcomes, using the results of an interesting natural experiment that occurred in Indonesia in the 1970s. Between 1973 and 1979 more than 61,000 primary schools were constructed. This amounts to approximately two new schools per 1,000 children. Enrollment rates increased for children aged between 7 to 12 from 69% in 1973, to 83% in 1978. (This contrasted with a decline in student enrollments before that period). Her estimates indicate that an additional school per 1,000 students led to between 0.12 and 0.19 additional years of education, and increased earnings by between 1.5% and 2.7%.

In another influential study, Card and Krueger (1992*b*) measure the effects of school quality on earnings using data from a natural experiment that resulted from school desegregation in the South. According to their estimates, improvements in the quality of Black schooling led to a 20% narrowing of the Black–White wage differential among southern men. Economic Application 34.2 outlines the main ideas.

ECONOMIC APPLICATION 34.2

School Quality and Black–White Relative Earnings

Card and Krueger (1992*b*) exploit the fact that racially segregated schooling led to enormous disparities in the educational resources available to Black and White children in the South. For instance, North Carolina (NC) was one of the most progressive states with regard to Black-schooling, and South Carolina (SC) was one of the least progressive.³⁵ Thus, in 1920, the South Carolinian Black student–teacher ratio was about 78:1 and in North Carolina it was about 55:1. At that time, the student–teacher ratio for White students was about 35 in both states.

By 1965 student–teacher ratios for both Blacks and Whites had converged to a common level of approximately 28 students per teacher in both states. The authors exploit the fact that the dramatic variation in school resources devoted to Black students in the two neighboring states potentially provides tremendous information about the significance of schooling investments. The narrowing of the relative student–teacher ratio (by 28%) between SC and NC from 1900 to 1940 increased relative Black earnings in SC by 5%.

This finding is consistent with other evidence available in the literature that indicates each 10% reduction in class size is associated with an increase in earnings of 0.4%–1.1% (see, for example, Card and Krueger 1992*a*). ■

Peer Effects

One of the primary motivations underlying the crafting of the original Coleman Report was establishing just how important a student’s peers are in the educational process. Angrist and Lang (2004) study the impact of Boston’s Metropolitan Council for Educational Opportunity (METCO) busing program on student performance.³⁶ This a long-running desegregation program that, by and large, sends Black inner-city children into the more affluent suburbs of Boston. In one year, the METCO program increased the proportion of Black students in one district from 7.5% to over 12.5%. Since, on average, METCO students have much lower test scores than suburban students their influx led to a reduction in average test scores in the more affluent neighborhoods affected by the program.

This raises the specter that the program might adversely affect non-METCO students. Nevertheless, the authors' findings suggest no overall decline in test results for the sample of non-METCO pupils as a whole. There is, however, some evidence that the program adversely affected *non*-METCO minority students (in particular, female third-graders), but these effects appear to have been short-lived.

Hanushek, Kain, Markman, and Rivkin (2003) find that, at all points in the test-score distribution, students benefit from having high-ability classmates. Sacerdote (2001) estimates peer effects using data on Dartmouth College roommates. His study exploits the fact that first-year room assignments are random. This allows him to identify whether peer effects are important determinants of students' GPA scores. The idea is that absent peer effects the GPA of a given Dartmouth student should be independent of the score of his (randomly assigned) roommate. His findings indicate, however, the presence of substantial peer effects. In a recent study, Zimmerman (2003) uses data from Williams College on an individual's grades, SAT scores, and the SAT scores of his (randomly assigned) roommate to estimate peer effects. In this case, he finds only modest peer effects.³⁷

34.5 Selectivity Bias: Recent Advances

Students are assigned to different size classes for a variety of different reasons. As we saw earlier, both low-ability and disadvantaged students could be assigned to small remedial classes to help them catch up (hypothesis **H1**). Alternatively, schools could attempt to nurture high-ability students by assigning them to small classes (hypothesis **H2**). In either case, the result is an empirical nightmare in terms of our attempts to sort out the consequences of something as seemingly simple as measuring the effects of class size on subsequent student performance. Several ingenious studies, however, have exploited exogenous changes in class sizes to shed light on the issue. They are the focus of this section.

Project STAR

Krueger (1999) examines the outcome of the Tennessee Student–Teacher Achievement Ratio (STAR) experiment—project STAR. This was a longitudinal study in which some 11,600 students and their teachers were, beginning in 1986, randomly assigned to one of three groups: (S) small classes (13–17 students per teacher), (R) regular classes (22–25 students per teacher), and (RA) regular/aide classes (which employed a full-time teacher's aide, but had 22–25 students per class).

The experiment took place in 80 different schools throughout the state. Each school was required to have at least one of each class type. The randomization procedure took place within schools. (Thus teachers in Nashville were not sent

packing to Chattanooga and vice versa).³⁸ Krueger’s findings indicate that students score more on tests in smaller classes than in larger ones.

One of the primary concerns in conducting good experimental work is recognizing that (unlike ants or bumblebees), people (here teachers and students) know that they are part of an experimental study. This can result in the following two problems:

- **The Hawthorn Effect** According to the *Hawthorne effect*, teachers respond—in an experimental setting—to being assigned to a small class by working hard.
- **The John Henry Effect** In contrast, according to the *John Henry effect*, teachers respond to the *bad luck* of (here) being assigned to a larger class by exerting more effort than they otherwise would to prove their mettle.

In his analysis of the data, Krueger (1999) addresses both the Hawthorne and John Henry effects but finds little evidence indicating that they affected the principal findings just reported.

Maimonides’ Rule

Economic Application 34.3 provides an overview of a recent ingenious study that was designed to isolate the effects of class size on student performance.

ECONOMIC APPLICATION 34.3

Maimonides’ Rule

The problem of determining optimal class size is a very old one indeed. As Angrist and Lavy (1999) note,

One of the earliest references to this topic is the Babylonian Talmud, completed around the beginning of the sixth century, which discusses rules for the determination of class size and teacher pupil ratios in bible study. “The great twelfth century Rabbinic scholar Maimonides interprets the Talmud’s discussion of class size as follows: ‘twenty-five children may be put in the charge of one teacher. If the number in the class exceeds twenty-five but is not more than forty, he should have an assistant to help with the instruction. If there are more than forty, two teachers must be appointed.’”³⁹

Angrist and Lavy (1999) exploit the fact that Maimonides’ rule has been used in Israel since 1969 to determine the division of classes among successive cohorts enrolled in public schools. The authors cleverly exploit the rule to obtain exogenous (i.e., independent-of-ability) variations in class size, which can be used to estimate the effects of class size on scholastic achievement.

The basic idea underlying their approach is simple but innovative. According to Maimonides' rule of 40, class size increases one for one with student numbers until 40 are enrolled. However, once there are 41 students the rule comes into effect and the class splits, giving (on average) $20.5 = 41/2$ students per class. This gives a *sawtooth* pattern of average class size when plotted against enrollments. (Readers should consult Figure 1 of their paper to see the striking sawtooth pattern for themselves.) The authors use this exogenous variation to estimate (using instrumental variable techniques) the effects of class size on student test performance.

The most obvious statistical danger in measuring the effects of class size on student performance is that low caliber students or those from disadvantaged backgrounds are often assigned to small remedial classes, in the manner described by **H1** on page 9. This can create a spurious correlation between small classes and low test scores. The procedure used by the authors is designed to ameliorate these problems by providing an exogenous variation in class size that is independent of (unmeasured) student ability. More specifically, according to Maimonides' rule, students are assigned to small classes of (say) 25 students because 50 students are enrolled in the school, and not because the students have low abilities. (Recall from the arguments in Appendix A that the two key conditions for a valid instrument are the discovery of a variable that is (IV_i) correlated with the variable of interest and (IV_{ii}) that is not correlated with any omitted variables.) The authors conclude,

The estimates show that reducing class size induces a significant and substantial increase in test scores for fourth and fifth graders, although not for third graders.⁴⁰ ■

Evidence from South Africa

One of the key problems in inferring the effects of class size on student performance is the potential endogeneity of the inputs. In particular, those parents who care about a good education may move closer to good schools to ensure their children are admitted to them.⁴¹ They may also devote effort at home to ensure that their children perform well at school. Although well intended, caring for children is reprehensible on statistical grounds. The reason is that it then becomes extremely difficult for us to ascertain whether any observed relationship between school resources and student outcomes reflects a causal link from the former to the latter, or whether it is due to (unobserved) parental tastes for the quality of their children's education.

Case and Deaton (1999) use a South African data set that ameliorates many of these difficulties. During the apartheid era:

- Blacks were severely limited in their residential choices (they could not move closer to good schools).

- Blacks had no say over school funding (they could not become politically active and vote to raise school expenditures).
- Funding disparities engendered enormous variation in class sizes, which ranged from between 20 and 80 students per class.

The authors note,

Our empirical analysis shows marked effects of school quality as measured by pupil–teacher ratios, on outcomes for Black children. Controlling for household background variables—which themselves have powerful effects on outcomes, but have no effect on pupil–teacher ratios—we find strong and significant effects of pupil–teacher ratios on enrollment, on educational achievement, and on test scores for numeracy.⁴²

For example, they find that reducing the average Black class size from 40 to 30 students would have increased educational attainment levels by about 0.52 years.

Inferring School Quality Using House Price Data

As stressed earlier in this chapter, a precondition for better understanding the prudence of alternative school reform proposals is measuring the value of higher-quality schools. Nevertheless, directly measuring the relationship between school quality and student outcomes is quite challenging. Urban economists, however, have recognized for many years that variations in house prices can provide a useful measure of the value of local amenities.⁴³ Indeed, the importance of proximity to a good school apparently is well known to both realtors and parents. Writing in 1996 Felicia Paik remarks:

Last April, Anthony Ackerman and his wife . . . purchased a new home in Houston. The main motivation for the move: to live in the Spring Branch school district, which produces college entrance test scores that rank in the top 5% in the country.

The Ackermans are really planning ahead; their daughter just turned one. “I’d say we bought this house at an \$80,000 premium if you compared it to a similar house in the Houston city school district,” says Mr. Ackerman, who paid \$190,000 for a 2,800-square-foot, four-bedroom home.⁴⁴

In an ingenious study, Black (1999) employs variations in house prices to indirectly impute the value that parents place on school quality. The key innovation in her study is that she gathers data on properties that are located on school attendance boundaries within given school districts. This implies that houses differ only with respect to the elementary school the child attends. In particular, given their close proximity to each other, the houses in her sample are similar in quality and other local amenities. Moreover, the fact that they belong to the same school district eliminates variation in tax rates and total school spending. Black finds that

parents are willing to pay 2.5% more for their house for each 5% increase in average test scores. Evaluated at the mean of her sample, this amounted to about \$4K (in 1993–1994 prices) for a 5% increase in test scores.

Model 34.2 is designed to help flesh out the links just described.⁴⁵

MODEL 34.2

A Model of House Pricing and School Quality

- A total of N identical families compete to live in $n = 1,000 < N$ identical houses located along a street that is 4 miles long. Each family has one and only one child.
- The (common) lifetime discounted value of owning a house is \$100K.
- At each end of the street there is a school labeled A and B respectively. Districting rules stipulate that school A serves families located in the first half of the street (house numbers 1, 2, 3, . . . , 500) and school B serves those in the second half (house numbers 501, 502, . . . , 1,000).
- School A is of a higher quality than school B . All parents value sending their children to A at \$10K and to B at \$2K.
- The housing market is perfectly competitive.

Remark: In the interests of simplicity, let's assume there are no property taxes and families do not care about the distance, d , to the school.

Our goal is to find out where people live and to determine the equilibrium values of their houses. The assumptions that people and houses are identical and the housing market is competitive are great simplifications. Most significant, they imply that, in equilibrium, the house price equals each buyer's reservation value, which is defined as its maximum willingness to pay for the house. To see why this is so, suppose that Dougal's current bid for a house that is served by school A is $p^A = \$90K < \$110K$. Then, under these circumstances, Betsy can offer \$100K (which beats Dougal's offer) and can accrue a surplus of $\$10K = \$110K - \$100K$ for herself. Yet, there is nothing to stop Norburt from acing Betsy's offer by proposing $p^A = \$105K$. This process continues ad infinitum until $p^A = \$110K$.

It follows that the market price for those houses indexed 1, 2, . . . , 500 is predicted to be $p^A = \$110K$, and $p^B = \$102K$ for houses 501, 502, . . . , 1,000. From our current perspective, the most salient feature of the equilibrium is that even though the two houses numbered 500 and 501 are next door to each other, there is a \$8K price difference between them! This difference reflects precisely the difference in the qualities of the two schools. In essence, Black (1999) is able to identify such a difference and use it to impute each parent's willingness to pay for incremental improvements in school quality.

Hoxby

Hoxby (2000) explores the effects of class size on student performance using U.S. data. She exploits changes in class size that result from exogenous-to-the-school-district variations in student populations.

The idea is that the number of students in each class is systematically related to the number of school districts within a given area. In turn, the number of school districts depends on the local geography of the area. Loosely speaking, for historical reasons, those areas in which walking from one point to another is the most arduous (because of natural barriers, such as streams or rivers) have the greatest number of districts (to ensure each student faced a reasonable walk to school). Hoxby carefully controls for this natural variation using extremely detailed geographic data. However, her findings indicate that changes in class size have a nugatory effect on student outcomes.

34.6 Higher Education

At a most basic level, one of the most fundamental questions in the area of higher education is,⁴⁶ Does the quality of the college that students attend influence their subsequent earnings? Providing a satisfactory answer to this question is important along several different dimensions.⁴⁷ The answer is important to policy makers, since they must decide whether to use scarce tax dollars to finance investments in higher education or to use them to foot the bill for myriad other competing public programs. Obviously, parents and students are also concerned with evaluating whether they are obtaining value for money from their educational investments. This concern is particularly relevant given the dramatic, and often painful, tuition hikes witnessed recently in the United States. Writing in 2003, Dobbs remarks,⁴⁸

Tuition costs at public colleges rose more rapidly last year [2002] than at any time over the past three decades, according to a report released yesterday. After adjusting for inflation, costs were up 13 percent for the year and 47 percent for the past decade. The annual report by the College Board, which collects data from more than 4,000 institutions, said tuition and fees also rose substantially last year at private colleges and universities, but at a slower rate than in the public sector. If room and board costs are included, the average student now pays \$26,854 a year to attend a private university, and \$10,636 to attend a public university in his or her own state.

Although admittedly large, the bitter pill of tuition hikes would perhaps be easier to swallow if they led to a commensurate, substantial, and demonstrable, increase in students' subsequent labor-market earnings.⁴⁹ As discussed next, although there are a few exceptions, a large and growing body of evidence points to a significant (positive) effect of college quality on an individual's subsequent earnings.⁵⁰

The Returns to Attending a More Selective College

The C student from Princeton earns more than the A student from Podunk not mainly because he has the prestige of a Princeton degree, but merely because he is abler. The golden touch is possessed not by the Ivy League College, but by its students.⁵¹

The basic approach to measuring the earnings effects of college quality involves estimating an equation of the form:

$$w = \beta_0 + \beta_1 X + \beta Q + \varepsilon \quad (34.16)$$

where Q is the quality of the college attended by a student (measured in, say, dollars of expenditure per student); $w \equiv \ln W$ the (natural) logarithm of the wage, W ; X represents individual characteristics (age, sex, family background and so on); ε is an error term; and β_0, β_1, β are parameters to be estimated.⁵²

The essential goal of the econometric exercise is to obtain an unbiased estimate of β —the rate of return to Q —given data on w , X , and Q . However, there are several obstacles that stand in the way of obtaining a reliable estimate of β . One of the most pernicious of them is selectivity bias.⁵³

The Empirical Evidence

Brewer, Eide, and Ehrenberg (1999) use sophisticated estimation methods to control for selectivity bias.⁵⁴ Their evidence suggests there is a substantial return from attending an elite college relative to attending a bottom-ranked public university. In one set of estimates—that controls for selectivity bias—the return from attending an elite college is 21.4% in increased annual earnings. Using standard OLS methods, the premium is estimated to be a gigantic 30.3%. The authors also find that the return to attending an elite private college has increased through time.

In an important study, Loury and Garman (1995) estimate a model of the effects of college years, selectivity standards, college major, and college performance on subsequent earnings.⁵⁵ They use data drawn from the National Longitudinal Study (NLS) of the high school class of 1972.⁵⁶ Their evidence suggests that college GPA and the choice of major both have large and significant effects on earnings. For instance, White students majoring in engineering or the physical sciences earned approximately 25% more than the average (White) student.⁵⁷ The figure is even more impressive for Black engineering majors and Black physical science students, as they earn some 32% more than the average Black student. Student performance is also important. According to their estimates, a one point increase in the GPA raises earnings by 9.5% for White and 25% for Black students, respectively.

Kermit, Black, and Smith (1995) use data from the National Longitudinal Survey of Youth (NLSY). They also control for unmeasured (by the econometrician)

student ability. Their estimates indicate that several dimensions of college quality positively affect young men's wages and sorting by more able students into better colleges has only a modest mitigating impact on the effect of college quality. They divide college quality into 5 distinct classes, ranging from *best* to *worst* and remark,

The quality effects are large. According to these estimates, men who attend colleges in the top fifth of the quality distribution earn wages about 18 percent higher than otherwise identical men who attend college in the bottom fifth of the quality distribution.⁵⁸

As in Loury and Garman (1995), the authors note that, in their sample, Blacks enjoy a substantial return from attending elite colleges.

Behrman, Rosenzweig, and Taubman (1996) study the impact of college quality on women's earnings, using a sample that includes identical and nonidentical twins born in Minnesota.⁵⁹ By using data on twins, the authors can isolate college-quality effects by controlling for (1) family background, (2) (unmeasured) ability, and (3) prior schooling. Their preferred estimates indicate that—controlling for family effects and individual abilities—Ph.D.-granting institutions, with modest enrollments and well-paid senior faculty, produce students who earn significantly more later in life. Their results suggest that where one attends college is important for one's future earnings. Indeed, the authors estimate that relative to no college, there is a 20.3% gain from a large public college (Mankato State), a 31.7% gain from a large public research university (University of Michigan at Ann Arbor), a 40.3% gain from an elite private teaching university (Wellesley), and 56.6% from an elite research university (University of Pennsylvania).

Hoxby (1998) estimates the returns to attending a more selective college, carefully controlling for student aptitude. Her evidence indicates that “[p]eople who invest in education at a more selective college generally earn back their investments several times over during their careers.”⁶⁰ In an interesting study, Dale and Krueger (2002) cleverly use elite-college rejects to overcome difficult econometric issues. They find a nugatory return from an elite college education for all but those from the poorest backgrounds.

Tuition Subsidies. Dynarski (2002) examines the effects of subsidizing the costs of education. In general, colleges and universities can be thought of as posting a sticker price—currently about \$30K in elite private schools. They then offer price reductions (i.e., scholarships) to many different groups, such as those from disadvantaged backgrounds and those with high academic merit. The government also subsidizes student participation in higher education by offering cheap loans and targeting specific groups through, for example, Pell grants. The effect of these programs is to render the task of evaluating the consequences of changes in the price of college on enrollments and student quality an especially challenging one. After

controlling for such factors, Dyranski estimates that each \$1K of tuition subsidy increases college enrollments by about 4%.⁶¹

Two- and Four-Year Colleges. Community colleges are an important component of U.S. postsecondary education. They enroll over half of first-time first-year students.⁶² Perhaps even more important, they enroll a disproportionate fraction of students from low-income and minority backgrounds. It is precisely this group of students whose behavior may be influenced by federal and state education policies, such as the provision of Pell grants and subsidized student loans. Although (annual) tuition costs at 2-year community colleges are a little cheaper than those at 4-year colleges, the opportunity cost—measured in terms of foregone earnings—of attending either type of institution is the same. Moreover, the opportunity cost dwarfs the tuition cost.

As a result of the opportunity-cost effect, the true cost of a year of full-time study is similar at both types of institutions. Indeed, according to Kane and Rouse (1995),

[I]n 1975 the average annual earnings of a male, 18–24 year old high school-graduate working full year and full time was \$20,845 (1991 dollars). In contrast, the average tuition . . . at public two-year institutions was \$701, and the average tuition at public four-year institutions was \$1,172.⁶³

Kane and Rouse (1995) find that the evidence points to similar annual rates of return at each type of institution that lie in the range 4%–6% per 30 credits (two semesters). Leigh and Gill (1997) report similar rates of return for returning adults. LaLonde (1995) also shows that subject matter is an important determinant of the rate of return, and there are substantive returns for pursuing technical courses, such as mathematics.⁶⁴

SUMMARY

- Currently the United States spends in excess of \$450 billion each year on primary and secondary education. Expenditures on higher education exceed \$250 billion.
- Since 1960, verbal SAT scores have witnessed a secular decline in the United States.
- The primary finding of the 1966 Coleman Report was that school inputs have a nugatory impact on student achievement relative to the importance of the student's family background and his peers.
- The educational production function is the primary workhorse used by economists to assess the effects of educational policies on educational outcomes.
- The rate of return to schooling equals the percentage change in some outcome (e.g., average SAT scores) resulting from a one dollar increase in educational investments.
- An accurate determination of the rate of return is critical to policy makers in their attempts to formulate prudent educational policies.

Nevertheless, a host of statistical (and conceptual) difficulties hinder attempts to obtain a satisfactory measure. One of the primary statistical problems results from omitted variable bias in general, and the inability to ascertain students' innate abilities in particular.

- There is relatively little work on the theory of education. A notable exception is Lazear (2001), who captures the effects of disruptive behavior on the learning environment. His principal finding is that class sizes may optimally adjust in a manner that obscures observing a relationship between class size and performance.

NOTES

1. Becker (1962), p. 25.
2. *A Nation at Risk* (1983, p. 1). The official title is *National Commission on Excellence in Education* (1983), and it is available at www2.ed.gov/pubs/NatAtRisk/index.html (accessed July 1, 2010).
3. Goldin (1999) provides an excellent overview of the expansion of secondary education in the United States.
4. Source: U.S. Department of Education Digest of Education Statistics, November 2003, Table 29. Available at <http://nces.ed.gov/Programs/digest> (accessed July 10, 2010).
5. Clotfelter (1999) provides an excellent and highly accessible overview of the main issues. See Winston (1999) for a discussion of the business aspects of higher education.
6. Sources: *Digest of Education Statistics* (1999, Table 3; 2002, Tables 227, 344, and 345). Available at <http://nces.ed.gov/Programs/digest> (accessed July 10, 2010).
7. Goldin and Katz (1999), p. 37.
8. Hanushek (1986), p. 1148.
9. Production functions are discussed in detail in Chapter 3.
10. Many data sets are cross-sectional, embodying information only on current performance and current inputs (i.e., not the cumulative inputs up to that point in time). For this reason, the educational production function is often written in the value-added form: $\Delta h_i = F(Q_i, H_i, X_i)$, where Δh_i is the improvement (value added) in student performance (measured over, say, a 2-year period), and all of the inputs are measured contemporaneously. See Summers and Wolfe (1977) for one of the first estimates of a value-added production function.
11. Early studies include Taubman and Wales (1974), Wachtel (1976), Ribich and Murphy (1975), and Johnson and Stafford (1973).
12. Hanushek (1986), p. 1142.
13. The outcome of the schooling process varies along many different dimensions. Some of these outcomes, such as creativity, may be very poorly mirrored by standardized test scores. Hanushek (1986, p. 1153) remarks, "One rather commonly held presumption is that better educated individuals are able to perform more complicated tasks or are able to adapt to changing conditions and tasks." Also see Heckman and Vytlačil (2001) and Heckman and Rubinstein (2001) for an analysis of the importance of noncognitive skills. Most interesting, Jacob and Levitt (2003) study how high-powered incentives can induce teachers to cheat. Their estimates indicate that such cheating occurs in a minimum of 4%–5% of classrooms annually.
14. Hanushek (1986), p. 1115.
15. Hanushek (1986), p. 1155.
16. The parameter σ measures (after multiplication by 100), the percentage change in the test score h that results from spending \$1 on increasing the teacher–student ratio q .
17. The notation $a \in A$ says that *ability, a, belongs to the set of abilities, A*. It is used to avoid confusing

- ability, a , with the indefinite article a . Also, in the interests of simplicity, we ignore the random disturbance, ε , and the other controls.
18. As we saw earlier, standard methods can also be used even if $a \in A$ is unobservable, provided that $a \in A$ and q are not correlated. Yet, there's the rub! In practice, both $a \in A$ and q are likely to be correlated for the reasons described shortly.
 19. Obviously, the econometrician sees only the combined difference $\hat{\sigma} \equiv \sigma - \alpha$, and not the individual components σ and α .
 20. See, for example, Iacovou (2003). Controlling for the endogeneity of class size, she shows that class size is strongly related to reading test-score results.
 21. Suppose that $Vp^n = \$10K$. No parent would be willing to pay the school more than \$10K in tuition because the family then makes a loss on the educational investment. Under competitive conditions, no family can pay the school less than \$10K because the school can always deny the student entry.
 22. To see this, note $p^n \equiv \exp[\ln p \cdot n]$. Differentiating with respect to n gives $\ln p \cdot \exp[\ln p \cdot n] \equiv \ln p \cdot p^n$ as claimed.
 23. Several excellent surveys exist in this area. Hanushek (1986) is a classic. Hanushek (1996) updates his 1986 report and offers a highly accessible account of the literature. Hanushek (2002) offers a splendid summary of the economics of education in general. Card and Krueger (1996) provide an excellent introductory account of more recent estimates that use earnings—as opposed to test score data—to measure student outcomes. See also Burtless (1996) and Blau (1996) for useful surveys of the major issues.
 24. The one ray of hope, as reported in Hanushek (1986, Table 8) is that teacher experience exerts a positive (and notably statistically significant) effect on student outcomes in 33 of the studies he considers (it exerts a negative and *significant* effect in only 7 of them). Yet, even here this may be a statistical artifact, “These positive correlations may result from more senior teachers having the ability to select schools and classrooms with better students” (ibid., p. 1162).
 25. Hanushek (1986), p. 1163.
 26. Card and Krueger (1996), p. 32.
 27. See Card and Krueger (1992a). Early work that used earnings data includes Welch (1966), Johnson and Stafford (1973), Ribich and Murphy (1975), and Wachtel (1976). Behrman and Birdsall (1983) show that both school *quantity* (years) and *quality* (resources) have important effects on subsequent earnings.
 28. Card and Krueger (1994), p. 2.
 29. See, for example, Bishop (1991); Murnane, Willett, and Levy (1995); Grogger (1996a); Grogger and Eide (1993); Neal and Johnson (1996); and Murnane, Willett, Braatz, and Duhaldeborde (2001).
 30. Card and Krueger (1992a), p. 36.
 31. Similar findings and concerns are echoed in Hanushek, Rivkin, and Taylor (1996), and in Deardon, Ferri, and Meghir (2002) (for UK data). These studies find little evidence of any effect school resources have on student outcomes.
 32. These include class size, teachers' salaries, and teachers' levels of education.
 33. Betts (1995), p. 231.
 34. Grogger (1996b) emphasizes that the differences between aggregate and less aggregated results could result from omitted variable bias. In particular, family background factors are correlated with school expenditures and subsequent earnings. It is difficult to control for these factors in studies that look at statewide measures of school quality. Controlling for these effects, his estimates indicate a modest effect of school expenditures and earnings: a \$1K increase in expenditures per pupil is predicted to increase earnings by 2.7% on average.
 35. Card and Krueger (1992b) use data from all of the southern states.
 36. See Angrist and Lang (2004); Hanushek, Kain, Markman, and Rivkin (2003); Sacerdote (2001), and Zimmerman (2003).

37. For example, a 100-point increase in his roommate's verbal SAT score increases a student's mean GPA score by 0.03. This is only 15% of the impact of a 100-point increase in his own SAT on his GPA score.
38. In practice, the experiment was far from ideal. To assuage parental complaints students in regular-size classes were, at the beginning of each year, randomly assigned between classes with and without a student aide. The snag is that students in small classes were not reassigned. This raises the possibility that students could be observed to do well in smaller classes relative to larger ones because of continuity with their classmates, rather than because of any inherent advantage of class size per se. In addition to this problem, about 10% of students were (nonrandomly) reassigned because of parental complaints and behavioral problems. Finally, the study was plagued by sample attrition. Over half of the students present in kindergarten were not observed in subsequent years. This raises the specter that in large classes, for example, poorly performing students were reassigned by their parents to better schools. See Krueger (1999, pp. 499–500) for a discussion of these issues.
39. Angrist and Levy (1999), p. 535.
40. Angrist and Levy (1999), p. 555.
41. Moreover, parents who care about education may engage in political action to increase school funding.
42. Case and Deaton (1999), p. 1050.
43. The classic reference is Oates (1969).
44. Felicia Paik, *The Wall Street Journal*, July 6, 1996.
45. See also Brasington (1999, 2002), who finds that test score performance, expenditures per pupil, and the student–teacher ratio are indeed reflected in house prices.
46. Dale and Krueger (2002, p. 1491).
47. See Clotfelter (1999) for an excellent overview of the major issues.
48. Michael Dobbs, *The Washington Post*, October 22, 2003.
49. On this point Brewer, Eide, and Ehrenberg (1999) note that several econometric studies indicate that the (male) college-wage premium increased by approximately 100% between 1975 and 1985, from 15% to 30%. Part of the recent increase in the relative tuition costs in public universities, reflects a certain catching up with private universities. As reported in Clotfelter (1999), between 1990 and 1995 the average tuition in private universities increased (in real terms) by about 150%; over the same period, tuition in public universities increased by only 15%–20%.
50. See, for example, Taubman and Wales (1974).
51. Hunt (1963, p. 56) as cited in Dale and Krueger (2002, p. 1523).
52. More generally, Q represents characteristics of the college the student attended. It includes a ranking of the college, the average SAT score, and school resources (e.g., faculty per student and annual expenditures on the library).
53. See Dale and Krueger (2002).
54. See Lee (1982) for an excellent discussion of the econometric methods used to overcome selectivity bias.
55. See Loury and Garman (1995). One of the first studies in the area was due to Wise (1974). James, Alsalam, Conaty, and To (1989) also estimate the effects of college quality on earnings. Their findings suggest that institutional characteristics have at most a modest effect on future earnings. What appears to really matter is the choice of major, the number of math credits, and students' GPA scores.
56. The data are based on a stratified sample that was designed to be representative of the U.S. population of high-school graduates circa 1972. The same individuals were re-interviewed in 1973, 1974, 1976, 1979 and 1986, giving an extremely rich source of data on a cohort of high-school students and their subsequent labor-market performance.
57. This figure is approximately the same as the one obtained by Andrisani and Daymont (1984), for male (White and Black) engineering students using the same data set.
58. Kermit, Black, and Smith (1995), p. 16.

59. See also Hoxby (1998).
60. Hoxby (1998), p. 17.
61. Ichimura and Taber (2002) also examine the effect of tuition subsidies. They find that subsidies have substantial effects. However, Heckman, Lochner, and Taber (1998) sound a cautionary note. The key innovation of their approach is dealing with the fact that wages are endogenous and depend on the numbers of students enrolled in college (which, in turn, depend on the tuition subsidies). They show that studies that ignore (anticipated) endogenous wage changes predict that enrollments are 10 times more responsive to price than those that do control for them. Rothschild and White (1995) develop a model of optimal scholarships. They examine the optimal tuition structure in an environment in which high-ability students improve the performance of their lower-ability peers. Thus a high-ability student simultaneously demands education and supplies his high skills to the university, which improves the performance of less able students. The authors show that this supply of skill effect can induce universities to bid for high-ability students using scholarships.
62. See Kane and Rouse (1995) and especially Kane and Rouse (1999) for a discussion of the educational role played by community colleges.
63. Kane and Rouse (1995), p. 610.
64. See also LaLonde (1995).

REFERENCES

- Andrisani, Paul J., and Thomas N. Daymont. (1984). "Job Preferences, College Major, and the Gender Gap in Earnings." *Journal of Human Resources* 19(3): 408–428.
- Angrist, Joshua D., and Kevin Lang. (2004). "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program." *American Economic Review* 5(94): 1613–1634.
- Angrist, Joshua D., and Victor Lavy. (1999). "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics* 114(2): 533–575.
- Becker, Gary S. (1962). "Investment in Human Capital: A Theoretical Analysis." *Journal of Political Economy* 70(5(2)): 9–49.
- Behrman, Jere R., and Nancy Birdsall. (1983). "The Quality of Schooling: Quantity Alone Is Misleading." *American Economic Review* 73(5): 928–946.
- Behrman, Jere R., Mark R. Rosenzweig, and Paul Taubman. (1996). "College Choice and Wages: Estimates Using Data on Female Twins." *Review of Economics and Statistics* 78(4): 672–685.
- Betts, Julian R. (1995). "Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth." *Review of Economics and Statistics* 77(2): 231–250.
- . (1996). "Do School Resources Matter Only for Older Workers?" *Review of Economics and Statistics* 78(4): 638–652.
- Bishop, John H. (1991). "Achievement Test Scores, and Relative Wages." In *Workers and Their Wages: Changing Patterns in the United States*. Ed. H. Koster Marvin. Washington, DC: American Enterprise Institute.
- Black, Sandra E. (1999). "Do Better Schools Matter? Parental Valuation of Elementary Education." *Quarterly Journal of Economics* 114(2): 577–599.
- Blau, Francine D. (1996). "Symposium on Primary and Secondary Education." *Journal of Economic Perspectives* 10(4): 3–8.
- Brasington, David M. (1999). "Which Measures of School Quality Does the Housing Market Value?" *Journal of Real Estate Research* 18(3): 395–413.
- . (2002). "The Demand for Local Public Goods." *Public Finance Review* 30(3): 163–187.
- Brewer, Dominic J., Eric R. Eide, and Ronald G. Ehrenberg. (1999). "Does It Pay to Attend an Elite

- Private College? Cross-Cohort Evidence on the Effects of College Type on Earnings." *Journal of Human Resources* 34(1): 104–123.
- Burtless, Gary. (1996). *Introduction and Summary. Does Money Matter? The Effects of School Resources on Student Achievement and Adult Success.* Washington DC: Brookings Institute.
- Card, David, and Alan B. Krueger. (1992a). "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." *Journal of Political Economy* 100(1): 1–40.
- . (1992b). "School Quality and Black-White Relative Earnings: A Direct Assessment." *Quarterly Journal of Economics* 107(1): 151–200.
- . (1994). "The Economic Returns to School Quality: A Partial Survey." Working Paper. Princeton, NJ: Princeton University, Department of Economics, Industrial Relations Section.
- . (1996). "School Resources and Student Outcomes: An Overview of the Literature and New Evidence From North and South Carolina." *Journal of Economic Perspectives* 10(4): 31–50.
- Case, Anne, and Angus Deaton. (1999). "School Inputs and Educational Outcomes in South Africa." *Quarterly Journal of Economics* 114(3): 1047–1084.
- Clotfelter Charles, T. (1999). "The Familiar but Curious Economics of Higher Education: Introduction to a Symposium." *Journal of Economic Perspectives* 13(1): 3–12.
- Coleman, John, E. Campbell, C. Hobson, J. McPartland, A. Mood, F. Weinfeld, and R. York. (1966). *Equality of Educational Opportunity.* Washington, DC: Government Printing Office.
- Dale, Stacy Berg, and Alan B. Krueger. (2002). "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables." *Quarterly Journal of Economics* 117(4): 1491–1527.
- Deardon, Lorraine, Javier Ferri, and Costas Meghir. (2002). "The Effect of School Quality on Educational Attainment and Wages." *Review of Economics and Statistics* 84(1): 1–20.
- Duflo, Esther. (2001). "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review* 91(4): 795–813.
- Dynarski, Susan. (2002). "The Behavioral and Distributional Implications of Aid for College." *American Economic Review, Papers and Proceedings* 92(2): 279–284.
- Goldin, Claudia. (1999). "Egalitarianism and the Returns to Education during the Great Transformation of American Education." *Journal of Political Economy* 107(6): S65–S94.
- Goldin, Claudia, and Lawrence F. Katz. (1999). "The Shaping of Higher Education: The Formative Years in the United States, 1890 to 1940." *Journal of Economic Perspectives* 13(1): 37–62.
- Grogger, Jeff. (1996a). "Does School Quality Explain the Recent Black/White Wage Trend?" *Journal of Labor Economics* 14(2): 231–253.
- . (1996b). "School Expenditures and Post-schooling Earnings: Evidence from High School and Beyond." *Review of Economics and Statistics* 78(4): 628–637.
- Grogger, Jeff, and Eric R. Eide. (1993). "Changes in College Skills and the Rise in the College Wage Premium." *Journal of Human Resources* 30(2): 280–310.
- Hanushek, Eric A. (1986). "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature* 24(3): 1144–1177.
- . (1996). "Measuring Investment in Education." *Journal of Economic Perspectives* 10(4): 9–30.
- . (2002). "Publicly Provided Education." In *Handbook of Public Economics*. Vol. 4. Ed. Alan J. Auerbach and Martin Feldstein. Amsterdam, North Holland: Elsevier, pp. 2047–2143.
- Hanushek, Eric A., John F. Kain, Jacob M. Markman, and Steven G. Rivkin. (2003). "Does Peer Ability Affect Student Achievement?" *Journal of Applied Econometrics* 18(5): 527–544.
- Hanushek, Eric A., S. G. Rivkin, and L. L. Taylor. (1996). "Aggregation and the Estimated Effects of

- School Resources." *Review of Economics and Statistics* 78(4): 611–627.
- Heckman, James J., Anne Layne Farrar, and Petra Todd. (1996). "Human Capital Pricing Equations with an Application to Estimating the Effect of Schooling Quality on Earnings." *Review of Economics and Statistics* 78(4): 562–610.
- Heckman, James J., Lance Lochner, and Christopher Taber. (1998). "General Equilibrium Treatment Effects: A Study of Tuition Policy." *American Economic Review (Papers and Proceedings)* 88(2): 381–386.
- Heckman, James J., and Yona Rubinstein. (2001). "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *American Economic Review* 91(2): 145–149.
- Heckman, James J., and Edward Vytlačil. (2001). "Identifying the Role of Cognitive Ability in Explaining the Level of Change in the Return to Schooling." *Review of Economics and Statistics* 83(1): 1–12.
- Hedges, L. V., R. D. Laine, and Greenwald. (1994). "Does Money Matter: A Meta-analysis of Studies of the Effects Differential School Inputs Have on Student Outcomes." *Educational Researcher* 23(3): 5–15.
- Hoxby, Caroline M. (1998). "The Return to Attending a More Selective College." Working Paper. Cambridge, MA: Harvard University.
- . (2000). "The Effects of Class Size on Student Achievement: New Evidence from Population Variation." *Quarterly Journal of Economics* 115(4): 1239–1285.
- Iacovou, Maria. (2003). "Class Size in the Early Years: Is Smaller Really Better?" *Education Economics* 10(3): 261–290.
- Ichimura, Hidehiko, and Christopher Taber. (2002). "Semiparametric Reduced-Form Estimation of Tuition Subsidies." *American Economic Review* 92(2): 286–292.
- Jacob, B. A., and Steven D. Levitt. (2003). "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating." *Quarterly Journal of Economics* 118(3): 843–877.
- James, Estelle, Nabeel Alsalam, Joseph C. Conaty, and Duc-Le To. (1989). "College Quality and Future Earnings: Where Should You Send Your Child to College?" *American Economic Review (Papers and Proceedings)* 79(2): 247–252.
- Johnson, George E., and Frank P. Stafford. (1973). "Social Returns to Quantity and Quality of Education." *Journal of Human Resources* 8(2): 139–155.
- Kane, Thomas J., and Cecilia E. Rouse. (1995). "Labor-Market Returns to Two- and Four-Year College." *American Economic Review* 85(3): 600–614.
- . (1999). "The Community College: Educating Students at the Margin between College and Work." *Journal of Economic Perspectives* 13(1): 63–84.
- Kermit, Daniel, Dan A. Black, and Jeffrey A. Smith. (1995). *College Quality and the Wages of Young Men*. Working Paper. London, Ontario: Department of Economics, University of Western Ontario.
- Krueger, Alan B. (1999). "Experimental Estimates of Education Production Functions." *Quarterly Journal of Economics* 114(2): 497–532.
- LaLonde, Robert J. (1995). "The Promise of Public Sector-Sponsored Training Programs." *Journal of Economic Perspectives* 9(2): 149–168.
- Lazear, Edward P. (2001). "Educational Production." *Quarterly Journal of Economics* 116(3): 777–803.
- Lee, Lung Fei. (1982). "Some Approaches to the Correction of Selectivity Bias." *Review of Economic Studies* 49(3): 355–372.
- Leigh, Duane E., and Andrew M. Gill. (1997). "Labor Market Returns to Community Colleges: Evidence for Returning Adults." *Journal of Human Resources* 32(2): 334–353.
- Loeb, Susanna, and John Bound. (1996). "The Effect of Measured School Inputs on Academic Achievement: Evidence from the 1920s, 1930s and 1940s Birth Cohorts." *Review of Economics and Statistics* 78(3): 653–664.
- Loury, Linda Datcher, and David Garman. (1995). "College Selectivity and Earnings." *Journal of Labor Economics* 13(2): 289–308.

- Murnane, Richard J., John B. Willett, M. J. Braatz, and Y. Duhaldeborde. (2001). "Do Different Dimensions of Male High School Students' Skills Predict Labor Market Success a Decade Later? Evidence from the NLSY." *Economics of Education Review* 20(4): 311–320.
- Murnane, Richard J., John B. Willett, and Frank Levy. (1995). "The Growing Importance of Cognitive Skills in Wage Determination." *Review of Economics and Statistics* 77(2): 251–266.
- National Commission on Excellence in Education. (1983). "A Nation at Risk: The Imperative for Educational Reform." Policy Report. Washington, DC: U.S. Department of Education.
- Neal, Derek, and William R. Johnson. (1996). "The Role of Premarket Factors in Black–White Wage Differences." *Journal of Political Economy* 104(5): 869–895.
- Oates, Wallace. (1969). "The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis." *Journal of Political Economy* 77(6): 957–971.
- Ribich, Thomas L., and James L. Murphy. (1975). "The Economic Returns to Increased Educational Spending." *Journal of Human Resources* 10(1): 56–77.
- Rothschild, Michael, and Lawrence J. White. (1995). "The Analytics of the Pricing of Higher Education and Other Services in Which the Customers Are Inputs." *Journal of Political Economy* 103(3): 573–586.
- Sacerdote, Bruce. (2001). "Peer Effects with Random Assignment: Results for Dartmouth Roommates." *Quarterly Journal of Economics* 116(2): 681–704.
- Summers, Anita, and L. Wolfe Barbara. (1977). "Do Schools Make a Difference?" *American Economic Review* 67(4): 639–652.
- Taubman, Paul, and Terence Wales. (1974). *Higher Education and Earnings*. New York: McGraw-Hill.
- Wachtel, Paul. (1976). "The Effects on Earnings of School Quality and College Investment Expenditures." *Review of Economic Statistics* 58(3): 326–331.
- Welch, Finis. (1966). "Measurement of the Quality of Schooling." *American Economic Review (Papers and Proceedings)* 56(1/2): 379–392.
- Winston, Gordon C. (1999). "Subsidies, Hierarchy and Peers: The Awkward Economics of Higher Education." *Journal of Economic Perspectives* 13(1): 13–36.
- Wise, David A. (1974). "Academic Achievement and Job Performance." *American Economic Review* 65(2): 350–366.
- Zimmerman, David J. (2003). "Peer Effects in Academic Outcomes: Evidence From a Natural Experiment." *Review of Economics and Statistics* 85(1): 9–23.