CHAPTER

13

Education production functions

Eric A. Hanushek

Hoover Institution, Stanford University, CESifo, IZA and NBER, Stanford, CA, United States

Glossary

Educational production function A function that relates various inputs to education including those of families, peers, and schools to the maximum level of student achievement that can be obtained

Fixed effects A form of statistical analysis that removes the average effects of a factor (such as individual schools) from the analysis; in the case of teachers, the fixed effect in models of achievement growth is often interpreted as a measure of teacher quality

Value-added In the context of education production functions, the value-added of an input would be the separate contribution of learning after allowing for other inputs and the base level of knowledge of the students

Overview

Much of the analysis in the economics of education flows from a simple model of production. The common inputs are things like school resources, teacher quality, and family attributes; and the outcome is some measure of student achievement—frequently but not always student test scores. Knowledge of the production function for schools can be used to assess policy alternatives and to judge the effectiveness and efficiency of public provided services. This area of research is, however, distinguished from many because the results of analyses enter quite directly into the policy process.

The attention to education production functions is driven largely by recognition that individual skills have significant payoffs in the labor market and elsewhere. Thus, a natural question is how skills can be developed and enhanced, leading to the analysis of how schools and other educational inputs enter into skill development.

This discussion focuses largely on evidence from the US where there is a lengthy history of analysis of education production functions. The focus of this work has changed over time, moving from standard inputs and resources to the effectiveness of teachers. This analysis has been aided by the development of much more extensive data bases on school performance that come from school accountability systems. There has also been wider analysis of educational production from other countries (Woessmann, 2016).

Measuring skills and human capital

Education production functions have their roots in the more general analysis of human capital, but the two different streams of analysis have largely diverged in the past. Most human capital analysis has strong and direct linkages

The Economics of Education, Second Edition
https://doi.org/10.1016/B978-0-12-815391-8.00013-6

to labor market outcomes and the determination of earnings. Education production functions, while ostensibly closely related, have focused more on the underlying determination of skills and human capital. Until recently, the two different foci have led to quite different perspectives on both the measurement of human capital and implicitly the fundamental modeling of economic outcomes.

Historically, human capital has been considered from the labor market perspective of the individual. In its simplest form, individuals make investments that develop their skills, and this stock of skills is optimized for the labor market. In this analysis, the most frequently employed measure of individual skills, or human capital, has been school attainment, or simply years of schooling completed.

There are several justifications for relying on years of schooling to measure individual skills. First, a prime motivation for the schools is the acquisition of knowledge and skills, and this justifies the heavy governmental investment in schools of nations around the world. Second, in the early development of human capital theory, Mincer (1970, 1974) developed a simple but elegant investment model for individuals that emphasized time spent in school. This theoretical development translated into one of the most successful empirical models—the "Mincer earnings function"—that relates individual earnings to school attainment and to labor market experience. The value of school attainment as a rough measure of individual skill has subsequently been verified by a wide variety of studies of labor market outcomes (Card, 2001). Third, reliance on years of school as the human capital measure is expedient. School attainment is very commonly measured in censuses and surveys, allowing, for example, estimation of Mincer earnings functions in 139 countries (see the review in Psacharopoulos & Patrinos, 2018).

However, the difficulty with this common measure of outcomes is that it simply counts the time spent in schools without judging what happens in schools—thus, it does not provide a complete or accurate picture of outcomes. It assumes a year of schooling produces the same amount of student achievement, or skills, over time and in every country. Importantly, it also assumes that schooling is the only input into skill development, ignoring the extensive contrary evidence discussed below. Finally, this measure of school outcomes ignores the extensive policy debates about ways to improve school quality.

A common extension in the investigation of individual human capital and labor market outcomes is the addition of cognitive skills as gauged by standardized test scores to the empirical models. Such skill measures have found their way into the literature slowly, because they have just been available for a limited number of surveys that include both achievement and labor market outcomes. Moreover, in these analyses the common interpretation of these measures is that they represent individual "ability" and thus can be added to a standard Mincer earnings function to correct for any school selection bias arising from people with higher ability continuing longer in school. Such investigations of the impact of cognitive achievement include, for example, Lazear (2003); Mulligan (1999); and Murnane, Willett, Duhaldeborde, & Tyler (2000)). In these, however, the general interpretation is still that school attainment is the measure of human capital.

A complementary line of research has considered aggregate human capital and how it affects national productivity and growth. The surge in empirical analyses of growth differences across countries begun in the early 1990s invariably included measures of school attainment to reflect the skills of the population (see Barro, 1991; Romer, 1990). In a subsequent comparison of alternative drivers of growth rate differences, Sala-i-Martin, Doppelhofer, and Miller (2004) found primary school enrollment to be among the strongest explanatory factors. Nonetheless, the estimated impacts of human capital and other inputs appeared very unstable, casting doubt on the line of empirical growth analyses (Levine & Renelt, 1992).

The parallel line of research into education production functions has concentrated on understanding the determinants of human capital. Test scores, or measures of cognitive skills more generally, have been interpreted as proxies for skills that are valued in the labor market and elsewhere and, as such, more immediate measures of human capital differences. As described below, the overall focus of this work has then been understanding how schools and other factors determine the skills of individuals.

But, this latter focus also leads to a very different interpretation of the prior labor market studies focused on school attainment. Moreover, reconsideration of the interpretation of measured cognitive skills in the prior human capital/labor market analyses helps to clarify the problems with these prior analyses and to point toward an alternative empirical approach. Specifically, these achievement measures have been interpreted as proxies for skills that are valued in the labor market, albeit the ability designation implies that these are fixed skills. An alternative interpretation is that the tested skills represent an explicit measure of human capital. If so, whether skills were determined by schools or by other inputs would in general not affect their use and interpretation in understanding variations in labor market outcomes. Thinking of skills and their measurement by test scores in this manner implies that school attainment is just one of a variety of inputs into an individual's skills.

The research considering cognitive skills as a direct measure of human capital goes a long way toward resolving some of the apparent anomalies in the prior labor market research. Using data for a representative sample of 23 countries that included test scores along with labor market information on individuals, Hanushek, Schwerdt, Wiederhold and Woessmann (2015) show that estimates of earnings functions in terms of achievement tests are readily interpreted in a human capital framework. Similarly, looking at long run growth in terms of cognitive skills helps resolve many of the difficulties with empirical growth models built on school attainment (Hanushek & Woessmann, 2008). Hanushek and Kimko (2000) demonstrate that quality differences measured by achievement have a dramatic impact on productivity and national growth rates. This is reinforced and extended in Hanushek and Woessmann (2012, 2015), who show that three-quarters of the variation in country growth rates can be explained by a simple growth model that focuses on cognitive skills. They call the aggregate test scores for countries "knowledge capital" to distinguish it from the school attainment measures that are frequently referred to as being synonymous with human capital.

These recent analyses of labor market relationships with standardized achievement test scores complete the linkage between human capital analysis from the individual labor market perspective and education production functions that seek to explain differences in test scores. This linkage provides the rationale for interpreting the results of education production estimates as indicating the longer run economic effects of schools and other inputs.

Basic production function estimates

Analysis of education production functions has a direct motivation. Because educational outcomes cannot be changed by fiat, much attention has been directed at inputs—particularly those perceived to be relevant for policy such as school resources or aspects of teachers.

Analysis of the role of school resources in determining achievement begins with the Coleman Report, the US government's monumental study on educational opportunity released in 1966 (Coleman et al., 1966). While controversial, that study's greatest contribution was directing attention to the distribution of student performance—the outputs as opposed to various school inputs such as spending per pupil or characteristics of teachers (Bowles & Levin,1968; Hanushek & Kain, 1972; Hanushek, 2016).

The underlying model that has evolved as a result of this research is very straightforward. The output of the educational process—the achievement of individual students—is directly related to inputs that both are directly controlled by policymakers (for example, the characteristics of schools, teachers, and curricula) and are not so controlled (such as families and friends and the innate endowments or learning capacities of the students). Further, while achievement may be measured at discrete points in time, the educational process is cumulative; inputs applied sometime in the past affect students' current levels of achievement.

Family background is usually characterized by such socio-demographic characteristics as parental education, income, and family size. Peer inputs, when included, are typically aggregates of student socio-demographic characteristics or achievement for a school or classroom. School inputs typically include teacher background (education level, experience, sex, race, and so forth), school organization (class sizes, facilities, administrative expenditures, and so forth), and district or community factors (for example, average expenditure levels). Except for the original Coleman Report, most empirical work has relied on data constructed for other purposes, such as a school's standard administrative records. More recent work has moved to use micro-data generated from school accountability programs. Statistical analysis (typically some form of regression analysis) is employed to infer what specifically determines achievement and what is the importance of the various inputs into student performance. Over time, attention has been increasingly directed at the statistical identification of factors that are causally related to student outcomes.

Measured school inputs

The central thrust of education production function estimation has changed over time. During the initial period—roughly 30 years starting with the Coleman Report—the analyses followed a common pattern of examining the impact on student learning of specific measures of school inputs. These studies focused on measured resources of the type typically included in school reports. The second period, following this initial estimation phase and carrying through to the present, moved to an examination of specific aspects of production, often with novel methods or data, and to concentration on teacher effects.

The state of knowledge about the impacts of basic resources is best summarized by reviewing available empirical studies from the first period. Most analyses of education production functions directed their attention at a relatively small set of resource measures, and this makes it easy to summarize the results (Hanushek, 2003). The 90 individual publications that appeared before 1995 contain 377 separate production function estimates. For classroom resources, only nine per cent of estimates for teacher education and 14% for teacher−pupil ratios yielded a positive and statistically significant relationship between these factors and student performance. Moreover, these studies were offset by another set of studies that found a similarly negative correlation between those inputs and student achievement. Twenty-nine per cent of the studies found a positive correlation between teacher experience and student performance; however, 71 percent still provided no support for increasing teacher experience (being either negative or statistically insignificant). Subsequent analysis of experience effects consistently indicates that increased experience for the first few years of teaching has a positive impact, but there is little to no additional impact past the initial teaching period (see Hanushek & Rivkin, 2012).

Studies on the direct effect of financial resources provide a similar picture, although here the analysis has been more controversial (see, for example, Hanushek,1994, 2003; Hedges, Laine, & Greenwald, 1994). These indicate that there is very weak support for the notion that simply providing higher teacher salaries or greater overall spending will lead to improved student performance. Per pupil expenditure has received the most attention, but only 27 percent of studies showed a positive and significant effect. In fact, seven per cent even suggested that adding resources would harm student achievement. It is also important to note that studies involving pupil spending have tended to be the lowest-quality studies as defined below, and thus there is substantial reason to believe that even the 27 percent figure overstates the true effect of added expenditure.

These studies make a clear case that resource usage in schools is subject to considerable inefficiency, because schools systematically pay for inputs that are not consistently related to outputs. These results of course do not indicate that money never matters or that money cannot matter. They instead point to the importance of *how* money is spent rather than *how much* is spent—a topic considered below.

Study quality

The previous discussions do not distinguish among studies on the basis of any quality differences. The available estimates can be reasonably categorized by a few objective components of quality. First, while education is cumulative, frequently only current input measures are available, which results in analytical errors and biased estimates of the effects of specific inputs. Second, schools operate within a policy environment determined almost always by higher levels of government. In the United States, state governments establish curricula, provide sources of funding, govern labor laws, determine rules for the certification and hiring of teachers, and the like. If these attributes are important—as much policy debate would suggest—they must be incorporated into any analysis of performance. The adequacy of dealing with these problems can thus be used as a simple index of study quality.

The details of these quality issues and approaches for dealing with them are discussed in detail elsewhere (Hanushek, 2003) and only summarized here. The first problem is ameliorated if one uses the "value added" versus "level" form in estimation. That is, if the achievement relationship holds across grades, it is possible to concentrate on the growth in achievement and on exactly what happens educationally between those points when outcomes are measured. This approach ameliorates problems of omitting prior inputs of schools and families, because they will be incorporated in the initial achievement levels that are measured (Hanushek, 1979). The latter problem of imprecise measurement of the policy environment can frequently be ameliorated by studying performance of schools operating within a consistent set of policies—for example, within individual states in the US. Because all schools within a state operate within the same basic policy environment, comparisons of their performance are not strongly affected by unmeasured policies (Hanushek, Rivkin, & Taylor, 1996).

If the available studies are classified by whether or not they deal with these major quality issues, the prior conclusions about research usage are unchanged (Hanushek, 2003). The best quality studies indicate no consistent relationship between resources and student outcomes. The studies finding strong resource effects, particularly for expenditure per pupil, are heavily concentrated in the group of lowest quality studies.

An additional issue, which is particularly important for policy purposes, concerns whether this analytical approach accurately assesses the causal relationship between resources and performance. If, for example, school decision-makers provide more resources to those they judge as most needy, higher resources could simply signal students known for having lower achievement. Ways of dealing with this include various regression discontinuity or panel data approaches. When done in the case of class sizes, the evidence has been mixed (Angrist & Lavy, 1999; Hoxby, 2000; Rivkin, Hanushek, & Kain, 2005).

More recent studies

The most significant innovation of recent years is the use of large administrative data bases. These data bases employ state or local records on individual student's performance and are most notable for tracking students across grades. Student performance is then related to that programs and personnel that each student is exposed to over time. These large scale databases, often following all students in a state over time, permit controlling for a wide range of influences on achievement through the introduction of fixed effects for schools, individuals, and time (see, for example, Rivkin et al., 2005 or Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2006). These fixed effects hold constant any systematic differences that do not vary within a category (such as constant differences among the sampled schools in terms of the selection of schools by families and teachers) and obtain estimates of various inputs from their variation within each of the schools. By eliminating systematic selection and sorting of students and school personnel, they can concentrate on specific causal factors that determine individual student outcomes.

An additional aspect of the growing state data bases is that students can now be traced to subsequent outcomes—university attendance, labor market experiences, criminal behavior, and more. This type of study, while just becoming more possible, involves linking data across state and federal programs where outside-of-school outcomes can be observed. Importantly, such studies also can validate general estimates of school and teacher effects by, for example, showing the direct linkages of estimated school factors on subsequent earnings for the students (Chetty, Friedman, & Rockoff, 2014).

There has also been continuing study of the effects of simply adding more funding on outputs. This work has in part been motivated by the introduction of production function estimates into court cases about the financing of schools (Hanushek & Lindseth, 2009). In fact, two recent studies use the imposition of court judgments against state funding rules to obtain estimates of the impact of added funding (Jackson, Johnson, & Persico, 2015; Lafortune, Rothstein, & Whitmore Schanzenbach, 2018). The estimation and interpretation of these is the subject of ongoing research.

A final alternative involves the use of random assignment experimentation rather than statistical analysis to break the influence of sample selection and other possible omitted factors. With one major US exception, this approach nonetheless has not been applied to understand the impact of schools on student performance. (Randomized trials have expanded much more rapidly in developing countries; see Duflo, Glennerster, & Kremer, 2007). The US exception is Project STAR, an experimental reduction in class sizes that was conducted in the US state of Tennessee in the mid-1980s (Word et al., 1990). To date, the use of randomized experiments has not had much impact on research or our state of knowledge about the impacts of resources. While Project STAR has entered into a number of policy debates, the interpretation of the results remains controversial because of concerns about the quality of the experiment (Krueger, 1999; Hanushek, 1999). The results of this experiment suggested a significant but small impact of lower class size but that all of the impact was concentrated in the first year of schooling (kindergarten or grade one). Smaller class sizes in later years had no additional impact on student outcomes.

Do teachers and schools matter?

Because of the Coleman Report and subsequent studies discussed above, many have argued that schools do not matter and that only families and peers affect performance. Unfortunately, these interpretations have confused measurability with true effects.

Extensive research since the Coleman Report has made it clear that teachers do indeed matter when assessed in terms of student performance instead of the more typical input measures based on characteristics of the teacher and school. The alternative approach to the examination of teacher quality concentrates on pure outcome-based measures of teacher effectiveness. The general idea is to investigate the "value-added of teacher" by looking at differences in growth rates of student achievement across teachers. A good teacher would be one who consistently obtained high learning growth from students, while a poor teacher would be one who consistently produced low learning growth. Early work relied upon very specialized samples of students (e.g., Hanushek, 1971; Murnane, 1975), but this has subsequently broadened out considerably (Hanushek & Rivkin, 2010).

The general research design is to estimate models of the growth in individual student achievement that can be attributed to various measured school and family factors and to differences in learning across the students with different teachers (see reviews in Hanushek & Rivkin, 2012; Jackson, Rockoff, & Staiger, 2014; Koedel, Mihaly, & Rockoff, 2015). The differences in student achievement growth across classrooms, which can be taken as a measure of teacher quality, appear to be consistent and very large (Hanushek & Rivkin, 2010). Hanushek (1992), for example, estimates that the variation in student outcomes from a good to a bad teacher can be as much as a full year of knowledge per academic year; in other words, while a poor teacher gets gains of 0.5 grade level equivalents during a school year, a good teacher gets gains of 1.5 grade level equivalents. Clearly, with a string of good or bad teachers, the implications for student performance could be very large.

More modern research into state administrative data bases have helped to refine the understanding of the importance of differences in teacher quality. For example, Rivkin et al. (2005) are able to provide rough bounds on the variation in teacher quality as seen within Texas (the source of their administrative database). By these studies, one standard deviation in teacher quality implies around a 0.15 standard deviation in the growth of student achievement. By this, having a series of good teachers (teachers at the 84 percentile of the quality distribution) instead of average teachers would lead to substantially different learning after just a few years. For example, 4–5 years of a good teacher could close the average achievement gap between low income and higher income students.

These estimates of magnitudes can be linked directly to studies of the economic impact of student achievement. Hanushek (2011a, 2011b) shows that a 75th percentile teacher each year generates over $400,000 in added income aggregated over a class of 30 students (compared to an average teacher). On the other hand, a 10th percentile teacher subtracts $800,000 in aggregate from a class of 30 students (again compared to an average teacher). Using a very different estimation approach that links teacher value added to the subsequent earnings of the specific students in the class, Chetty et al. (2014) confirm the order of magnitude of these teacher impacts.

These results can also be reconciled with the prior ones. These differences among teachers are simply not closely correlated with commonly measured teacher characteristics (Hanushek, 1992; Rivkin et al., 2005). Moreover, teacher credentials and teacher training do not make a consistent difference when assessed against student achievement gains (Boyd et al., 2006;

Kane, Rockoff, & Staiger, 2006). Finally, teacher quality does not appear to be closely related to salaries or to market decisions. In particular, teachers exiting for other schools or for jobs outside of teaching do not appear to be of higher quality than those who stay (Hanushek, Kain, O'Brien, & Rivkin, 2005).

The analysis of teacher value-added demonstrates the linkage between the research and policy discussions and decisions. Teacher value-added measures have been actively discussed in terms of teacher evaluations. In reviewing US state policies, the National Council on Teacher Quality (2017) finds that by 2017 39 states require teacher evaluations that include objective measures of student achievement growth, although the exact form and weight placed on these varies widely.

This linkage with policy also heightens attention to the studies, and there has been extensive analysis of the properties of value-added estimates including their stability, their accuracy, and their implications for the nature of teaching; see, for example, Braun, Chudowsky and Koenig (2010), Hanushek and Rivkin (2012), and Haertel (2013).

Benefits and costs

Throughout most consideration of the impact of school resources, attention has focused almost exclusively on whether a factor has an effect on outcomes that is statistically different from zero. Of course, any policy consideration would also consider both the magnitude of the impacts and the costs of change. For magnitude of impact, even the most refined estimates of, say, class size impacts does not give very clear guidance. The experimental effects from Project STAR indicate that average achievement from a reduction of eight students in a classroom would increase by about 0.2 standard deviations, but only in the first grade of attendance in smaller classes (kindergarten or first grade); see Word et al. (1990); Krueger (1999). Hoxby (2000) in her regression discontinuity estimation for Connecticut schools finds no systematic effect of class size. Rivkin et al. (2005), with their fixed effects estimation, find effects half of Project STAR in grade four and declining to insignificance by grade seven.

From a policy perspective the magnitude of alternative estimates is at best small. In order to be relevant for policy, it is necessary to compare the outcomes of any change with its costs. Most educational research ignores such comparisons and neglects any consideration of costs.

It is easy to see the importance of cost considerations when put in the context of the debates over class size reduction. In economic terms the potential impacts of class size reduction are very small when contrasted with the costs of such large class size reductions, which typically involve some of the most expensive policy changes currently contemplated. The relevant alternative policy would be to compare the gains from spending on class size reduction with the potential gains from improving the quality of teachers.

Some conclusions and implications

The existing research suggests inefficiency in the provision of schooling. It does not indicate that schools do not matter. Nor does it indicate that money and resources never impact achievement. The accumulated research surrounding estimation of education production functions simply says there currently is no clear, systematic relationship between resources and student outcomes. The general conclusion from the existing work is that *how* resources are used is generally more important than *how much* is used. At the same time, more modern research into the determinants of student achievement strongly indicates that teacher quality differences are the most significant part of differences across schools.

References

Angrist, J. D., & Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics, 114*(2), 533–575.

Barro, R. J. (1991). Economic growth in a cross section of countries. *Quarterly Journal of Economics, 106*(2), 407–443.

Bowles, S., & Levin, H. M. (1968). The determinants of scholastic achievement–an appraisal of some recent evidence. *Journal of Human Resources, 3*(1), 3–24.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2006). How changes in entry requirements alter the teacher workforce and affect student achievement. *Education Finance and Policy, 1*(2), 176–216.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica, 69*(5), 1127–1160.

Chetty, R., Friedman, J. N., & Rockoff, J. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *The American Economic Review, 104*(9), 2633–2679.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., et al. (1966). *Equality of educational opportunity*. Washington, D.C.: U.S. Government Printing Office.

Duflo, E., Glennerster, R., & Kremer, M. (2007). Using randomization in development economics research: A toolkit. In *Handbook of development economics*. North Holland.

Hanushek, E. A. (1971). Teacher characteristics and gains in student achievement: Estimation using micro data. *The American Economic Review, 60*(2), 280–288.

Hanushek, E. A. (1979). Conceptual and empirical issues in the estimation of educational production functions. *Journal of Human Resources, 14*(3), 351–388.

Hanushek, E. A. (1992). The trade-off between child quantity and quality. *Journal of Political Economy, 100*(1), 84–117.

Hanushek, E. A. (1994). Money might matter somewhere: A response to Hedges, laine, and Greenwald. *Educational Researcher, 23*(4), 5–8.

Hanushek, E. A. (1999). Some findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size effects. *Educational Evaluation and Policy Analysis, 21*(2), 143–163.

Hanushek, E. A. (2003). The failure of input-based schooling policies. *Economic Journal, 113*(485), F64–F98.

Hanushek, E. A. (2011a). The economic value of higher teacher quality. *Economics of Education Review, 30*(3), 466–479.

Hanushek, E. A. (2011b). Valuing teachers: How much is a good teacher worth? *Education Next, 11*(3).

Hanushek, E. A. (2016). What matters for achievement: Updating coleman on the influence of families and schools. *Education Next, 16*(2), 22–30.

Hanushek, E. A., Guido, S., Simon, W., & Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review, 73*, 103–130.

Hanushek, E. A., & Kain, J. F. (1972). On the value of equality of educational opportunity as a guide to public policy. In F. Mosteller, & D. P. Moynihan (Eds.), *On equality of educational opportunity* (pp. 116–145). New York: Random House.

Hanushek, E. A., Kain, J. F., O'Brien, D. M., & Rivkin, S. G. (2005). The market for teacher quality. In *NBER working paper No. 11154*. Cambridge, MA: National Bureau of Economic Research.

Hanushek, E. A., & Kimko, D. D. (2000). Schooling, labor force quality, and the growth of nations. *The American Economic Review, 90*(5), 1184–1208.

Hanushek, E. A., & Lindseth, A. A. (2009). *Schoolhouses, courthouses, and statehouses: Solving the funding-achievement puzzle in America's public schools*. Princeton, NJ: Princeton University Press.

Hanushek, E. A., & Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *The American Economic Review, 100*(2), 267–271.

Hanushek, E. A., & Rivkin, S. G. (2012). The distribution of teacher quality and implications for policy. *Annual Review of Economics, 4*, 131–157.

Hanushek, E. A., Rivkin, S. G., & Taylor, L. L. (1996). Aggregation and the estimated effects of school resources. *The Review of Economics and Statistics, 78*(4), 611–627.

Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature, 46*(3), 607–668.

Hanushek, E. A., & Woessmann, L. (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth, 17*(4), 267–321.

Hanushek, E. A., & Woessmann, L. (2015). *The knowledge capital of nations: Education and the economics of growth*. Cambridge, MA: MIT Press.

Hedges, L. V., Laine, R. D., & Greenwald, R. (1994). Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher, 23*(3), 5–14.

Hoxby, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. *Quarterly Journal of Economics, 115*(3), 1239–1285.

Jackson, C. K., Johnson, R. C., & Persico, C. (2015). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *Quarterly Journal of Economics: Forthcoming*.

Jackson, C. K., Rockoff, J. E., & Staiger, D. O. (2014). Teacher effects and teacher related policies. *Annual Review of Economics, 6*, 801–825.

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York city. *Economics of Education Review, 27*(6), 615–631.

Koedel, C., Mihaly, K., & Rockoff, J. E. (2015). Value-added modeling: A review. *Economics of Education Review, 47*, 180–195.

Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics, 114*(2), 497–532.

Lafortune, J., Rothstein, J., & Whitmore Schanzenbach, D. (2018). School finance reform and the distribution of student achievement. *American Economic Journal: Applied Economics, 10*(2), 1–26.

Lazear, E. P. (2003). Teacher incentives. *Swedish Economic Policy Review, 10*(3), 179–214.

Levine, R., & Renelt, D. (1992). A sensitivity analysis of cross-country growth regressions. *The American Economic Review, 82*(4), 942–963.

Mincer, J. (1970). The distribution of labor incomes: A survey with special reference to the human capital approach. *Journal of Economic Literature, 8*(1), 1–26.

Mincer, J. (1974). *Schooling, experience, and earnings*. New York: NBER.

Mulligan, C. B. (1999). Galton versus the human capital approach to inheritance. *Journal of Political Economy, 107*(6, pt. 2), S184–S224.

Murnane, R. J. (1975). *Impact of school resources on the learning of inner city children*. Cambridge, MA: Ballinger.

Murnane, R. J., Willett, J. B., Duhaldeborde, Y., & Tyler, J. H. (2000). How important are the cognitive skills of teenagers in predicting subsequent earnings? *Journal of Policy Analysis and Management, 19*(4), 547–568.

National Council on Teacher Quality. (2017). *State teacher policy yearbook, 2017*. Washington: National Council on Teacher Quality.

Psacharopoulos, G., & Patrinos, H. A. (2018). Returns to investment in education: A decennial review of the global literature. *Education Economics, 26*(5), 445–458.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417–458.

Romer, P. (1990). Human capital and growth: Theory and evidence. *Carnegie-Rochester Conference Series On Public Policy, 32*, 251–286.

Sala-i-Martin, X., Doppelhofer, G., Ronald, I., & Miller. (2004). Determinants of long-term growth: A bayesian averaging of cassical estimates (BACE) approach. *The American Economic Review, 94*(4), 813–835.

Woessmann, L. (2016). The importance of school systems: Evidence from international differences in student achievement. *The Journal of Economic Perspectives, 30*(3), 3–32.

Word, E., Johnston, J., Pate Bain, H., DeWayne Fulton, B., Boyd Zaharies, J., Lintz, M. N., et al. (1990). *Student/teacher achievement ratio (STAR), Tennessee's K-3 class size study: Final summary report, 1985-1990.* Nashville, TN: Tennessee State Department of Education.

Further reading

Ehrenberg, R. G., Brewer, D. J., Adam, G., & Douglas Willms, J. (2001). Class size and student achievement. *Psychological Science in the Public Interest, 2*(1), 1–30.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature, 24*(3), 1141–1177.

Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review, 30*(3), 466–479.

Hanushek, E. A., & Woessmann, L. (2015). *The knowledge capital of nations: Education and the economics of growth*. Cambridge, MA: MIT Press.

Mishel, L., & Rothstein, R. (Eds.). (2002). *The class size debate*. Washington, DC: Economic Policy Institute.

Moe, T. M. (2011). *Special interest: Teachers unions and America's public schools*. Washington, DC: Brookings Institution Press.

Peterson, P. E. (2010). *Saving schools: From Horace Mann to virtual learning*. Cambridge, MA: Belknap Press of Harvard University Press.