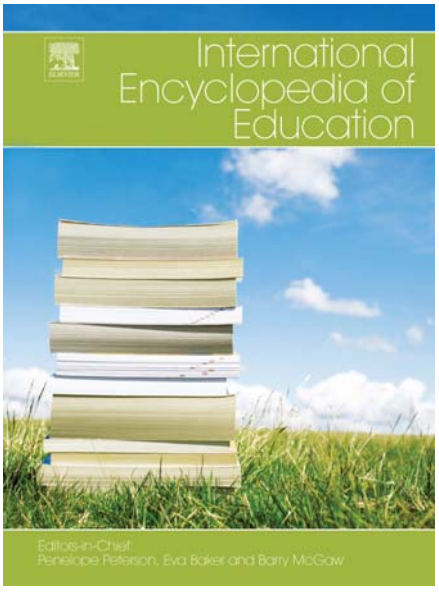

Hanushek E A (2010), Education Production Functions: Developed Country Evidence. In: Penelope Peterson, Eva Baker, Barry McGaw, (Editors), International Encyclopedia of Education. volume 2, pp. 407-411. Oxford: Elsevier.

Education Production Functions: Developed Country Evidence

E A Hanushek, Stanford University, Stanford, CA, USA

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 others because the results of analyses enter quite directly into the policy process.

Historically, the most frequently employed measure of schooling has been attainment, or simply years of schooling completed. The value of school attainment as a rough measure of individual skill has been verified by a wide variety of studies of labor-market outcomes (e.g., Mincer (1970) and Psacharopoulos and Patrinos (2004)). 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.

Recent direct investigations of cognitive achievement find significant labor-market returns to individual differences in cognitive achievement (e.g., Lazear (2003), Mulligan (1999), and Murnane *et al.* (2000)). Similarly, society appears to gain in terms of productivity; Hanushek and Kimko (2000) demonstrate that quality differences in schools have a dramatic impact on productivity and national growth rates. (A parallel line of research has employed school inputs to measure quality but has not been as successful. Specifically, school input measures have not proved to be good predictors of wages or growth.) The economic implications of differences in cognitive achievement appear to be very large (Hanushek, 2005).

As 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, the 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 and Levin, 1968; Hanushek and Kain, 1972).

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 both inputs – ones that are directly controlled by policymakers (e.g., the characteristics of schools, teachers, and curricula) and others that 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 sociodemographic characteristics such as parental education, income, and family size. Peer inputs, when included, are typically aggregates of student-sociodemographic 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 (e.g., 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. 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.

Measured School Inputs

The state of knowledge about the impacts of resources is best summarized by reviewing available empirical studies. Most analyses of education production functions have 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 9% 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 percent of the studies found a positive correlation between teacher experience and student performance; however, 71% still provided no support for increasing teacher experience (being either negative or statistically insignificant). Studies on the effect of financial resources provide a similar picture. 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% of studies showed a positive and significant effect. In fact, 7% 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% figure overstates the true effect of added expenditure. More recent studies do not change this picture about how resources relate to student outcomes.

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.

Study Quality

The previous discussions do not distinguish among studies on the basis of any quality differences. The available estimates can be 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. Second, schools operate within a policy environment set almost always at 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. As all schools within a state operate within the same basic-policy environment, comparisons of their performance are not strongly affected by unmeasured policies (Hanushek *et al.*, 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 and Lavy, 1999; Hoxby, 2000; Rivkin *et al.*, 2005).

The most significant innovation of recent years is the use of large administrative databases. These databases 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 those 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, e.g., Rivkin *et al.* (2005) or Boyd *et al.* (2006)). These fixed effects hold constant any systematic differences that are constant among the 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.

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 exception, this approach nonetheless has not been applied to understand the impact of schools on student performance. The exception is the logislated Student/Teacher Achievement Ratio progs (or 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.

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 do 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) and 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 4 and declining to insignificance by grade 7.

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 (see the range of cost estimates in Brewer *et al.* (1999)). 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, the subject of the next section.

Do Teachers and Schools Matter?

Due to 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 total teacher effects 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.

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 mean differences in learning across the students with different teachers. The differences in student-achievement growth across classrooms, which can be taken as a measure of teacher quality, appear to be very large. Hanushek (1992) 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 databases have helped refine the understanding of the importance of differences in teacher quality. For example,

Rivkin *et al.* (2005) and Hanushek *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 high-income students.

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 *et al.*, 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 *et al.*, 2005).

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. 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.

See also: Cost–Benefit Analysis and Cost–Effectiveness Analysis; Data; Education Production Functions: Concepts; Education Production Functions: Evidence from Developing Countries; Empirical Research Methods in the Economics of Education; Human Capital; Teacher Quality in Education Production; The Economics of Class Size.

Bibliography

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

Bowles, S. and 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., and Wyckoff, J. (2006). How changes in entry requirements alter the teacher workforce and affect student achievement. *Education Finance and Policy* **1**(2), 176–216.

Brewer, D. J., Krop, C., Gill, B. P., and Reichardt, R. (1999). Estimating the cost of national class size reductions under different policy alternatives. *Educational Evaluation and Policy Analysis* **21**(2), 179–192.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., *et al.* (1966). *Equality of Educational Opportunity*. Washington, DC: US Government Printing Office.

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. (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. (2005). The economics of school quality. *German Economic Review* **6**(3), 269–286.

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

Hanushek, E. A., Kain, J. F., O'Brien, D. M., and Rivkin, S. G. (2005). The market for teacher quality. *Working Paper No. 11154*, Cambridge, MA: National Bureau of Economic Research February.

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

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

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.

Kane, T. J., Rockoff, J. E., and Staiger, D. O. (2006). What does certification tell us about teacher effectiveness? *Evidence from* New York City. *Working Paper No. 12155*, Cambridge, MA: National Bureau of Economic Research (April).

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

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

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.

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., Willett, J. B., Duhaldeborde, Y., and 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.

Psacharopoulos, G. and Patrinos, H. A. (2004). Returns to investment in education: A further update. *Education Economics* **12**(2), 111–134.

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

Word, E., Johnston, J., Bain, H. P., *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

Burtless, G. (ed.) (1996). *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*. Washington, DC: Brookings.

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

Friedman, B. M. (ed.) (2003). *Inequality in America: What Role for Human Capital Policies?* Cambridge, MA: MIT Press.

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. (2002). Publicly provided education. In Auerbach, A. J. and Feldstein, M. (eds.) *Handbook of Public Economics*, pp 2045–2141. Amsterdam: Elsevier.

Hanushek, E. A. and Welch, F. (eds.) (2006). *Handbook of the Economics of Education,* vols. 1 and 2. Amsterdam: North-Holland.

Mishel, L. and Rothstein, R. (eds.) (2002). *The Class Size Debate.* Washington, DC: Economic Policy Institute.

Moe, T. M. (ed.) (2001). *A Primer on America's Schools.* Stanford, CA: Hoover Institution Press.

Peterson, P. E. (ed.) (2003). *Our Schools and Our Future: Are We Still at Risk?* Stanford, CA: Hoover Press.