ABSTRACT: This paper provides a literature review on the economics of school quality. It includes a concise analysis of the main trends and findings in the research that has been conducted over the last several years worldwide. Policy implications and priorities for future research are also identified. The literature under review pertains to the United States and other developed countries, as well as to developing countries.

1. Economic Measures of School Performance

The strong and positive link between years of schooling and earnings of students once they enter the labor market is by now one of the best established facts in labor economics. A recent review of research by Psacharopolous (1994) establishes that an additional year of schooling has larger effects on earnings for workers in developing countries, and that within developing countries the returns to education decline with the level of schooling. In light of the observed returns to an extra year of schooling, the imposition of compulsory school attendance laws in both developed and developing countries during the last 100 years and the associated expenditures are widely viewed as an excellent public investment.

Recently, researchers have examined whether the substantial observed returns to additional years of schooling depend on the level of funding that schools receive. Put differently, would an increase in resources per pupil make as attractive an investment as requiring students to attend one additional year of schooling?

Researchers have studied the impact of school spending in a number of ways. The most common method involves testing for a correlation between students’ test scores and the resources of the schools they attend, after controlling for confounding factors such as family background. In the United States, the landmark Coleman Report (1966) established that in American schools the most important determinants of student performance were family background and the student’s peer group. Surprisingly, variations in school resources accounted for little of the variation in student performance. The intervening decades have witnessed a massive effort to prove or disprove the
contention that school resources have little impact on test scores. A recent review by Hanushek (1996) shows that the Coleman finding has stood the test of time remarkably well. For instance, Hanushek reports that of 277 estimates in the literature on the impact of the teacher-pupil ratio, only 15 showed a positive and significant impact on student performance, compared to 13 that showed a negative and significant link. Results for other measures of school resources are similarly mixed, although stronger in some cases.

Hedges, Laine and Greenwald (1994) present a meta-analysis of the test score literature and claim to find strong effects of school resources. Hanushek (1994) provides a rebuttal. He questions the validity of the technique used by Hedges et al. on several grounds. In particular, he notes that if their estimate of the impact of school resources were accurate, then test scores in the United States should have skyrocketed during the 1970’s and 1980’s when spending per pupil rose significantly. In reality, student achievement in the United States improved very little over this period.

Hanushek (1995) reviews research on the impact of school spending in developing countries. Based on 100 studies in developing countries, he reports little evidence that the teacher-pupil ratio or teacher salaries are positively and significantly related to student performance. But fully 35 of 63 studies found a positive significant effect of teachers’ education, and 22 out of 34 studies reported a positive and significant link between school facilities and student performance. The overall pattern suggests that school spending might matter more in a developing country context than in countries such as the United States. Fuller and Clark (1994) report similar evidence. One obvious explanation is diminishing returns. (The phrase ‘diminishing returns’ refers to the often observed pattern in industry in which the effectiveness of an input is particularly high at low levels of application, but decreases as more and more of the input is purchased.)

The research on test scores provides an extremely valuable insight into what types of school spending matter most. But economists have long noted that test scores themselves explain only a small part of the overall variation in outcomes such as earnings or the years of education students obtain (Mason and Griliches 1972). These latter measures, earnings and education, are perhaps more meaningful indicators of adult success than test scores. What, then, do we know about the long-term effect of school spending on student’s later educational attainment and earnings? More broadly, do we even know if there are significant differences in outcomes between students who attended different schools?

2. Variations in School Quality: Evidence from Earnings and Test Scores

One way of determining whether schools vary in quality is to examine whether test scores or earnings of individuals differ systematically across the schools that they attended. Economists tend to focus on earnings of workers rather than test scores for two reasons. First, earnings serve as a proxy for the underlying productivity of the individual, and so provide clues as to the long-term productivity consequences of the types of schooling that workers obtained when they were young. Second, unlike test scores, earnings provide a good indicator of the person's overall economic welfare.
A number of papers have shown that students attending different public schools in the United States have systematic variations in their earnings, after controlling for personal traits of the students. (Betts 1995; Grogger 1996a) The test-score literature provides similar signs that schools vary in quality. Hanushek, Kain and Rivkin (1998) provide evidence based on test scores of individual students in Texas that schools vary significantly in how quickly their students learn. These authors go further, and show that within schools, much of the variation in rates of learning between one grade and the next appears to depend on variations in teacher quality. Similarly, Murnane (1975) shows that individual teachers differ systematically in how quickly their students learn, even after controlling for student traits.

These results suggest that, in the United States at least, schools vary significantly in quality, and that within schools, teachers vary in their effectiveness. Existing research has clearly established these points. However, researchers are just now beginning to understand the underlying causes of these variations in school quality and teacher effectiveness. For policy purposes, it is important to know that schools vary in quality and that within schools the quality of education provided in different grades can differ. But these findings alone provide no guidance as to how to improve the quality of education. To answer this second question we need to examine the impact of different types of school spending on long-term student outcomes.

3. School Resources and Earnings

Betts (1996a) reviews the evidence from the United States on whether and how school resources affect students’ earnings later in life. The evidence is mixed. The most typical approach to studying this question has involved testing for a link between a person’s earnings and the level of school resources in his or her state of birth. These studies typically, but not always, have found that school resources such as the teacher-pupil ratio and teacher salaries are positively and significantly related to earnings. Examples include Rizzuto and Wachtel (1980) and Card and Krueger (1992). A second set of papers has tested for a relation between earnings and spending per pupil in the person’s school district. This research is somewhat more mixed but often finds a positive link between spending and later earnings. Examples include Jud and Walker (1977), Link and Ratledge (1975), Ribich and Murphy (1975) and Wachtel (1975). Finally, some recent studies have tested for a relation between a person’s earnings and the level of resources in the school that the student actually attended. This literature reveals little impact of school resources on students’ later earnings; examples include Betts (1995) and Grogger (1996a).

A key feature of this research is the variety of measures of school resources that researchers have used. The most common measure of school inputs has been spending per pupil, typically measured at the district or state level. But almost as many studies have used the teacher-pupil ratio, or its reciprocal, the pupil-teacher ratio, as the measure of school inputs. The third most commonly studied measure of school resources is teacher salaries. Roughly ten studies have used at least one of the following additional measures of school inputs: teacher education, books per student, teacher experience and length of the school year.
The use of overall spending per pupil provides a useful initial analysis of "whether spending matters", at least given the way schools in the various samples spent their budgets during the time period under study. But Betts (1996a) and Murnane et al. (1991, p. 7) argue that a more useful question might be how education dollars might best be spent, on class size reduction, increasing teachers’ salaries and so on. The latter types of analyses are arguably more useful for policy makers.

Accounting for the Variations in Results

Betts (1996a) finds three strong patterns in this literature, both in the level of statistical significance of the school resources and the size of the estimated effect. These patterns appear to hold regardless of which measure of school input is used. Studies that find smaller and less significant effects tend to:

(1) examine school resources at the actual school attended, rather than in the district or state,

(2) examine earnings of workers under 30 years of age, and

(3) examine earnings of workers educated in the last half of the twentieth century.

Researchers have interpreted these patterns in a number of ways. Hanushek, Rivkin and Taylor (1996) argue, based on test score data, that studies which aggregate beyond the school are subject to "aggregation bias". Aggregation bias refers to incorrect inferences that can result when the researcher uses aggregated data in the presence of non-linear relationships between, in our case, school spending and student outcomes. The authors argue that even without such non-linear relations omitted variable bias will increase when data are aggregated to the geographic level of the most important omitted variables. Since in the United States state governments play an increasingly important role in education policy, one might then suspect that omitted state-level variables are the most important. The authors show that if this is the case, then statistical estimates of the impact of school resources that are based on data aggregated to the state level will be biased upward. They then use High School and Beyond data to show that exactly such a pattern emerges in models of test scores and years of schooling obtained. The authors’ findings suggest that the school-level studies should give us the most accurate picture of how school spending affects students’ later earnings, while the state-level literature should give the least accurate picture due to upwardly biased estimates.

On the other hand, Card and Krueger (1996) speculate that measurement error biases the results in the school- and district-level studies downward. Their argument is that if one asks a school principal about the pupil-teacher ratio in his or her school, or spending per pupil in the district, the principal won’t provide accurate information. If this were so, then using state-level measures of school inputs can "average out" the errors from school to school. Statistically, the estimated impact of school inputs should rise when one uses a state-level proxy for school inputs. Grogger (1996b) tests these two competing hypotheses against each other and concludes that aggregation bias is likely to be the more
important of the two explanations. In other words, the school-level literature is more likely to give an accurate picture of the impact of school spending than is the state-level literature. Hanushek, Rivkin and Taylor (1996) also examine the hypothesis of measurement error in the school-level studies and reject it in favor of the hypothesis of aggregation bias. Overall, these results strongly suggest that school-level studies will provide more reliable estimates of the impact of school resources than will studies based on aggregation over large geographical areas, such as states.

The fact that the less pessimistic studies tend to examine older workers hints that school spending may have an increasing impact on workers’ wages as they get older and settle into career jobs. Betts (1996b) tests this hypothesis explicitly and rejects it. He instead finds evidence that in some cases the impact of school spending might wear off slightly as workers age.

The third pattern, that smaller effects of school resources are observed for samples of American workers educated more recently, leads Betts (1996a) to speculate that there may have been a structural change over time. One explanation for the possibly decreasing effectiveness of American schools over time is that diminishing returns have set in as school spending rose in the United States over the last half century.

Two recent studies do not fit neatly into the pattern that "school spending matters if spending is measured at the state level but not if spending is measured at the school level". Altonji and Dunn (1996) use a fixed-effect approach to control for unobserved family traits. This technique completely removes, or "controls for", all variations between families that are the constant across siblings. This technique can be helpful if, for instance, some unmeasured trait of the family, such as wealth, is associated with both the children’s earnings years later and the level of school resources that they received while in school. In effect, the authors examine how variations in the actual high schools attended by siblings within a family are related to the siblings’ later earnings. They find a positive effect, although one must wonder whether the fact that some parents send different siblings to different schools is a purely random phenomenon.

A second paper that breaks with the past literature, by Heckman, Layne-Farrar and Todd (1996), questions the use of state-level data that in earlier work have given the strongest indications that school spending matters. The authors note that the earlier state-level literature assumes that every additional year of schooling increases earnings by the same percentage. Yet several authors, notably Hungerford and Solon (1987), have shown that in truth earnings levels increase discontinuously at certain points that mark graduation milestones, such as 12 and 16 years of schooling. Heckman, Layne-Farrar and Todd (1996) incorporate these well-known "sheepskin effects" into their model. They find that after allowing for non-linearities in the relation between earnings and education, school resources increase earnings only for the minority of students who later obtain college degrees. Their paper also questions the "identification procedures" used in the state-level literature. (Identification in this case refers to the sources of variation in the data that are used by the researcher to estimate the impact of school resources on earnings. Heckman and his co-authors find that the practice in some earlier work of relying on people who
migrate from their state of birth to other states when they become workers may introduce bias into the estimated effect of school spending.)

**How Big is the Impact of School Resources?**

Betts (1996a) calculates the internal rate of return to increased school spending and to increasing the teacher-pupil ratio. The calculations suggest that the internal rate of return is on the order of 0.55 to 2.55% for spending per pupil. Using the teacher-pupil ratio, the state-level estimates suggest an internal rate of return of 1.45 to 2.35%. However, the internal rate of return, based on the overall estimated impact of the teacher-pupil ratio from estimates at all levels of aggregation, is not defined, meaning that even at a zero rate of discount expenditures designed to boost the teacher-pupil ratio are never recovered through subsequent wage gains for students. Even the highest of these rates of return is very low relative to the real interest rate. It is important to understand that this finding suggests that the returns to increased school resources in the United States are small, given the way in which school currently spend education dollars. It does not guarantee that more effective ways of spending cannot be found.

More recently, Peltzman (1997) makes the same observation, noting that a better investment than reducing class size might be to buy bonds for students at the start of their school years and to return the proceeds to the students upon graduation from school. (This thought experiment is apparently intended by the author not as a policy prescription, but as a cautionary tale about the very small impact of school resources on earnings.)

Betts (1996a) also compares the rates of return to increased school spending with the rate of return to an additional year of high school. He estimates that the internal rate of return to the latter is about 10.95%, far above the 0.55 to 2.55% internal rate of return to additional school spending. One implication of this is that strengthening compulsory school attendance laws might be a more cost-effective strategy than increasing spending per pupil further in the United States. Ironically, in most American states the school-leaving age has been static, at about 16 years of age, for the last three decades, while real spending per pupil has almost tripled (Betts 1998a).

**Identifying the Effects of School Quality by Examining Earnings of Immigrants to the United States**

Bratsberg and Terrell (1997) provide an interesting twist to the literature by asking whether the returns to education among immigrants to the United States depend on average levels of school resources in the immigrants’ original home countries. The authors find that several measures of school resources are positively correlated with the returns to education among immigrants. The estimated effects are close to those in the literature that uses Americans’ state of birth to assign an estimate of school resources. Betts and Lofstrom (forthcoming) also examine the impact of source county traits on immigrants’ earnings in the United States. They find that the pupil-teacher ratio is negatively related to immigrants’ earnings, but only for those immigrants who had obtained some postsecondary education before leaving their home country. This finding
is similar to the finding by Heckman, Layne-Farrar and Todd (1996) that average school resources in Americans’ state of birth are positively linked to earnings only for those who have obtained postsecondary education. Overall, these two papers suggest that school spending outside the United States has a measurable positive influence on earnings of immigrants once they enter the United States, although the effect might be limited to immigrants with some postsecondary education.

4. School Resources and Educational Attainment

An intermediate path between the extremes of using test scores of young students and earnings of students years after they have graduated is to examine the educational attainment of students. Betts (1996a) reviews the literature on how school inputs affect educational attainment. The number of published studies on this topic is smaller than the number of studies that analyze earnings (Betts reports 13 papers on educational attainment compared to 22 studies on earnings). There also exists considerable heterogeneity in the way in which these studies measure educational attainment (e.g. years of schooling versus graduation).

Interestingly, similar patterns emerge to those found in the research on earnings. Research that uses state-level measures of school resources finds a significant effect more often than studies that measure resources at the level of the school actually attended. The former approach also yields larger estimated effects. For example, 100% of the studies that use state-level measures of spending per pupil find a significant link to educational attainment. This compares to 83% of the district-level estimates and 0% of the lone estimate produced at the school level. The same pattern emerges for other measures of school inputs. For instance, 100% of the state-level and district-level studies of the teacher-pupil ratio suggest a positive link to educational attainment, while only 33% of the school-level studies find a similar link. All of the state-level studies found a significant relation between teachers’ salaries and educational attainment, but 0% of the district- and school-level studies found a significant relation. Similarly, the estimated impact of school spending is about 10 times smaller in the district-level studies than in the state-level studies. (Betts, 1996a, does not report the average size of the effect for the school-level study because the paper in question does not provide the required data.) Overall, these studies suggest that the link between school spending and students’ educational attainment might be positive, but is rather weak and does not appear to be systematic.

5. Standards, Incentives and Competition

Given the mixed results in the American literature on the impact of school spending, and the small rates of return to school spending in even the most optimistic branches of the literature, many researchers in the United States have turned their attention to the role of non-monetary factors in determining school quality. Researchers both in the United States and other countries have asked whether educational standards and testing could affect student achievement. Yet other studies examine whether schools would become more efficient if competition between schools grew. There are a number of ways of
increasing such competition, but all hinge in some way upon the notion of allowing parents to choose between schools for their children.

The bulk of this literature focuses on the impact of incentives on test scores, and often the analysis is based on fairly crude correlational analysis conducted across fairly highly aggregated geographical areas. Other studies are conducted at less aggregated levels.

A Lack of Incentive to Study

John Bishop has written a series of detailed articles providing indirect evidence that American students lag their counterparts in other developed countries because students in the United States lack an incentive to study. Two examples from Bishop’s (1996a, b, c; 1997) surveys provide the flavor of his logic. He finds that in Canada, provinces that have instituted graduation exams students have considerably higher test scores. Students in these provinces also report doing more homework and watching less television. As a second example, he notes that students from New York have routinely obtained higher scores on the Scholastic Aptitude Test (now known simply as the SAT) than students from most other states. He attributes this to New York’s long history of statewide graduation exams, arguing that they provide students with an adequate incentive to learn during the high school years. He provides numerous similar examples from cross-country comparisons as well.

Costrell (1994) and Betts (1998b) provide theoretical analyses of the impact of educational standards. Both papers make the point that administrators will never find it easy to change standards, because any change in standards, either an increase or a decrease, will produce both winners and losers. Costrell’s article presents the most fully articulated model of the political economy of standards to date. His other findings include the result that if educational standards are set locally, rather than in a centralized way, they will tend to be set less stringently. Betts (1998b) assumes that students vary in ability, and that employers cannot fully determine the ability of job applicants. In this case, an egalitarian society, that is, one that expresses particular concern about its least well off members, may in fact want to set fairly high standards rather than low standards. The intuition for this result is that in a society such as the United States, in which only a small portion of students drop out before graduating, the stigma attached to dropping out will be much higher than in a society where dropping out is more routine. This engenders a bigger wage gap between these workers and more highly educated workers than would otherwise arise.

With these complications in mind, what do we know about how variations in educational standards affect student outcomes in the real world? Betts (1996c and 1997) provides evidence that in the United States teachers who assign more homework, and schools that have higher grading standards, produce students whose test scores grow quickly relative to scores of other students. Both of these factors explain more of the variation in gains in student achievement than do spending-related factors including class size, teacher education and teacher experience. Cooper (1989) reviews the existing literature on the impact of homework on student learning. Overall, this literature provides quite strong and
consistent evidence that homework is positively related to gains in student achievement. However, the results are much more consistent at the high school level than at the primary level.

One perennial topic of debate among educators is the value of grade retention, that is, the policy of requiring students to repeat a grade if they have not mastered the curriculum in a given year. Analysis of studies in the United States suggests that students who are held back a year are more likely to drop out of school, unless during the subsequent year students receive additional help (Holmes 1989).

The Chicago Public Schools recently introduced what, in an American context, is a radical program. Students who are sufficiently far behind national norms on achievement tests are required to attend summer school, and then re-take the tests. Those who do not improve to an acceptable level must repeat the grade. In the first year of this program, substantial minorities of each cohort of Chicago students were required to attend summer school. Of these students, roughly one half improved sufficiently to proceed to the next grade. The average gain in test scores over the summer was very high Betts (1998a). However, the large summer gains in test scores may in part reflect "teaching to the test". It will take some time to learn if the effects of the program are long lasting.

Betts (1998a) also shows, using representative national data-sets, that if similar programs were imposed throughout the United States large fractions of students would fail. For instance, suppose that the standard for passing on to the next grade in mathematics was that a student had to obtain a test score as high as the average score among students two grades less advanced. In such a case, where a Grade 9 student would be held back if he or she did not score at the average level for a Grade 7 student, roughly 25% of students would fail. This fact points to a key concern for all educators: the dispersion in student achievement within grade levels, at least in the United States, is extremely large relative to the average gain in achievement from one grade to the next.

*Competition between Schools for Students*

A central tenet of neoclassical economics is that free markets improve efficiency because open competition between suppliers of a good or service lowers the cost and/or increases the quality provided to the consumer. In contrast, a monopolist or an "oligopoly", that is, a small group of suppliers, can restrict supply, driving up the price of the good or service, or, at the same price, reducing the quality of the good or service. In many developed countries, public schooling more closely represents a monopoly or oligopoly, in that parents’ choice of school for their children is dictated almost solely by the neighborhood in which they live. So, the argument goes, allowing parents to choose freely between schools will force all schools to maximize the quality of education they provide for a given level of funding. Otherwise they risk losing students to other schools.

Of course, the claim that greater competition in the "market for students" will improve student outcomes hinges upon the assumption that there exist no "market imperfections"
such as imperfect information. See Fels (1996) and Aoki and Feiner (1996) for a review of these issues.

There are a number of ways in which the degree of school choice could be increased in a country such as the United States. The recent spread of charter schools in the United States has increased the degree of choice afforded parents, and at the same time has reduced the regulatory oversight of schools. This solution increases the degree of competition within the public school system. A second way of increasing school choice involves granting parents vouchers that permit them to send their children to either public or private schools.

Evidence on the importance of inter-school competition for students is limited, but suggests that additional competition improves outcomes for students. For instance Hoxby (1994a) shows that the presence of private schools in an area appears to improve the efficiency of nearby public schools that must compete for students from the same area. Hoxby (1994b) shows that in American cities with more than one public school district, student outcomes are significantly better, perhaps because of parents’ implied ability to abandon one district in favor of another by moving.

Evidence on what existing forms of school choice have achieved is still relatively sparse. Rouse (1998) provides a balanced review of what researchers have learned from small-scale choice programs in the United States, such as the school voucher program implemented in Milwaukee, Wisconsin. Considerable disagreement surrounds this program, but overall, the statistical evidence suggests that disadvantaged students attending private schools through the Milwaukee voucher program have improved educational outcomes relative to similar students left behind in the public schools. However, the effects are modest. Moreover, it remains to be seen whether the added competition provided by Milwaukee’s voucher plan has spurred the public school system itself to improve its effectiveness. Enrollment caps placed on the Milwaukee voucher program will probably limit its effectiveness as a "spur to competition" in the public schools. Hoxby (1996a) and Kane (1996) provide a worthwhile exchange on whether private school vouchers are likely to prove an effective reform.

Some of the most interesting research programs designed to study the impact of school competition and choice have been conducted in developing countries. Patrinos and Ariasingam (1997) review the worldwide literature on demand-side financing of education and then summarize a number of World Bank projects designed to give parents greater access to schooling for their children. This diverse set of programs ranges from direct financing of community schools in Chad to a student loan program in Jamaica and a full-blown private-school voucher program in Columbia. Most of these programs have not yet been analyzed formally, but evidence from Bangladesh suggests that in a USAID program subsidies given to poorer families led to dramatic improvement in enrollment rates in participating schools, on the order of 13%. In a study funded by the World Bank, Jimenez and Sawada (1998) provide evidence on the EDUCO school-choice program in El Salvador. Under this program, parents in certain areas can opt to send their children to an EDUCO school. These schools are governed by a local parent council.
Sawada find that, after adjusting for the lower socioeconomic status of students in these schools, EDUCO students did not differ significantly in math achievement but had higher attendance rates and higher language test scores than did students in regular schools. These results suggest that the average quality of schooling may have improved somewhat with the introduction of school choice and decentralized control.

**Teacher Unions**

Eberts and Stone (1984) and Hoxby (1996b) test whether teacher unionization in the United States affects the productivity of school spending. Both provide evidence pointing in this direction. For instance, Eberts and Stone (1984) report that while unionized and non-unionized schools are about equally effective in promoting student achievement, unionized schools do so at a 15% higher cost. Hoxby shows that the impact of school inputs on student outcomes is insignificant in unionized schools, but highly significant in the minority of schools at which teachers are not unionized. The causes of this negative correlation between teacher unionization and the effectiveness of school spending are not entirely clear. One interpretation is that teachers’ unions capture economic "rents" from school administrations.

**6. Policy Implications**

At one level, the key policy debate in education, especially in developing countries, concerns breadth versus depth. Should governments expand accessibility to a given level of education, be it primary in some countries or even postsecondary in other more developed countries, or should they instead focus on providing the best possible education to a smaller pool of students? Psacharopolous’ (1994) aforementioned review of the literature on the returns to an extra year of schooling strongly suggests that in developing countries, providing universal access to primary education is more important than providing the highest level of education to a smaller group of students. This conclusion follows from his observation that the returns to an extra year of schooling appear to decline with the level of schooling.

A closely related policy question concerns whether to spend on additional years of schooling for young people, or to increase the quality of education, for a given level of schooling, by increasing spending per pupil. The extensive American literature suggests that the former policy has a much higher payoff than the latter. For example, Betts and Roemer (1998) address the relative effectiveness of changing school spending and raising the school-leaving age as policy tools to reduce the earnings gap between black and white men in the United States. They conclude that raising the school-leaving age is about three times as cost-effective as increasing school spending. However, they find that raising the school-leaving age by itself does little to narrow the earnings gap, in part because relatively few people in the United States drop out of high school.

Increasing the average years of schooling through stronger laws on compulsory attendance -- and enforcement of these laws -- could prove relatively more effective in developing countries than in the United States. Such policy changes might work well
given higher drop-out rates in many developing countries, combined with the evidence provided by Psacharopolous (1994) that the returns to additional schooling are strong in developing countries, particularly at lower levels of education.

The developing-country literature on school spending, as surveyed by Hanushek (1995), points to the same conclusion that additional spending might be quite ineffective relative to expanding access to education. However, it seems that a greater proportion of the studies in developing countries reveal statistically significant effects of school resources on student outcomes than is true in the literature from the United States. Based on the proportion of developing-country studies that find significant effects, it appears that expenditures on facilities and teacher training are more likely to increase student outcomes significantly than are other expenditures such as reducing class size. See Hanushek (1995) and a reply by Kremer (1995) for the beginnings of an exchange on "what matters most" in developing countries’ schools.

One recent experimental study of Kenyan education highlights these mixed findings. The study found that expenditures for textbooks and uniforms reduced drop-out rates relative to a control group of students in schools that did not participate (Kremer et al. 1997). At the same time class sizes increased considerably as parents decided to enroll their children at the schools that received the additional funds. Overall, the innovations had little effect on test scores. Since the overall change in funding for textbooks and uniforms and higher class sizes was about zero, the authors conclude that Kenyan schools could reduce drop-out rates without lowering test scores or increasing spending. The schools could do this by increasing class size and using the savings to pay for textbooks and to reduce the fixed costs of sending children to school. Glewwe et al. (1998, 1999) provide more experimental evidence from Kenya, concluding that textbook provision had little effect on test scores apart from the top fifth of students, and that the provision of wall charts, at least in initial analyses, do not seem to improve academic achievement.

Fuller et al. (1997), in a study of education in Brazil, come to similar conclusions to those of Kremer et al. (1997), arguing that allowing class size to "float upwards" and using the savings for additional classroom resources might improve students’ rate of learning without increasing overall expenditures.

Glewwe (1999) reports on a detailed statistical study of test scores and earnings in Ghana. The study finds that a number of school traits are strongly correlated with gains in math and reading scores. In terms of cost effectiveness, provision of blackboards in classrooms represents the best investment, followed by, in order, repairs to leaky roofs and provision of textbooks. Other measures of school inputs, such as average years of teacher experience, training and education, did not bear a significant relation to gains in either test score. Given the small sample sizes involved in the analyses, the finding that any school inputs were significantly related to student learning is impressive. Glewwe and co-authors also show that these test scores are related to workers' earnings (both in the relatively small employee sector and to a lesser extent in the larger self-employed sector), implying that investment in the three resources mentioned above should affect students' earnings later in life.
Overall, then, evidence that "school spending matters" in studies from developing countries is stronger than in studies from the United States, but it is still quite mixed. There is not yet a clear winner in the "quality versus quantity" debate. But overall, evidence from developing countries makes a more compelling case that it is spending on quantity, rather than quality, that is likely to have a stronger and more systematic payback.

There appears to be widespread agreement that improving incentives in schools could do much to improve student performance. However, many practical issues need to be resolved before stricter accountability becomes a reality. For instance, key issues related to potential cultural or gender bias in standardized tests and measurement error in standardized tests require clarification. Only then are parents and educators likely to feel entirely comfortable with using test scores for: (1) making difficult decisions about student remediation and grade retention and (2) making teacher pay contingent upon students’ performance in standardized tests. See Koretz (1996) and Meyer (1996) for detailed discussion of these issues.

Perhaps the policy issue of most far-reaching significance concerns the degree of school choice that families require. Evidence from relatively small-scale voucher programs in the United States provides mixed evidence, but on the whole suggests that vouchers can improve outcomes for disadvantaged students who participate. Evidence from a series of large-scale subsidy or choice programs suggested by the World Bank will in time provide ample evidence about the impact of school choice and inter-school competition for students on both the level and variations in student achievement.

7. Topics for Further Research

Many issues remain partly or fully unresolved. One of the key unresolved statistical issues is the concern that school spending may be correlated with demographic traits of families and communities, for which researchers often have only crude proxies. For example, family and peer group traits appear to be strongly related to students’ own test scores and later earnings. To the extent that existing studies have not fully controlled for these traits, and to the extent that they are positively related with the level of spending in neighborhood schools, past studies may have overestimated the impact of school spending on student outcomes. One approach to this problem has been to move away from aggregated data-sets toward ones that provide detailed demographic data on students, their families, and the overall student body in the schools actually attended by students. A more expensive, but arguably more compelling, solution involves running controlled experiments. See for instance Mosteller, 1995, for an analysis of a controlled experiment involving class size reduction in Tennessee. This experiment suggests that class size reductions can increase student achievement in primary school, but that the bulk of the effect occurs during the first year in which the student enrolls in a smaller class. The aforementioned study by Kremer et al. (1997) provides a smaller scale example from Kenya. Ideally, a long-term experiment could be arranged that would test for persistent effects of school spending by following students into the labor market.
Typically, research into school quality has examined the link between school resources and one or more of the following indicators: test scores, years of schooling obtained and earnings. (This survey has focused on the latter two outcomes.) There are many other ways that one might examine the effectiveness of school spending. After all, preparing students to participate effectively in the labor market represents only one of many rationales for public education. It would be worthwhile to test for a relation between school resources and broader measures of student outcomes, such as job satisfaction, health and civic participation. This point seems particularly applicable to developing counties, where a number of critical intermediary relationships govern the link between school spending and students’ life chances. To give just one example, is there any link between school spending and students’ rates of child-bearing and child-raising? Glewwe (1997) provides a critical review of the small but growing literature on this important question. He concludes that there is considerable evidence that years of schooling are negatively related to women’s fertility. In contrast, very few studies have yet addressed the impact of school quality or school prices on fertility even indirectly, making it difficult to draw any conclusions at this point in time.

On the whole, existing research has focused much more heavily on the question "Does a given type of spending matter?" than on the questions "How big is the effect?" and "What are the most cost-effective reforms?" Betts (1996a) concludes that the overall returns to school expenditures have been rather modest in the United States. In the United States, expenditures on teacher training and lengthening the school year might well prove more effective than reducing class size. As already noted, limited evidence from developing countries hints that spending on facilities and teacher training is likely to be more effective than reducing class size.

One area ripe for further investigation concerns the question of diminishing returns to school spending. Betts (1996a) reviews scattered and inconclusive American evidence on this question. In an unpublished paper, Betts and Johnson (1997) analyze state-level data in the United States and find evidence of strongly diminishing returns to spending on reducing the pupil-teacher ratio, but only mildly diminishing returns and increasing returns to spending on teacher salaries and spending on lengthening the school year, respectively. The fact that a greater proportion of developing-country studies than American studies find that school inputs significantly affect student outcomes implies that an increase in school spending is most helpful in relatively impoverished areas and schools. Given the large variations in school spending across developing countries, this question would seem particularly amenable to study in a developing country context, particularly if it was posed in terms of which types of school spending matter most.

We also have a great deal to learn about the impact of higher standards on student (and teacher) performance. Bishop has done much to piece together work on this subject. Much more can and should be done. Similarly, we need to develop a much better understanding of the role of tracking, ability grouping and curriculum in determining the level of student outcomes, and the degree of variation in student outcomes.
Overall, existing work on the impact of school resources on earnings does not provide an overwhelming case for additional expenditures, especially in American schools. Limited evidence suggests that empowering parents with school choice programs, backed by subsidies to needy families, may have a role to play in school improvement in both developing and developed nations. The current evidence remains scant, but will likely grow considerably in coming years as results emerge from programs in the United States and the many projects supported by the World Bank.

Acknowledgements

The author is Associate Professor of Economics at the University of California, San Diego, and Senior Fellow at the Public Policy Institute of California. He thanks Harry Patrinos and Paul Glewwe for helpful suggestions.

8. References


