The Use of Cost-Benefit Analysis in Guiding Investments in Human Capital in Elementary and Secondary School

August 2006

Clive R. Belfield

clive.belfield@qc.cuny.edu
Economics Department, Queens College,
City University of New York, Flushing, NY 11367

Abstract:
This report reviews the literature on cost–benefit analysis of elementary and secondary school investments. A range of investments are considered: policies such as improving teacher quality, reducing class and school size; programs such as pre-school, Head Start, grade retention, and other specific interventions; and reforms, such as school choice, accountability, and whole-school reform. With the exception of pre-school, the economic evidence is extremely thin, with almost no cost–benefit ratios. Areas that may offer potential for further inquiry are teacher productivity, special education, and the educational needs of immigrant students. Economic analysis can be valuable in clarifying where investments in education should be made.

Report prepared for Professor David Weimer, University of Wisconsin–Madison

1. Introduction

In this paper I draw on the wealth of literature in the economics of education to consider public policy in regard to the costs and benefits of incremental investments in primary and secondary education. I identify the additional investments that are perceived to be socially desirable and offer a general evaluation of the evidence for them. The investments are divided into three types: policies, programs, and reforms. In conjunction, I discuss the most important applications to date of cost-benefit analysis in primary and secondary education. Finally, I suggest investments that might be promoted as socially desirable; these suggestions are based on my reading of the evidence interpreted via a simple value function which incorporates efficiency and equity criteria. I restrict my discussion to the U.S. education system.

To begin, and taking a general overview of the large body of evidence, it is important to set out a number of points about economic analysis of human capital investments.

One conclusion stands out: the private benefits to investing in additional education are substantial.

The private earnings advantage from extra years of schooling is clear. Carneiro and Heckman (2003, 148-149) state that "By now there is a firmly established consensus that the mean rate of return to a year of schooling, as of the 1990s, exceeds 10% and may be as high as 17 to 20%."[1] This return is not significantly biased by endogeneity or confounding and it does not appear to differ much by race, ability level, or gender (Card, 1999; Barrow and Rouse, 2005). This fact has led researchers to conclude that an additional year of attainment would convey benefits to almost any student (particularly for an education system where 20% of students do not graduate from high school).

There are also private earnings advantages from higher achievement: persons with higher test scores earn more. As an approximation, the impact on earnings of a one standard deviation increase in test scores is equivalent to that from having one more year of schooling. But, the magnitude of the gain from higher achievement is open to debate, in part because test score advantages in elementary school do not perfectly correlate with advantages by graduation. Hanushek (2006) argues that the link is strong, reporting four estimates such that the earnings premium from a 1 standard

[1] Labor economists also contend that the supply of educated workers is low, although the impact on wages is the subject of some debate (Lemieux, 2006).

deviation increase in test scores is 12%. Heckman and Vytlacil (2001) contend that the link is weak. Most recent estimates by Rose (2006) show an almost-zero effect for males, with females obtaining a 9% advantage where test scores are 1 standard deviation higher.

Yet, income gains are only a fraction of the private advantages from education. A number of studies have identified both monetary and non-monetary advantages from being more highly educated (for examples, see Wolfe and Zuvekas, 1997; for the range of powerful health-related effects, see Cutler and Lleras-Muney, 2006).[2] At issue is the cost of generating achievement and attainment gains, not whether there will be significant economic benefits from so doing. Given the high levels of public subsidy for education, private economic calculations are relatively simple: a successful investment in elementary or secondary education should yield a net positive return.

However, economic analyses may yield different conclusions depending on whether the public or private perspective is adopted. From the public perspective, the costs are higher (schools are largely publicly funded) and the benefits are lower (income tax gains rather than income gains).[3] But, this distinction (high private / lower public) is partly because crime and welfare effects are not counted (see Lochner and Moretti, 2004). The public benefits may be as high as or higher than the private benefits (for the social costs of a high school dropout, see Cohen, 1998). Also, there is no consensus on whether to refer to the strictly fiscal gains or the social gains (i.e. public and private) for policy purposes. However, most research at least supports current levels of public funding for schools.[4] More often, the tendency is for proposals to increase public funding, if only to meet equity goals such as equitable or adequate funding for all families.

Beyond these conclusions, it is necessary to acknowledge that cost–benefit analysis and research on the economic value of specific educational investments is far

[2] Offsetting these additional benefits are the illiquidity and irreversibility of human capital investments. However, little is known about how important either of these factors is.
[3] The evidence of public externalities from increased education levels (e.g in raising economic growth) is also not strong (Taylor, 1999).
[4] Libertarian economists might argue that – given the weak progressivity of the tax code and the fact that parents are rational agents and leaving aside some redistribution on equity grounds – income transfers or tax relief should be given to parents so that they can invest in education independently. However, the effect of short-run changes in income on educational attainment is not especially strong. The effect of an increase in permanent income is somewhat stronger, but not dramatically larger than the effects of educational interventions (Taylor et al., 2004). The effect of socioeconomic status on educational outcomes is strong, but status encompasses more than income. Plus, many parents are credit-constrained (at least when their children are young).

from compelling. Put simply, the available evidence does not give much guidance on what investments are optimal under a given set of circumstances (Mervis, 2004).[5]

One reason is that economic analysis is often not wanted. Policymakers (and ideologues) do not want studies that undermine reforms that look effective but might not be cost-effective. Another reason is that genuine economic analysis is often not done.[6] Levin (2001) describes as "rhetorical" the methods used in the vast majority of educational studies that claim to be about cost-effectiveness, i.e. they make claims with no data on costs or effects. Most studies focus on the impacts of an educational investment (King Rice, 2002), presuming that positive impacts will mean positive net returns.

In defence, researchers might perhaps argue that cost–benefit analysis is more complex in practice than in theory (with the latter being clearly set out in Levin and McEwan, 2002).

Costs data are often confidential, and information on marginal costs is often unavailable. Some resources are provided 'in-kind' and from various funding sources. Also, the costs of specific educational interventions will differ across the sites where they are implemented. A year of attainment can be approximately bounded at $6,000-$10,000 (at least as an average cost). But the cost of raising achievement scores by one standard deviation is unknown; it may vary dramatically according to how it is produced.

Identifying impacts is also problematic. Educational programs have diverse benefits, including cognitive and non-cognitve skills (Heckman et al., 2006), and weighing their relative values is not straightforward. Educational outcomes are primarily a function of family background, with school effects secondary (Levin and Belfield, 2002). Again, marginal impacts are needed but average impacts are most commonly

---

[5] Three quotations from economists are illustrative of the thinness of the research base:

> "Estimating the costs of achieving improvements in the teacher force is generally impossible based directly on current data. We simply have limited experience with any policies that alter the incentives for hiring and retaining high quality teachers (and which also evaluate the outcomes)" (Hanushek, 2006, 459).
>
> "There is no shortage of policy proposals. There is, however, a shortage of empirical evidence on the efficacy of the proposed policies" (Carneiro and Heckman, 2002, 87).
>
> "The evidence generated from local providers may fall far short of standards for convincing evidence… raising questions about the decision-making of local providers based on this evidence, and the quality of the evidence provided to policymakers" (Neumark, 2006, 315).

[6] This is also the case at the supranational level. On the less-than-complete economic analysis performed by the World Bank, see Vawda et al. (2003).

the only ones available. Resources to perform research trials on education programs – to ensure all biases are adjusted for – are often not available (Mosteller and Boruch, 2002). And the effects of such interventions will vary across student groups and depend on whether students are willing to participate (Grissmer, 2002). An effective program may easily be ineffective if it is not implemented properly.[7] A targetted program will be much more effective than a universal one, delivered to students who would have prospered regardless. But programs may vary in how accurately they are or can be targetted.

Consequently, there are very few studies with clearly reported rates of return or net present values. And there are even fewer where different investment opportunities are compared directly (exceptions being Aos et al., 2004; Levin et al., 1987). Disappointingly, there is not only a paucity of rigorous cost–benefit studies but also uncertainty as to the external validity and generalizability of any findings.[8] This leads to the unavoidable conclusion that the selection of educational investments is based on criteria other than whether they yield a high rate of return.

For some, the failure to apply economic criteria may be desirable. But most evaluations do not make their values implicit. This hinders the application of cost-benefit analysis. If, say, one places a heavy weight on raising the academic achievement of disadvantaged children, then this might justify public funding for a range of programs even if they are high cost. Their achievement is valued at more than the cost of raising funds from taxpayers. Indeed, most educational investments appear to be advocated for on equity grounds, not efficiency grounds. But, without any information on the costs of such programs it is impossible to say how much weight is being placed on raising achievement levels for disadvantaged children. Efficiency and equity are not substitute criteria.[9]

Notwithstanding these caveats, the review below offers some conclusions about which elementary and secondary school investments might be socially desirable using

[7] Case studies of career academies and high-school reforms show significant variance in implementation (an early example is given in Stern et al., 1989).
[8] For example, research on the High Scope / Perry Pre-school program, which scores well on many methodological criteria, only includes samples of at-risk African American children in Michigan in the 1960s. The pertinence of any findings to, say, an early education program for immigrant children in California in the 2000s is an open question.
[9] Economic analysis depicts efficiency and equity as subject to a trade-off. For example, if one adopts a simple cost–benefit approach where the highest return investment is selected, from the public perspective greater educational investments should be made for boys over girls. Boys commit more crime, earn more and work more (and so pay more taxes). Yet, most educational programs are gender-neutral.

cost–benefit analysis. I then consider what other investments might appear to offer as high or higher returns.

2. Desirable Investments in Human Capital

In this review of investments that may pass cost-benefit tests, three types are distinguished: policies, programs, and reforms. Policies are changes that influence the technology of education in a broad way and are not linked to a specific program or template, such as reducing class size. Programs are more narrowly focussed interventions or treatments that could be applied to some children or in some settings, such as summer school. Reforms are systematic and large-scale changes to the educational organizations or institutions with the intention of affecting many aspects of the education process. Examples of reforms would be rigorous exit-based exams or school choice.

In reviewing the empirical evidence, studies are appraised according to methodology. Experimental research is given priority but other methods are included (otherwise, the review would be extremely brief). Quasi-experimental research is considered more valid than controlled observational studies. Consequently, this review is general, rather than a narrow investigation of a few promising interventions (see also Heckman, 2000).

2.1 Policies

Review of the literature shows several policies which are commonly cited as desirable educational investments.

(i) Raising teacher quality

Research evidence clearly establishes that some teachers are more effective than others at raising student achievement (Wayne and Youngs, 2003; King Rice, 2003) and that the cumulative effect of having more effective teachers over the K-12 years is significant (Hanushek and Rivkin, 2004; Nye et al., 2004; Rivkin et al., 2005). From this evidence, researchers have concluded that investing in high quality teachers is a worthwhile investment.

However, it is not certain that investing in high quality teachers would pass a cost–benefit test. The characteristics of high quality teachers are not easily identifiable and so the practical policy is to pay teachers more as part of their baseline salary.[10]

[10] Performance-related pay has been less than successful, for various reasons (Ballou, 2001). Allocating teachers to schools is also far from straightforward (Lankford et al., 2002).

Presumably higher salaries will generate a more talented teaching force (Temin, 2003). Loeb and Page (2000) estimated the effects of raising teacher salaries on high school graduation using state data with a ten-year time lag. A ten percent increase in teacher salaries across the K-12 years would raise the high school graduation rate by five percentage points. Given the economic benefits of graduating from high school, this investment is likely to pass a cost-benefit test. An alternative to paying teachers more is to hire better quality teachers and fire lower quality ones (or wait for them to retire). Hanushek (2006) estimates this alternative would generate significant economic benefits, even as it would take a long time for the teaching pool to become high quality.

(ii) Reducing class size

Perhaps the most commonly mentioned educational policy is that of reducing class size. The canonical reference is to the Tennessee Project STAR (Mosteller, 1995). Students were randomly assigned to larger classes that averaged 22 students to smaller ones that averaged about 15 students for up to four years duration, kindergarten to third grade. Tests were administered on word skills, reading, and math. Longitudinal follow-up of students found that those in smaller classes for more years had higher test scores and were more likely to graduate from high school than students assigned to larger classes. The largest effects were found for minorities and students from the lowest socioeconomic backgrounds (Finn and Achilles, 1999; Finn et al., 2005).[11] Implementing this class size reduction policy would increase the high school graduation rate by 11 percentage points, suggesting that the policy would pass a cost–benefit test.

An economic evaluation by Krueger (1999, 2002) finds that the internal rate of return to Project STAR is around 6%. This is above a threshold rate of return that might be used to evaluate public projects (Moore et al., 2004). However, it is a social rate of return (adding public and private returns together), although it only counts labor market returns. Other economists have questioned whether the estimate of 6% is overly optimistic (Carneiro and Heckman, 2002).[12] The organizational changes required to

[11] Effect size gains across the subjects were 0.15-0.26 across kindergarten and 3rd grade. After kindergarten, students in smaller classes were 0.5/1.6 months ahead in reading/math. For 1st grade, the effects were 1.3/2.8 months. For 2nd grade, the gain was 3.9/3.5 months; and 4.5/2.6 months for 3rd grade. The effects are cumulative, i.e. each year of smaller classes produces a boost in test scores (Nye, 2001). Minority students gain most from smaller classes in reading (but not math). Girls gain most in math (but not reading). The effects do not vary by ability (Nye et al., 2002). After fifth grade, those in smaller classes were about five months ahead.

[12] A national analysis of the cost implications of reduced class size by Brewer et al. (1999) found that costs would increase by approximately $250 per student (2005 dollars) to reach class sizes of 20 or $1,400 to reach class sizes of 15. Compared to average per-pupil spending, this

simulate the circumstances of the STAR treatment are extremely large. (The experiment itself cost $12 million). For a school district, going from 22 students per class to 15 students per class would increase the numbers of classes by 47%, with a corresponding increase in expenditures. (Expenditures may not rise proportionately: some costs are independent of class size. However, most costs are for teachers' salaries; if these costs are only 60% of total costs, expenditures will still be increased by 27%). This ratio does not account for the strong likelihood that, in order to attract extra teachers of equivalent quality, salaries will have to rise.[13] It does not include planning and administration costs. Finally, this change must be sustained over multiple years: the average duration for a participant in STAR was 2.3 years.

Researchers are divided as to whether such an investment is worth making, both in absolute terms and relative to other educational investments. The Krueger estimate of 6% is a moderately good return on investment, although such a return should be considered in light of the magnitude of the initial outlay. And, the costs are incurred by the school district (public costs) and the benefits are reaped by the students (private benefits). Because benefits are greater for students from more disadvantaged backgrounds a targeted class size reduction would appear most efficient. In contrast, Prais (1996) has argued that the academic gains from STAR could be achieved by adding 2-3 extra days to the academic year, and that this would be considerably less expensive than reducing class size. A simpler comparison is that class size reduction would raise achievement by approximately half a year in grade equivalence, but require approximately 1.1 years of expenditures (=2.3$\times$0.47). Relative to other investments, reducing class size appears to be expensive.

(iii) Reducing school size

A newly popular educational policy is to reduce school size. The rationale is that students are not engaged and motivated to learn in large, impersonal settings. Academic achievement will be enhanced through higher quality and more focused instruction, increased student engagement, and more parental involvement. However, there is very little evidence that small schools are more effective than larger ones,

---

represents a 3% to 18% increase in funding. This study only considered operational costs (i.e. instruction), and not facilities costs (i.e., reorganization to accommodate more classrooms). It also assumed new teachers could be hired without raising wages. These assumptions understate the costs of class size reduction.

[13] Finding new teachers may be expensive; in California's class size reduction initiative newly recruited teachers were less experienced and less qualified (Ogawa et al., 1999).

controlling for factors such as urbanicity or demographics (Darling-Hammond et al., 2006).

Existing evidence shows inconsistent results, but most of this is from simple cross-sectional comparisons which do not control for endogeneity. However, a high-quality study for elementary schools in Indiana by Kuziemko (2006) finds statistically significant academic advantages for smaller schools. Kuziemko (2006, Tables 5-7) reports larger schools have lower attendance rates and lower math scores and slightly lower language scores. However, doubling the size of each elementary school would reduce math scores by only 4 percentage points. Whether or not this would pass a cost–benefit test will depend on the economies of scale, i.e. whether large schools have lower or higher unit costs.

2.2 Programs

Programs are educational interventions involving a specific technology and implementation plan. They may be large-scale or small-scale.[14]

(i) Publicly-funded pre-school

Expanding pre-school provision is possibly the most compelling investment on economic grounds. The evidence is based on high-quality research methods with full cost–benefit analyses from both the private and public perspective; and it is almost completely consistent in identifying impacts (see the review of evidence in Barnett and Belfield, 2007). The three most frequently cited programs are the High Scope/Perry Pre-School program, the Chicago Child–Parent Centers, and the Abecedarian program (Belfield et al., 2006; Reynolds et al., 2002; Barnett, 2002). Each yields returns over the lifetimes of participants that easily exceed costs. The Perry program generates total benefits that are 8.8 times the costs, with public benefits that are 7.2 times the costs. Ratios for the Chicago program are 10.1 (total benefits) and 6.9 (public benefits). And ratios for the Abecedarian program are 3.8 and 2.7 respectively.

Each of these model programs was targetted to at-risk children, but cross-sectional research indicates that all children benefit from pre-school (Magnuson and Waldfogel, 2005). Even if the benefits for advantaged families are dramatically below

[14] I have not considered interventions either that are behavioral (e.g. Big Brothers / Big Sisters), but are education-related insofar as they may raise achievement or attainment, or that are school-based but not strictly educational (e.g. drugs education classes). On the latter, see Aos et al. (2004) for a full tabulation. But, if these reduce public expenditures on crime or health they may be more socially desirable than interventions that raise achievement (see Caulkins et al., 2004).

those for at-risk children, there would still be economic gains from expanded pre-school programs (Karoly and Bigelow, 2005).

(ii) Head Start

Head Start, the federal government's largest comprehensive child development intervention, is intended to improve children's cognitive, social, emotional, and physical development, as well as offer support to parents. Early analyses raised the possibility that Head Start would have only transitory effects (Currie and Thomas, 1995). The best estimates of its effects on children's cognitive and socio-emotional development are provided by a recent national randomized trial of Head Start. The magnitude of estimates for immediate effects of one year of Head Start is fairly small, from less than 0.10 to 0.24 for standardarized measures of language and cognitive abilities. This echoed the findings from the Early Head Start randomized trial in which cognitive and language effects were about 0.10, or smaller. Both Head Start and Early Head Start randomized trials yield small decreases in anti-social behavior, about 0.10 in size, and there is no evidence at all in Head Start studies of negative effects on social and emotional development.

Importantly, Head Start does appear to generate long-term gains. Garces et al. (2002) found increased rates of high school graduation and college attendance for whites, with no clear effect on earnings by age 23. African American children in Head Start were less likely to be booked or arrested for a crime. However, it is not clear whether these benefits are sufficient to offset the costs of the Head Start.[15]

(iii) Other programs

A range of other programs have been considered.

One example of a potentially promising program is summer school (as suggested by Carneiro and Heckman, 2002). Summer school's promise is based partly on the same arguments used for more attainment and partly on the fact that children 'lose ground' during the time away from school. However, the economic case for summer school investments has not been established. Using an experimental field trial in Baltimore, Borman and Dowling (2006) show that summer school is effective: after two successive summer schools, the treatment group is approximately 0.5 standard deviations ahead of the control group in test scores. These effects are close to the

[15] A recent paper by Ludwig and Miller (2006) shows slight improvement in attainment from participation in Head Start. However, it also identifies an improvement in mortality which, given the economic value of a life, be sufficient to justify investment in Head Start.

effect size of 0.19 reported in a meta-analysis of summer school by Cooper et al. (2000). Annual costs for the Baltimore program are estimated at $815 per student, with an additional $700 in in-kind resources. Back-of-an-envelope calculations suggest that this program is cost-effective.[16]

A second example is grade retention. Here, perhaps, there is the most agreement: grade retention is a very inefficient investment. It imposes additional costs on a school system. And it generates adverse effects on the retainees (Temple et al., 2003); a recent study by Hong and Raudenbush (2005) finds that those held back in kindergarten learn less than if they had been promoted (see also the review by Jimerson, 2001).

A final example is peer tutoring, which scores well on cost-effectiveness grounds. Wolfe and Tefft (2004) list this as a 'most promising' intervention in their review. However, this intervention is primarily attractive because it encourages children to teach others and so imposes very low public costs. If children's time as tutors is valued at zero, any peer-tutoring program that is effective will yield a very high rate of return.

2.3 Reforms

Policies and programs are proposed with the expectation of increased funding. In contrast, reforms are proposed with the intention of changing how existing resources are allocated. Such reforms are predicated on the belief that simply throwing more resources at the existing system will not yield high returns. Rather, organizational and institutional change is necessary.[17]

(i) School choice and competition

School choice and competition are reforms that seek to foster markets for educational services. Believers in school choice argue that it is a revolutionary change to the education system and that the effects will be dramatic (Hoxby, 2003).

However, the effects of school choice and competition, although positive, have been found to be only modest (Teske and Schneider, 2001). Positive but very moderate results have been found from introducing competitive pressures into education markets (Belfield and Levin, 2002). Voucher programs also yield only slight achievement gains

[16] Costs for two years would be $3,000 with benefits of 0.5 standard deviations. Using the metric of 1 standard deviation equal to one year of attainment, spending is $3,000 to obtain 0.5 of a year of attainment. In unit costs, the program is about the same cost as extra attainment, which we have argued is a good investment if it can be generated.

[17] To quote again from Carneiro and Heckman (2002, 159): "Marginal improvements in school quality are likely to be ineffecitve in raising lifetime earnings and more fundamental changes are required if we hope to see a significant improvement in our educational system".

for participants.[18] The randomized field trials for vouchers in New York, Dayton, and Washington DC found small test score gains after three years (Howell and Peterson, 2002, Table 6-1). Yet, the gains are only reported for the treatment group of students who use the voucher (not those offered it). There are no effects in New York and Dayton, with achievement gains for voucher users in Washington DC in the second (of three) years, and the effects are not cumulative (rising with the duration of participation). However, there are strong impacts for African American children across all three years in New York and the second year in Dayton and in Washington.

Most evidence from expanded public school choice points to the same conclusion of small achievement gains from placement in a choice school. The Chicago lotteries analyzed by Cullen et al. (2005, Table 6) show no gains from winning the lottery in terms of any of the following educational outcomes across ninth and tenth grades: dropping out, reading, algebra, English, geometry, course credits, and absences; but they do find lower involvement with the criminal justice system. (For charter schools, there is no gain and possibly even some fall in achievement).

Even as competition and choice conveys slight positive benefits, these may not be sufficient to justify the costs of re-organization. These costs are not easily identified. Typically, the voucher payments are lower than the per-student expenditure in the schools. Thus, if there are any positive benefits, voucher programs would immediately pass a cost–benefit test. However, there are other costs as well as the voucher payment that need to be taken into account and these will reduce the returns to voucher reforms (see Levin and Driver, 1997).

(ii) Accountability standards and exit exams

One reform that is typically considered to be a good investment is the introduction of rigorous exit-based exams. This reform appears attractive because it might be low cost: schools already impose some form of assessment so a replacement should not be expensive, and tougher exams mean students will have to work harder (and their time is not a cost to the public purse).

In their evaluation of the A+ Accountability Program in Florida, Figlio and Rouse (2005) find modest results. Using data on over 180,000 students, they compare the

[18] Compared to a random sample of Milwaukee school children and a low income sample, Witte (1999) finds no achievement gains for voucher users from the Milwaukee Parental Choice Program (Greene et al. (1998) find positive impacts). Rouse (1998) compared voucher users with those applicants who were not offered a voucher (by chance); this analysis also compensated for continuation in the program (as well as student fixed effects). Voucher users show small but positive effect size gains of 0.08-0.12 per year for math but zero for reading.

performance of students in schools eligible for vouchers against students in schools that just avoided being eligible (i.e., the schools were graded F in one or two subjects rather than all three). They find achievement gains for voucher students from low-performing schools but with much of the improvement attributable to teaching to the high-stakes test (and to student characteristics). An analysis of state-level scores by Hanushek and Raymond (2005) also finds positive impacts from stricter accountability regimes.

Accountability frameworks may therefore help in raising achievement for some students. However, the benefits of exit-based exams vary significantly across student groups (Dee and Jacob, 2006). Plus, imposing exit-based tests may discourage students from accumulating attainment, reducing education levels for those who expect to fail the test. The net effect is likely to be gains for those pushed to study harder and losses for those who drop out early. Importantly, Dee and Jacob (2006) find only very small impacts on labor market outcomes (earnings and employment) as a result of stronger exit-based exams. Thus, the aggregate effect is likely to be small.

(iii) Whole-school reform

Whole-school reforms are often advocated as a way to change the culture and organization of a school to ensure greater learning. Overall, economic analysis of whole-school reform is incomplete, despite the substantial cost involved in implementing them. Many of the challenges in conducting economic analysis (identified above) are especially pertinent to whole-school reform (see Levin, 2002).

One example of whole-school reform which has been evaluated is Success for All. This reform focuses on promoting early school success among educationally at-risk students. Success for All includes materials, training, and professional development to implement a schoolwide program for grades K-5 to ensure every child will reach third grade on time. It serves approximately 1 million children in 2,000 schools. The evaluation by Borman and Hewes (2002) shows Success for All may be a good investment because it raises test scores by 8th grade, reduces special education placement, and reduces rates of grade retention. With a per student investment of approximately $3,100 over four years of schooling, the effect size gains in reading and math were 0.29 and 0.1 respectively.

An example of reform at the high school level is First Things First, which emphasizes small learning communities, instructional improvement, and teacher advocacy for each student (Quint et al., 2005). Small learning communities require that schools or sub-units of schools with which students and faculty are affiliated are limited

to no more than 350 students. Key teachers remain together for several years and each student is matched with a staff member who meets with the student regularly, monitors student progress, and works with parents to support student success. Instructional improvement focuses on high expectations and rigorous curriculum. In a research study using interrupted time-series data First Things First generated higher graduation rates as well as benefits in terms of student attendance and test scores in mathematics and reading. The high school graduation rate was increased by 16 percentage points as a result of the intervention. Although costs data are not available, this improvement in graduation rates is sufficiently large that this reform is likely to pass a cost–benefit test.

3. Important Educational Investments

On this reading of the literature, there appears to be a reasonable evidence base justifying investments in high-quality pre-school, with other investments described as 'promising'. Beyond this, researchers have focussed on identifying a common set of features that lead to increased high school graduation rates and educational success. These features are: (1) small school size where students and staff are known to each other; (2) high levels of personalization such that students' personal and academic needs are addressed; (3) high academic expectations with a rigorous curriculum and assessment; (4) strong counseling for students facing personal and educational challenges; (5) parental engagement to support school programs; (6) extended-time school sessions so that learning is maintained; and (7) competent and appropriate personnel with both credentials and commitment to the school and its mission. However, the cost of educational programs with such features is unknown.

Therefore, it is not appropriate to give a list of desirable investments as this would be mere speculation. But it is reasonable to speculate about areas where researchers might find high returns. This approach to identifying desirable investments is to look for 'low-hanging fruit', i.e. where a lot of money is spent but relatively little evaluation is performed. An alternative approach is to consider investments which force students to work harder. From the public perspective this is very low cost because students are not compensated for their effort.

Clearly, a lot of educational investment must be spent on teaching personnel, so reforms to the teaching profession should be investigated further. As well as research on how better teachers raise student achievement, further investigation should look at

what makes teachers more productive (e.g. how absenteeism and turnover rates can be reduced, or how job satisfaction can be enhanced).

A second priority area is special education. Approximately 15% of students are in special educational programs and the vast majority of these students do not have severe disabilities. Special education expenditures represent one-quarter of all public expenditures. Almost no research is conducted on which types of special education services are most cost-effective or on whether students are appropriately enrolled in special education programs. Yet one of the arguments for pre-school is that it reduces placement in special education programs, yielding a considerable public saving which alone offsets about half the cost of pre-school (Conyers et al., 2003). It is unlikely that pre-school is the only intervention with such impacts.

Finally, it is necessary to recognize that the returns will vary across different student groups and that those with the lowest education levels are likely to benefit the most from additional educational investments. The U.S. population now has a large proportion of immigrants with relatively low education levels (a large fraction have not completed high school and a non-trivial proportion have less than 9th grade schooling by age 20). Given the high economic returns to basic skills, more research should be performed on interventions that help those with very low levels of education.

4. Conclusion

Although the above review does not identify many policies, programs, or reforms that have support from economic evaluation, some conclusions are possible.

First, a general presumption is that it is more efficient to invest early (pre-school and elementary school) rather than late (high school and remedial training). This presumption has two parts. One is that, simply abstracting from the technology of education and how public funds are allocated, the returns will be higher for early investments. Human capital accumulation is dynamic such that higher level skills cannot be obtained without the foundation of earlier, lower level skills. The other is that many researchers would argue that existing investments are not sufficiently allocated to the early years (Carneiro and Heckman, 2003, 90). These presumptions give further credibility to pre-school programs, which currently have some of the strongest (and positive) economic evaluations.

Second, educational investments must have long-term impacts in order to yield a positive return. This persistence has been established for years of attainment, which are

correlated with earnings over decades. But, the persistence of achievement gains has not been fully established. A program which boosts test scores in elementary school by one standard deviation may not necessarily convey impacts in later adulthood. Also, higher returns are obtained if there impacts across multiple dimensions beyond the labor market (such as health effects or welfare dependency). Although public investments in education can pass a cost-benefit test based only on the income tax gains, the margin for investment is considerably smaller if this is the only source of gain.

Third, given the empirically identified economic benefits to the state of high school graduation (higher tax payments, lower expenditures on criminal justice, health, and welfare), a large fraction of investments that are effective in this domain should be cost-effective. This is not a 'blank check' to invest in programs to reduce the number of drop-outs, but there is probably a lot of room to be 'generous' in public funding of education and to experiment with multiple approaches simultaneously (e.g. smaller classes with high quality teachers).

Fourth, given the heterogeneity in students' learning proficiency, it is very unlikely that a single type of investment will yield the highest returns in each situation. Some economists such as Hanushek (2004) have gone further, raising the question: What if there are no 'best practices'? But, this appears to confuse the search for the most efficient investments with the search for investments that yield a positive return. Some practices will be better than others. It is the case that little is known about whether to invest in new programs or simply to invest more in existing programs. But, it is likely that the highest returns will be found in situations where students currently have the least resources.

Finally, improvements should be made in the economics of evaluating education reforms. Cost-benefit analysis is an important evaluative method. Indeed, it may be a superior method, given the ad hoc elements included in other methods (for a discussion, see Adler and Posner, 2000). More cost-benefit analyses should be performed and more economic metrics (such as the monetary value of achievement gains) should be developed and systematized. Such analysis is not a substitute for making decisions on equity grounds.

References

Adler, MD and EA Posner. 2000. *Cost-Benefit Analysis. Legal, Economic and Philosophical Perspectives*. Chicago: University of Chicago Press.

Aos, S et al. 2004. *Benefits and Costs of Prevention and Early Intervention Programs for Youth*. Washington State Institute for Public Policy, www.wsipp.wa.gov/rptfiles/04-07-3901.pdf

Ballou, D. 2001. Pay for performance in public and private schools. *Economics of Education Review*, **20**, pp.51-61.

Barnett, WS and CR Belfield. 2007. Early childhood education. *Future of Children*, forthcoming.

Barrow, L and CE Rouse. 2005. Do returns to schooling differ by race and ethnicity? *American Economic Review*, **95**, 211-218.

Belfield, CR and HM Levin. 2002. The effects of competition on educational outcomes: a review of the U.S. evidence. *Review of Educational Research*, **27**, 279-341.

Belfield, CR, Nores, M, Barnett, WS and L Schweinhart. 2006. The Perry Pre-School 40-year follow-up cost–benefit analysis. *Journal of Human Resources*, **42**, 215-246.

Borman, GD and GM Hewes. 2002. The long-term effects and cost-effectiveness of Success for All. *Educational Evaluation and Policy Analysis*, **24**, 243-266.

Borman, GD and NM Dowling. 2006. Longitudinal achievement effects of multiyear summer school: Evidence from the Teach Baltimore randomized field trial. *Educational Evaluation and Policy Analysis*, **28**, 25-48.

Brewer, DJ, Krop, C, Gill, BP and R Reichardt. 1999. Estimating the costs of national class size reductions under different policy alternatives. *Educational Evaluation and Policy Analysis*, **21**, 179-192.

Card, D. 1999. Causal effect of education on earnings. In O Ashenfelter and D Card (Eds) *Handbook of Labor Economics*. Amsterdam: North-Holland.

Carneiro, P and JJ Heckman. 2002. Human capital policy. In JJ Heckman and A Krueger (Eds) *Inequality in America. What Role for Human Capital Policies?* MIT Press: Cambridge, MA.

Caulkins, JP, Pacula, RL, Paddock, S and J Chiesa. 2004. What we can and cannot expect from school-based drug prevention. *Drug and Alcohol Review*, **23**, 79-87.

Cohen, MA. 1998. The monetary value of saving a high-risk youth. *Journal of Quantitative Criminology*, **14**, 5-33.

Conyers, LM, Reynolds, AJ, and S-R Ou. 2003. The effects of early childhood intervention and subsequent special education services: Findings from the Chicago Child-Parent Centers. *Educational Evaluation and Policy Analysis*, **25**, 75-95.

Cooper, H, Charlton, K, Valentine, JC, and L Muhlenbruck. 2000. Making the most of summer school: A meta-analytic and narrative review. *Monographs of the Society for Research in Child Development*, **65**, serial no. 260.

Cullen, JB, Jacob, BA and S Levitt. 2005. The effect on school choice on student outcomes: Evidence from randomized lotteries. *Journal of Public Economics*, **89**, 729-760.

Currie, J and D Thomas. 1995. Does Head Start make a difference? *American Economic Review*, **85**, 341-364.

Cutler, D., and A Lleras-Muney. 2006. Education and health: evaluating theories and evidence. NBER Working Paper.

Darling-Hammond, L, Milliken, M, and P Ross. 2006. High school size, structure, and content: What matters for student success? Working paper, Brookings Institution.

Dee, T and B Jacob. 2006. Do high school exit exams influence high school achievement and labor market performance? NBER Working Paper, #12199.

Figlio, D and CE Rouse. 2006. Do accountability and voucher threats improve low performing schools? *Journal of Public Economics*, **90**, 239-255.

Finn, JD and CM Achilles. 1999. Tennessee's class size study: Findings, implications, misconceptions. *Educational Evaluation and Policy Analysis*, **21**, 97-109.

Finn, JD, Gerber, SB and J Boyd-Zaharias. 2005. Small classes in the early grades, academic achievement, and graduating from high school. *Journal of Educational Psychology*, **9**, 214-223.

Garces, E, Thomas, D, and J Currie. 2002. Longer-term effects of Head Start. *American Economic Review*, **92**, 999-1012.
Greene, JP. Peterson, PE, and J Du. 1998. School choice in Milwaukee: A randomized experiment. In Peterson, PE and BC Hassel (eds). 1998 *Learning from School Choice*. Brookings Institution, Washington.
Grissmer, D. 2002. Cost-effectiveness and cost-benefit analysis: The effect of targetting interventions. In Levin, HM and PJ McEwan (Eds) *Cost-Effectiveness Analysis and Educational Policy*. AEFA Yearbook: Eye on Education, Larchmont, NJ.
Hanushek, EA and ME Raymond. 2005. Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, **24**, 297-327.
Hanushek, EA and SG Rivkin. 2004. How to improve the supply of high-quality teachers. In D Ravitch (Ed.) *Brookings Papers on Education Policy 2004.* Washington, DC: Brookings Institution Press.
Hanushek, EA. 2004. What if there are no 'best practices'? *Scottish Journal of Political Economy*, **51**, 156-172.
Hanushek, EA. 2006. Alternative school policies and the benefits of general cognitive skills. *Economics of Education Review*, **25**, 447-462.
Heckman, JJ and E Vytlacil. 2001. Identifying the role of cognitive ability in explaining the level of and change in the return to schooling. *Review of Economics and Statistics*, **83**, 1-12.
Heckman, JJ, Stixrud, J and S Urzua. 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. NBER Working Paper, w12006.
Heckman, JJ. 2000. Policies to foster human capital. *Research in Economics*, **54**, 3-56.
Hong, GL and SW Raudenbush. 2005. Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics. *Educational Evaluation and Policy Analysis*, **27**, 205-224.
Howell, WG and PE Peterson. 2002. *The Education Gap. Vouchers and Urban Public Schools.* Washington: Brookings.
Hoxby, CM (Ed.). 2003. *The Economics of School Choice.* University of Chicago and NBER Press: Chicago.
Jimerson, SR. 2001. Meta-analysis of grade retention research: Implications for practice in the 21st century. *School Psychology Review*, **30**, 420-437.
Karoly, L and JH Bigelow. 2005. *The Economics of Investing in Universal Pre-School Education in California.* RAND, Pittsburgh, PA.
King Rice, J. 2002. Cost analysis in education policy research: A comparative analysis across fields of public policy. In Levin, HM and PJ McEwan (Eds) *Cost-Effectiveness Analysis and Educational Policy*. AEFA Yearbook: Eye on Education, Larchmont, NJ.
King Rice, J. 2003. *Teacher Quality: Understanding the Effectiveness of Teacher Attributes.* Washington, DC: Economic Policy Institute.
Krueger, AB. 1999. Experimental estimates of education production function. *Quarterly Journal of Economics*, **114**, 497-532.
Krueger, AB. 2002. Inequality: Too much of a good thing? In JJ Heckman and A Krueger (Eds.) *Inequality in America. What Role for Human Capital Policies?* MIT Press: Cambridge, MA.
Kuziemko, I. 2006. Using shocks to school enrollment to estimate the effect of school size on student achievement. *Economics of Education Review,* **25**, 63-75.
Lankford, H, Loeb, S and J Wyckoff. 2002. Teacher sorting and the plight of urban schools. *Education Evaluation and Policy Analysis*, **24**, 37-62.
Lemieux, T. 2006. Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *American Economic Review*, **96**, 461-498.
Levin, HM. 2001. Waiting for Godot: Cost-effectiveness analysis in education. In RJ Light (Ed.) *Evaluations that Surprise.* Jossey-Bass: San Francisco.
Levin, HM. 2002. Issues in designing cost-effectiveness comparisons of whole-school reforms. In Levin, HM and PJ McEwan (Eds) *Cost-Effectiveness Analysis and Educational Policy*. AEFA Yearbook: Eye on Education, Larchmont, NJ.
Levin, HM and CR Belfield. 2002. Families as contractual partners in education. *UCLA Law Review*, **46**, 1799-1824.

Levin, HM and C Driver. 1997. Costs of an educational voucher system. *Education Economics*, **5**, 303-311.
Levin, HM, Glass, GV and GR Meister. 1987. Cost-effectiveness of computer-aided instruction. *Evaluation Review*, **11**, 50-72.
Levin, HM and PJ McEwan. 2001. *Cost-Effectiveness Analysis* (2nd Edition). New York: Sage Publications.
Levin, HM and PJ McEwan (Eds). 2002. *Cost-Effectiveness Analysis and Educational Policy*. AEFA Yearbook: Eye on Education, Larchmont, NJ.
Lochner, L and E Moretti. 2004. The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *American Economic Review*, **94**, 155-189.
Loeb, S and ME Page. 2000. Examining the link between teacher wages and student outcomes: the importance of alternative labor market opportunities and non-pecuniary variation. *The Review of Economics and Statistics*, **82**, 393-408.
Ludwig, J and DL Miller. 2006. Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *Quarterly Journal of Economics*, forthcoming.
Magnuson, KA and J Waldfogel. 2005. Early childhood care and education: Effects on ethnic and racial gaps in school readiness. *Future of Children*, **15**, 169-188.
Masse, LN and WS Barnett. 2002. A benefit-cost analysis of the Abecedarian Early Childhood intervention. In Levin, HM and PJ McEwan (Eds). *Cost-Effectiveness and Educational Policy*. Eye on Education: Larchmont, NJ.
Mervis, J. 2004. Education research – Meager evaluations make it hard to find out what works. *Science*, 304, 1583.
Moore, MA, Boardman, AE, Vining, AR, Weimer, DL and DH Greenberg. 2004. Just give me a number! Practical values for the social discount rate. *Journal of Policy Analysis and Management*, **23**, 789-812.
Mosteller, F and R Boruch (Eds). 2002. *Evidence Matters: Randomized Trials in Education Research.* Brookings Institution: Washington, DC.
Mosteller, F. 1995. The Tennessee study of class size in the early school grades. *The Future of Children*, **52**, 113-127.
Neumark, D. 2006. Evaluating program effectiveness: A case study of the School-to-Work Opportunities Act in California. *Economics of Education Review*, **25**, 315-326.
Nye, B, Hedges, LV, and S Konstantopoulos. 2001. The long term effects of small classes in early grades: Lasting benefits in mathematics achievement at grade 9. *Journal of Experimental Education*, **69**, 245-257.
Nye, B, Hedges, LV, and S Konstantopoulos. 2002. Do low achieving students benefit more from small classes? Evidence from the Tennessee class size experiment. *Educational Evaluation and Policy Analysis*, **24**, 201-217.
Nye, B, Konstantopoulos, S and LV Hedges. 2004. How large are teacher effects? *Educational Evaluation and Policy Analysis*, **26**, 237-257.
Ogawa, RT, Huston, D, and DE Stine. 1999. California's class size reduction initiative: Differences in teacher experience and qualifications across schools. *Educational Policy*, **13**, 659-673.
Prais, SJ. 1996. Class size and learning: The Tennessee experiment – what follows? *Oxford Review of Education*, **22**, 399-414.
Quint, J, Bloom, HS, Rebeck Black, A and L Stephens with TM Akey. 2005. *The Challenge of Scaling Up Educational Reform: Findings and Lessons from First Things First.* New York: Manpower Development Research Corporation.
Reynolds, AJ, Temple, JA, Robertson, DL and EA Mann. 2002. Age 21 cost-benefit analysis of the Title I Chicago Child-Parent Centers. *Educational Evaluation and Policy Analysis*, **24**, 267-303.
Rivkin, SG, Hanushek, EA and JF Kain. 2005. Teachers, schools, and student achievement. *Econometrica*, **73**, 417-458.
Rose, H. 2006. Do gains in test scores explain labor market outcomes? *Economics of Education Review*, **25**, 430-446.
Rouse, CE. 1998. Private school vouchers and student achievement: An evaluation of the Milwaukee Parental Choice Program. *Quarterly Journal of Economics*, **113**, 553-602.

Stern, D, Dayton, C, Paik, I-W, and A Weisberg. 1989. Benefits and costs of dropout prevention in a high school program combining academic and vocational education: Third year results from replications of the California Penisula Academies. *Educational Evaluation and Policy Analysis*, **11**, 405-416.

Taylor, BA, Dearing, E and K McCartney. 2004. Incomes and outcomes in early childhood. *Journal of Human Resources*, **39**, 980-1007.

Taylor, LL. 1999. Government's role in primary and secondary education. *Federal Reserve Bank of Dallas Economic Review*, http://dallasfed.org/research/er/1999/er9901b.pdf.

Temin, P. 2003. Low pay, low quality. *Education Next*, **3**, 8-13.

Temple, JA, Reynolds, AJ, and S Ou. 2003. Grade retention and school dropout: Another look at the evidence. In Walberg, HJ, Reynolds, AJ, and MC Wang (Eds). *Can Unlike Students Learn Together?: Grade Retention, Tracking, and Grouping.* Greenwich, CT: Information Age.

Teske, P and M Schneider 2001. What research can tell policymakers about school choice. *Journal of Policy Analysis and Management*, **20**, 609-632.

Vawda, AY, Moock P, Gittinger JP, and HA Patrinos. 2003. Economic analysis of World Bank education projects and project outcomes. *International Journal of Educational Development*, **23**, 645-660.

Wayne, AJ and P Youngs. 2003. Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, **73**, 89-122.

Witte, JF. 1999. *The Market Approach to Education*. Princeton University Press: Princeton.

Wolfe, BL and N Tefft. 2004. Child interventions that may lead to increased economic growth. Working paper, Pew Charitable Trusts.

Wolfe, BL and S Zuvekas. 1997. Non-market outcomes of schooling. *International Journal of Education Research*, **27**, 491-502.