Evaluation and Program Planning 32 (2009) 74–82

Contents lists available at ScienceDirect

Evaluation and Program Planning

journal homepage: www.elsevier.com/locate/evalprogplan

Shifting the bell curve: The benefits and costs of raising student achievement

Stuart S. Yeh *

University of Minnesota, Educational Policy and Administration, College of Education and Human Development, 330 Wulling Hall, 86 Pleasant Street, S.E., Minneapolis, MN 55455, USA

ARTICLE INFO

Article history:
Received 26 December 2007
Received in revised form 15 July 2008
Accepted 22 August 2008

Keywords:
Cost–benefit analysis
Educational finance
Mathematics
Reading
Economics of education
Evaluation
Educational productivity

ABSTRACT

Benefit–cost analysis was conducted to estimate the increase in earnings, increased tax revenues, value of less crime, and reductions in welfare costs attributable to nationwide implementation of rapid assessment, a promising intervention for raising student achievement in math and reading. Results suggest that social benefits would exceed total social costs by a ratio of 28. Fiscal benefits to the federal government would exceed costs to the federal treasury by a ratio of 93. Social benefits would exceed costs to each state treasury by a ratio no lower than 286, and fiscal benefits would exceed costs to each state treasury by a ratio no lower than 5, for all but two state treasuries. Sensitivity analyses suggest that the findings are robust to a 5-fold change in the underlying parameters.

1. Introduction

In this paper, I estimate the benefits and costs of nationwide implementation of a promising intervention for raising student achievement. I use standard benefit–cost techniques to estimate the increase in earnings, increased tax revenues, value of less crime, and reductions in welfare costs that would be attributable to the intervention. The intervention that is proposed has largely been overlooked, and the results reported here are significantly more positive than the results reported by previous investigators regarding alternative interventions (Brady et al., 2005; Levin, Belfield, Muennig, & Rouse, 2007). The sensitivity analyses demonstrate that the findings are robust to a 5-fold change in the underlying parameters, including the arbitrary assumption that the effect of the intervention weakens after grade 5 and primarily maintains the gains achieved in grades 1–5.

1.1. Effects on educational attainment and earnings

While it is well-established that basic skills in math and reading predict educational attainment and earnings (Currie & Thomas, 2001; Murnane, Willet, & Levy, 1995; Neal & Johnson, 1996; O'Neill, 1990; Winship & Korenman, 1999), only one analysis has accounted for the reciprocal effects of cognitive skills and schooling (Winship & Korenman, 1999). Using data from the National Longitudinal Survey of Youth (NLSY), involving a national sample of 12,686 individuals, Winship and Korenman (1999) found that a 1.5 S.D. difference in basic math and reading skills is ultimately associated with a 1.945 year difference in educational attainment and a 53.7% difference in annual earnings, controlling for a host of covariates including ability.

Thus, a growing body of research indicates that strong preparation in math and reading, captured by math and reading test scores, can overcome socioeconomic disadvantages and "... is *the* major determinant of differences in educational attainment" between advantaged and disadvantaged young people (Bowen, Kurzweil, & Tobin, 2005, p. 224). This conclusion is not altered even after considering the affordability of college. There is no evidence that students are being forced to enroll in inexpensive colleges that are inappropriate for their level of preparedness (Hoxby, 2000). Instead, students from high and medium-high income families who have low SAT scores and high school grades are being replaced by highly prepared students from low income families (Hoxby, 2000). Less than 8% of all students are prevented from enrolling by their inability to pay (Avery & Hoxby, 2004; Carneiro & Heckman, 2003) and federal Pell grants have no significant impact on enrollment (Kane, 1999), leading Bowen et al. (2005) to conclude that "family finances have a fairly minor direct impact on a student's ability to attend a college" (p. 91). Math and verbal SAT scores "are much more important factors in the college [application] process than financial variables such as family income" (Spies, 2001, p. 17).

* Tel.: +1 612 624 0285.
E-mail address: yehxx008@umn.edu.

0149-7189/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.evalprogplan.2008.08.006

Furthermore, once students from disadvantaged socioeconomic backgrounds enroll, they do not underperform their more advantaged counterparts (controlling for SAT scores), suggesting that expansion of the college population to include a greater proportion of disadvantaged students would not depress the earnings of college graduates, as long as this expansion is the result of improved academic preparation—as envisioned by the proposed intervention: "When we [regress rank-in-class on SAT scores], we find that (perhaps as expected) SAT scores explain much of the variation in rank-in-class" (Bowen et al., 2005, p. 118). The significance of academic preparation, as measured by SAT scores, extends to graduation rates: When SAT scores are controlled, there is only a 4.7 percentage point difference in graduation rates between students in the top and bottom income quartiles (Bowen et al., 2005). Furthermore, after controlling for SAT scores, there is no difference between advantaged and disadvantaged students in their rates of attainment of lucrative law and business degrees (Bowen et al., 2005). In summary, these findings suggest not only that student achievement in reading and math is highly predictive of future educational attainment and earnings, but interventions that target reading and math achievement are promising ways to improve educational attainment and earnings.

2. Rapid assessment[1]

The difficulty is identifying an effective way to raise student achievement. However, a promising approach involves the implementation of systems where student performance in math and reading is rapidly assessed 2–5 times per week. The concept of rapid assessment is embodied by *Reading Assessment*,[2] a popular program designed to encourage students to read books at appropriate levels of difficulty while alerting teachers to learning difficulties and encouraging teachers to provide individualized tutoring or small group instruction. This is achieved through a system of frequently assessing each student's reading comprehension and monitoring each student's reading level. First, books in the school's library are labeled and shelved according to reading level. Second, students select books to read based on their interests and their reading levels, according to the results of the *STAR Reading* test, a norm-referenced computer-adaptive test (Renaissance Learning, no date). This helps students to avoid the frustrating experience of choosing a book that is too difficult. After finishing a book, the student takes a computer-based quiz, unique to the book, that is intended to monitor basic reading comprehension (*Rapid Assessment Corporation* has created more than 100,000 quizzes). Similarly, *Math Assessment* is a popular program that provides individualized, printed sets of math problems, a system of assessing student performance on those problems, and a scoring system where students and teachers receive rapid, frequent feedback on student performance upon completion of every set of problems.

A detailed review of research regarding the effectiveness of rapid formative assessment is available elsewhere (Yeh, 2007). To summarize, two large randomized experiments, involving a total of 2643 students, evaluated the effectiveness of the *Reading Assessment* program over 9-month periods (Nunnery, Ross, & McDonald, 2006; Ross et al., 2004). The average effect size was 0.279 S.D., with unusually disadvantaged populations of students. The only randomized study of *Math Assessment*, involving 1880 students, found an effect size of 0.324 S.D. over a 7-month period after controlling for treatment integrity (Ysseldyke & Bolt, 2007). The only national, peer-reviewed quasi-experimental evaluation of *Math Assessment*, involving 2202 students, found an average effect size of 0.392 S.D. over one semester (18 weeks) (Ysseldyke & Tardrew, 2007). The mean effect size, across the two studies of *Math Assessment*, is 0.358 S.D. Averaging the effect size estimates for *Reading* and *Math Assessment* produces an overall effect size of 0.319 S.D., suggesting that a 1.5 S.D. improvement in student achievement might be obtained over a period of 5 years.

3. Costs of rapid assessment

A detailed cost analysis of *Reading Assessment* and *Math Assessment*, two widely implemented variants of rapid assessment whose characteristics match the characteristics of effective feedback systems, is available elsewhere (Yeh, 2007). Large fixed costs of $22,809.60 incurred by every building of 500 students at start-up, including costs of software, teacher and administrator training, and scanners, were amortized over the life of the program, assumed to be 7 years. Ongoing costs include fees that provide access to 100,000 book quizzes for every student, plus access to *Math Assessment* grade level libraries tagged to state standards for grades 1 through 7, as well as multiple subject area libraries for the secondary grades (pre-algebra, algebra 1, algebra 2, geometry, probability and statistics, pre-calculus, calculus, basic math, chemistry, and physics). The annual cost in 2006 dollars, including both fixed and ongoing costs, averaged $22.27 per student, or $28.31 per student adjusted for the opportunity costs of teacher training time and opportunity costs created by large upfront fixed costs.[3]

The effect size for rapid assessment (0.319 S.D. per year) suggests that improving student achievement by 1.5 S.D. for every American student could be achieved by implementing rapid assessment in grades 1 through 5. However, it is likely that the effects of a 5-year intervention would fade over time without booster shots every year. Furthermore, student mobility, the influx of English language learners, and the need to close achievement gaps by the No Child Left Behind deadline of 2014 create a need for ongoing intervention to bring all students up to grade level. Implementing the intervention in grades 1 through 12 would serve to maintain gains by students in grades 1 through 5, accommodate the needs of immigrants and English language learners, and boost the achievement of students who have already advanced beyond grade 5. Thus, it would be desirable to implement the intervention throughout grades 1–8 for all students, and through high school for the bottom 40% of all students. Therefore, the expected annual cost per student is $28.31, incurred from ages 6 through 13, and $11.32 incurred from ages 14 through 17, or a total of $226.19 after discounting at a rate of 3.5% to age 6, when the intervention begins.

3.1. Extra college costs

For the majority of students who receive the proposed intervention, a 1.945 increase in years of schooling associated with a 1.5 S.D. increase in AFQT test scores creates costs to society of 1.945 years of college plus foregone wages. Weighted by the proportion of students enrolled in degree-granting public, private

[1] This section is adapted from a paper (Yeh, 2007) comparing the effects of rapid assessment with the effects of increased educational expenditure, vouchers, charter schools, and increased educational accountability.

[2] *Reading Assessment*, *Math Assessment*, and *Rapid Assessment Corporation* are pseudonyms used to avoid the appearance that the author endorses the assessment software.

[3] Cost figures reflect the actual operating experience of schools in a typical district, verified by the researcher through classroom observation of operating procedures as well as teacher and administrator interviews in 8 schools (spanning elementary, middle, and high school levels). While a school of 500 students taking 2–5 assessments per week in math and reading suggests that 2000–5000 assessments are processed weekly, the burden on teachers is minimal because students scan their own bubble sheets, the software scores each assessment, and summary reports are available to teachers and administrators electronically.