University of Minnesota-Twin Cities

From the SelectedWorks of Stuart S Yeh

2010

The Cost-Effectiveness of NBPTS Teacher Certification

Stuart S Yeh, University of Minnesota-Twin Cities



Available at: https://works.bepress.com/stuart_yeh/6/

Evaluation Review 34(3) 220-241 © The Author(s) 2010 Reprints and permission: sagepub.com/journalsPermissions.nav DOI: 10.1177/0193841X10369752 http://er.sagepub.com



The Cost-Effectiveness of NBPTS Teacher Certification

Stuart S. Yeh¹

Abstract

A cost-effectiveness analysis of the National Board for Professional Teaching Standards (NBPTS) program suggests that Board certification is less cost-effective than a range of alternative approaches for raising student achievement, including comprehensive school reform, class size reduction, a 10% increase in per pupil expenditure, the use of value-added statistical methods to identify effective teachers, and the implementation of systems where student performance in math and reading is rapidly assessed 2–5 times per week. The most cost-effective approach, rapid assessment, is three magnitudes as cost-effective as Board certification.

Keywords

teacher certification, assessment, achievement, cost-effectiveness, economics of education

¹ University of Minnesota, Minneapolis, MN, USA

Corresponding Author:

Stuart S. Yeh, University of Minnesota, 330 Wulling Hall, 86 Pleasant Street, S.E., Minneapolis, MN 55455, USA Email: yehxx008@umn.edu Prominent educational researchers have concluded that teacher quality is perhaps the most important factor predicting student achievement (Ferguson 1998; Goldhaber, Brewer, and Anderson 1999; Goldhaber 2002; Hanushek, Kain, and Rivkin 1999; Wright, Horn, and Sanders 1997). To illustrate, Hanushek (1992) found that a high-quality teacher can increase learning by an entire grade level equivalent above the amount contributed by a low-quality teacher. This conclusion is currently unchallenged by other educational researchers. Thus, improvements in teacher quality are believed to be a promising approach for raising student achievement. However, unless there is a way for building principals to identify high-quality teachers, it is difficult to select those teachers to improve student achievement.

Research regarding identifying factors is hampered by the difficulty of disentangling those factors from the contribution of other factors influencing student achievement. This difficulty is best addressed through random assignment of teachers and students to classrooms. Thus, it is significant that the only study of teacher quality using random assignment concluded that the main observable characteristics of teachers—years of experience and level of education—are essentially uncorrelated with gains in student achievement: "Neither teacher experience, nor teacher education explained much variance in teacher effects (never more than 5%)" (Nye, Konstantopoulos, and Hedges 2004, 249). A previous review of the research concluded:

While few would disagree with the claim that better teachers would improve the quality of education and student achievement, the task of identifying and hiring effective teachers is not simple. For example, one study found that only 3% of the contribution teachers made to student learning was associated with teacher experience, degrees attained, and other readily observable characteristics (Goldhaber 2002). In other words, years of experience and graduate degrees do not have a strong impact on student achievement. Students whose teachers have master's degrees do not outperform students whose teachers have only a bachelor's degree (Hanushek et al. 2005; Wenglinsky 2001; Grissmer et al. 2000). Meta-analyses and reviews of research suggest that there is no clear relationship (Greenwald, Hedges, and Laine 1996; Hanushek 1989; Wilson and Floden 2003) or close to zero relationship (Hanushek 1996; Hedges, Laine, and Greenwald 1994; Hanushek et al. 2005) between teachers' level of educational attainment and student achievement. The relationship between teacher experience and student achievement is also inconsistent (Wilson and Floden 2003). First- and second-year teachers are less effective, on average, than more experienced teachers, but after this period, experience has little effect (Hanushek et al. 2005; Wenglinsky 2001; Grissmer et al. 2000). Therefore, if readily observable characteristics such as graduate credentials and years of experience are weakly related to teacher quality, it may be extremely difficult for a school principal to identify and hire good teachers. (Yeh 2006, 147)

If readily observable characteristics are inadequate for the purpose of identifying strong teachers, perhaps a sophisticated certification exam, in combination with portfolio evidence including video recordings of teacher–student interactions, can discriminate effectively. This approach may be more sensitive to the attributes that actually make teachers successful in the classroom.

This article focuses on a cost-effectiveness analysis of the National Board for Professional Teaching Standards (NBPTS) teacher certification program. NBPTS is an independent organization established in 1987 with the goal of advancing the quality of teaching and learning (National Board for Professional Teaching Standards 2006a). NBPTS developed professional standards for teaching, then contracted with the Educational Testing Service (and, later, Pearson Educational Measurement) to create a voluntary system to certify teachers who meet those standards (Educational Testing Service 2004; National Board for Professional Teaching Standards 2006a). Perhaps the clearest indication that a cost-effectiveness analysis of the NBPTS program is needed is that in January, 2004, the U.S. Congress directed the National Research Council (NRC), which is the research arm of the National Academy of Sciences, to conduct an evaluation, including a cost-effectiveness analysis. Subsequently, the NRC wrote an evaluation report (Hakel, Koenig, and Elliott 2008) but found that insufficient information was available to conduct the cost-effectiveness analysis. This article seeks to fulfill the request by the U.S. Congress for a cost-effectiveness analysis of the NBPTS program using cost information that was not available to the NRC.

The chair of the NRC committee that wrote the report concluded that "Earning NBPTS certification is a useful 'signal' that a teacher is effective in the classroom," where classroom effectiveness is defined in terms of the level of student achievement that is associated with specific teachers (The National Academies 2008). While NBPTS seeks to achieve outcomes in addition to improvements in student achievement, as measured through standardized tests, the report noted that little research is available regarding the effects of NBPTS on student motivation, breadth of achievement,

attendance and rates of promotion, or possible spillover effects on the teaching practices of NBPTS teacher colleagues, school and district educational standards, and the quality of school- and district-level teacher professional development (Hakel, Koenig, and Elliott 2008). Findings regarding the effects of NBPTS as a form of professional development are mixed (Hakel, Koenig, and Elliott 2008). Finally, research regarding the effects of NBPTS certification on teacher mobility and career paths is not rigorous enough to permit causal inferences (Hakel, Koenig, and Elliott 2008). Thus, there is insufficient evidence to evaluate the effect of NBPTS on the field of teaching and the education system (Hakel, Koenig, and Elliott 2008). While the available evidence suggests that NBPTS certification may provide a means of identifying highly skilled teachers, this evidence does not provide sufficient information about the effect of the NBPTS certification process as a form of professional development or the effect of the program on teacher recruitment and retention (Hakel, Koenig, and Elliott 2008).

At present, there is insufficient evidence to suggest that NBPTS certification has broad effects beyond serving as a signal of teacher effectiveness according to standardized measures of student achievement. Thus, it is appropriate to focus on the cost-effectiveness of the NBPTS certification system with regard to the impact on student achievement if uncertified teachers are replaced with certified teachers. I applied standard cost-effectiveness techniques to evaluate the relative cost-effectiveness of a range of alternative approaches for raising student achievement, including two alternative methods of identifying effective teachers (Yeh 2009b; Yeh and Ritter 2009).

Cost-effectiveness analysis provides standard techniques and a conceptual framework to assist policy makers in comparing dissimilar interventions for raising student achievement (Levin and McEwan 2001; Levin and McEwan 2002). For example, Levin, Glass, and Meister (1987) evaluated the cost-effectiveness of four dissimilar interventions (computer-assisted instruction, cross-age tutoring, class size reduction, and increases in instructional time) to assess the relative impact of each intervention per dollar, per pupil. Importantly, there is no requirement that the interventions share common features—only that they share the ultimate goal of raising student achievement. The ultimate rationale for NBPTS certification is that teachers who are certified produce higher levels of student achievement; therefore, studies that judge NBPTS certification based on that metric are appropriate. The theory of action underlying NBPTS certification is that it is possible to improve achievement by replacing weak teachers with strong teachers. The quality of teaching is an intermediate, rather than a final goal. Society cares about the quality of teaching to the extent that it improves student outcomes. While a critic might argue that standardized measures of student achievement are imperfect, and only capture one of many outcomes that result from NBPTS certification programs, standardized measures of achievement are the only widely used measures of student achievement that permit comparisons across a wide range of approaches for raising student achievement. In practice, comparative cost-effectiveness analyses can only be conducted using this type of measure. Thus, to reject standardized measures of student achievement is to reject the application of cost-effectiveness analysis to the field of education.

While cost-effectiveness analysis is not new, certain features may be unfamiliar. Importantly, the evaluation of effectiveness is separable from the evaluation of costs (Levin 1988). Thus, estimates of effectiveness can be derived from published evaluations and then combined with estimates of costs derived through the ingredients approach (Levin and McEwan 2001). The information regarding effectiveness and cost is standardized in ratios that facilitate comparisons across interventions. Widely dissimilar interventions are likely to result in widely disparate effectiveness-cost ratios. This information is valuable to policy makers who are grappling with decisions about the most cost-effective way to efficiently allocate scarce taxpayer dollars to improve student achievement.

It is important to complete a rigorous cost-effectiveness analysis because the results may contradict popular beliefs-among researchers as well as policy makers. As Levin, Glass, and Meister (1987) demonstrated, interventions that are resource-intensive, such as tutoring, may in fact be more costeffective than popular alternatives such as class size reduction. Thus, it is essential to conduct and publish rigorous cost-effectiveness studies even when "back of the envelope" calculations suggest that a particular intervention may not be cost-effective. While a quick glance at recent evaluations of the NBPTS certification program may suggest that it is unlikely to be cost-effective, it is vital to conduct a thorough cost-effectiveness study because many educational researchers and policy makers may assume, in the absence of a rigorous study, that teacher certification is a promising, effective, and cost-effective approach for raising student achievement. In the absence of a study that addresses this question directly, scarce public resources may be misallocated, the achievement of disadvantaged students may remain depressed, and schools that are not making "adequate yearly progress" may continue to be sanctioned under the No Child Left Behind Act of 2001.

Costs of NBPTS Certification

NBPTS certification is a lengthy, highly demanding process. Applicants for certification are required to submit a portfolio to NBPTS involving four entries (National Board for Professional Teaching Standards 2006b). Three are classroom based, where video recordings of teacher-student interaction and examples of student work serve as supporting documentation. A fourth entry relates to the candidate's accomplishments outside of the classroom with families, the community or colleagues-and how they affect student learning. Each entry requires some direct evidence of teaching or school counseling as well as a commentary describing, analyzing, and reflecting on this evidence. Following submission of the portfolio, candidates are tested on their content knowledge through six 30-minute exercises, specific to the candidate's chosen certificate area, at 1 of 300 NBPTS computerbased testing centers across the United States. Applicants are scored on a scale of 75 to 425, incorporating both the portfolio and the assessment center exercises, and they must earn a score of at least 275 to achieve certification (Goldhaber, Perry, and Anthony 2003).

Rice and Hall (2008) provide the best available estimate of the full social cost of the NBPTS certification program, with detailed estimates of costs for all resources devoted to program administration and infrastructure, information and recruitment, group meetings, portfolio development, the NBC application fee, mentor training, research, development, and dissemination, averaging \$25,665.37 (adjusted to 2006 dollars) per participant, per year. Importantly, Rice and Hall (2008) estimated the large opportunity costs of the time teachers devote to the development of their portfolios, as well as the uncompensated time of staff including teacher mentors, librarians, child care providers, and administrators who provide support. These opportunity costs constitute the bulk of the social cost of the NBPTS certification program.

However, while Rice and Hall (2008) assumed that the certification process averages 1 year, it actually takes 2 years (Sanders, Ashton, and Wright 2005; Goldhaber and Anthony 2006). Because the estimate by Rice and Hall (2008) of the program's cost per graduate is based on calculations where total annual costs for all candidates are divided by the annual number of participants, their headcount, using the total number of (first- and second-year) participants in any given year, shrinks the cost per graduate by half (see p. 348, footnote 12). Correcting for this underestimate suggests a cost per graduate of \$51,330.74. Amortized over an average teaching career of 7.86 years as an NBPTS-certified teacher, and averaged over 20 students

per classroom, the annual cost of NBPTS certification per student is \$326.53.¹ The cost is underestimated to the extent that stricter certification requirements deter high-quality applicants from teaching in the public schools, raising barriers to entry that increase labor costs (Angrist and Guryan 2004).

A limitation of the NBPTS certification program is that a substantial percentage of all teachers have taught for less than 5 years and, thus, would have not reached a point where they could have accumulated the minimum 3 years of teaching experience required for NBPTS certification (National Board for Professional Teaching Standards 2006c) nor finished a certification process that takes an average of 2 years beyond the 3-year probationary period (Sanders, Ashton, and Wright 2005). Thus, any conceivable benefits of NBPTS certification can never be realized for the large proportion of teachers in the workforce with 5 or less years of teaching experience.

Nationally. 20% of fourth grade teachers, 23% of eighth grade math teachers, and 22% of eighth grade reading teachers have less than 5 years of teaching experience (Stancavage et al. 2006). Averaging the eighth grade figures and weighting the fourth and eighth grade percentages by total fourth grade (3,611,638) and eighth grade (3,824,670) enrollment (U.S. Department of Education 2007) suggests that, overall, about 21.3% of all teachers have less than 5 years of teaching experience. Thus, (hypothetical) federal requirements that all eligible teachers apply for NBPTS certification, and all teachers who fail the NBPTS exam be replaced with NBPTS-certified teachers, would only benefit the 78.7% of students who are taught by teachers with 5 or more years of teaching experience. Furthermore, if all eligible teachers apply for NBPTS certification and teachers who fail the exam are replaced with teachers who are certified, the process of teacher replacement would take a minimum of 8 years, including 6 years to replace teachers (who fail the NBPTS exam) through normal teacher attrition, in addition to the 2 years required for teachers to complete the application and examination process.²

Effects of NBPTS Certification

Several studies have found that teachers who eventually receive NBPTS certification were more effective than other teachers before they began the certification process, but something about the certification process reduces their productivity and their performance never returns to pre-application levels (Clotfelter, Ladd, and Vigdor 2007; Goldhaber and Anthony 2007; Harris and Sass 2007). NBPTS certification may identify teachers who are

initially more effective than uncertified teachers, but that information may be gained at the cost of reduced teacher performance. Thus, it is important to distinguish the signaling effect, measured by the deviation of the performance of certified teachers compared to the performance of never certified teachers at a point *after* certification, in contrast to the effect of the certification process on the human capital of certified teachers, measured from pre- to postcertification.

Seven large-scale studies offer the necessary power to detect effects, if they exist, and either controlled for student or school fixed effects or used hierarchical linear modeling (HLM; Goldhaber and Anthony 2007; Cavalluzzo 2006; Sanders, Ashton, and Wright 2005; Clotfelter, Ladd, and Vigdor 2007, forthcoming; Harris and Sass 2007; Ladd, Sass, and Harris 2007). These studies provide the best available estimates of the signaling and human capital effects of NBPTS certification.

The largest study involved NBPTS teachers in Florida, 1,517 certified in reading and 1,256 in math, and controlled for student and school fixed effects (Harris and Sass 2007). While teachers certified before 2001 were more productive after certification than teachers who were never certified, the most recent data, for teachers certified in 2002 and 2003, indicate that the certification process actually reduced teacher performance to the level of teachers who were never certified, with effect sizes (from precertification to postcertification) averaging -0.069 SD in math and -0.115 SD in reading. The researchers concluded that NBPTS-certified teachers in Florida were more effective than other teachers before they start the certification process, but their relative productivity fell during the certification process and never returned to pre-application levels. After certification, the performance of teachers certified in 2002 and 2003 was nearly identical to the performance of never certified teachers, with average signaling effect sizes of 0.000 SD in math and -0.002 SD in reading.

A second study involved an unspecified number of NBPTS-certified teachers but included all teachers in grades 3, 4, and 5 in North Carolina for the years 1995 to 2004 (Clotfelter, Ladd, and Vigdor 2007). After controlling for student fixed effects, the certification process reduced teacher performance, with effect sizes (from 2 years prior to certification to postcertification) averaging -0.008 SD in math and -0.015 SD in reading, based on the researchers' preferred models. Although the reductions in performance were not statistically significant at conventional levels, they fit the pattern reported by Harris and Sass (2007) as well as Goldhaber and Anthony (2006). After certification, NBPTS teachers remained somewhat

more effective than never certified teachers, with average signaling effect sizes of 0.032 SD in math and 0.017 SD in reading.

A third study analyzed data from Florida and North Carolina using common specifications and time periods and controlling for both student and school fixed effects (Ladd, Sass, and Harris 2007). The certification process had mixed effects on teacher performance (from 2 years prior to certification to postcertification), with average effect sizes of 0.016 *SD* in math and a negative -0.022 *SD* in reading. After certification, NBPTS teachers remained somewhat more effective than never certified teachers, with average signaling effect sizes of 0.050 *SD* in math and 0.023 *SD* in reading. McCaffrey and Rivkin (2007) drew upon the same data and analyses and reported identical results (see the coefficients for their gain model with student and school fixed effects in table 4).

A fourth study, involving 42 NBPTS-certified teachers, controlled for school fixed effects and identified a sample of students who were apparently randomly assigned to classrooms within their schools (Clotfelter, Ladd, and Vigdor 2006). This identification was based on six chi-square tests to determine whether the distribution of students within each classroom, based on the students' observed characteristics, fit the random pattern that would be expected if students were indeed randomly assigned. After controlling for lagged student achievement, NBPTS-certified teachers were less effective in math than never certified teachers, with a negative signaling effect size of -0.035 SD but marginally more effective in reading, with a positive effect size of 0.005 SD.

A fifth study, sponsored by NBPTS, involved an unspecified number of teachers from two large North Carolina school districts, applied HLM, controlled for various teacher characteristics, and found signaling effect sizes of 0.018 *SD* in math and -0.038 *SD* in reading, compared to teachers who never applied for certification (Sanders, Ashton, and Wright 2005).

A sixth study involved 303 of North Carolina's NBPTS-certified teachers and found, prior to certification, that teachers who eventually became certified were significantly more effective than teachers who never applied for certification or applied and failed (Goldhaber and Anthony 2007). However, the relevant comparison is effectiveness after the certification process is completed. In 5 of 6 specifications, the signaling effect for NBPTS-certified teachers after their 1st year of certification was not statistically different (at an α level of .05) than the effect for teachers who never applied for certification, and it is significantly negative in the sixth specification. Thus, it appears that NBPTS-certified teachers were no more effective than never certified teachers after the 1st year of certification. The most rigorous

Study	Effect Size (SD)				
	Reading	Math	Average		
Harris and Sass (2007)	-0.002ª	0.000ª	-0.001		
Clotfelter, Ladd, and Vigdor (2007)	0.017 ^b	0.032 ^b	0.025		
Ladd, Sass, and Harris (2007)	0.023 ^c	0.050 ^c	0.037		
Clotfelter, Ladd, and Vigdor (2006)	0.005 ^d	–0.035 ^d	-0.015		
Sanders, Ashton, and Wright (2005)	-0.038 ^e	0.018 ^e	-0.010		
Goldhaber and Anthony (2007)	0.004 ^f	-0.098 ^f	-0.047		
Cavalluzzo (2006)	_	0.063 ^g	0.063		
Average	0.002	0.004	0.007		

Table 1. Signaling Effect Sizes for NBPTS-Certified Teachers in Comparison Wi	ith
Teachers Who Were Never Certified by NBPTS	

^a Coefficients from table 4 converted to effect sizes using the means of the standard deviations reported on page 18. ^b The average of the coefficients from Models 4 and 5 (see table 6). ^c Coefficients from table 5 (specifications control for both student and school fixed effects). ^d Coefficients from the model controlling for lagged achievement, see table 7. ^e Average of effect sizes ("Cert vs. Never") across Grades 4, 5, 6, 7, and 8 (see table 3d). ^f Coefficients for "Past NBCT," from table 2, Models 6 and 12 (which control for student fixed effects), converted to effect sizes using the reported standard deviations of 9.94 in reading and 12.34 in math. ^g Coefficients from Model 15, which controls for student fixed effects.

specification controlled for student fixed effects and found a small positive signaling effect size of 0.004 SD in reading and a negative effect size of -0.098 SD in math.

The seventh study, regarding math achievement in the Miami-Dade County Public Schools, compared the performance of 61 NBPTS-certified teachers to teachers who never applied for certification and found a signaling effect size of 0.063 *SD*, controlling for student fixed effects and a range of teacher and school characteristics (Cunningham and Stone 2005; Cavalluzzo 2006).

To summarize, there is a small signaling effect of NBPTS certification, and effects on human capital are either mixed or negative. The average signaling effect size across the seven key studies is 0.002 *SD* in reading and 0.004 *SD* in math (Table 1). This represents the average gain in student achievement of replacing an existing teacher with an NBPTS-certified teacher (the effect is diluted because some teachers in the general population are already teaching at the NBPTS level and would pass the NBPTS exam if they applied while others would fall below the NBPTS standard and would fail the NBPTS exam).

While other studies have investigated the relationship between NBPTS certification and student achievement (Bond et al. 2000; Vandevoort, Amrein-Beardsley, and Berliner 2004; McColskey et al. 2006; Stone 2006), none involved a sample with more than 35 NBPTS-certified teachers, none controlled for student fixed effects, and the only study that used HLM involved a small sample of 25 NBPTS-certified teachers and failed to find any impact on student achievement (McColskey et al. 2006). These studies are limited by key methodological weaknesses (Education Commission of the States 2006; Cunningham and Stone 2005).³

Relative Cost-Effectiveness

If the average effect size across all students for NBPTS certification is 0.002 *SD* in reading and 0.004 *SD* in math, and the average cost per student is \$326.53, the effectiveness-cost ratio (effect size in standard deviation units divided by annual cost per student in dollars) for NBPTS certification is 0.000006 in reading and 0.000012 in math.

For comparison, Table 2 lists the annualized effect sizes, annual costs per student, and effectiveness-cost ratios for a range of alternative approaches for raising student achievement, including the proposal by Gordon, Kane, and Staiger (2006) to use value-added statistical methods to identify and replace ineffective probationary teachers with new teachers, a cheaper version of the proposal by Gordon, Kane, and Staiger (2006) using 5th year college graduates, establishing a minimum SAT score of 1,000 for new teacher applicants, voucher programs, charter schools, a 10% increase in per pupil expenditure, increased educational accountability (defined as the implementation of high school-level exit exams), comprehensive school reform, class size reduction, high-quality preschool, and rapid assessment, where student performance in math and reading is rapidly assessed 2–5 times per week.

The effectiveness-cost ratios for NBPTS teacher certification are at the low end of the ratios in Table 2, suggesting that it is not a cost-effective approach for raising student achievement. An alternative approach for improving the quality of the teaching force is the proposal by Gordon, Kane, and Staiger (2006) to identify and replace the bottom quartile of novice teachers at the end of their second year of teaching using value-added assessments of teacher performance. Their proposal is one of the featured proposals by the Hamilton Project at the Brookings Institution. The significance of the proposal by Gordon, Kane, and Staiger (2006) is that it converts the general strategy of value-added teacher assessment into a

	Effect Size (SD) ^a			Effectiveness-Cost Ratio ^c	
	Reading	Math	Cost ^b	Reading	Math
NBPTS teacher certification	0.002	0.004	\$326.53	0.000006	0.000012
Proposal by Gordon, Kane, and Staiger (2006)	0.057	0.057	\$624.72	0.000091	0.000091
Gordon et al. (using 5th year grads)	0.057	0.057	\$512.27	0.000111	0.000111
Minimum SAT score of 1,000	0.004	0.015	\$894.10	0.000004	0.000017
Voucher programs	0.032	0.080	\$9,646.0	0.000003	0.000008
Charter schools	0.009	0.001	\$8,086.30	0.000001	0.000000
10% increase in spending	0.083	0.083	\$1118.83	0.000075	0.000075
Increased accountability	0.051	_ 0.062	\$2025.97	0.000253	0.000253
Comprehensive school reform	0.510	0.510	\$217.83	0.002341	0.002341
Class size reduction					
Nye, Hedges and Konstantopoulos, (2001)	0.104	0.090	\$1,379.28	0.000075	0.000065
Finn et al. (2001)	0.120	0.129	\$1,379.28	0.000087	0.000094
Perry preschool	0.150	0.155	\$12,147.03	0.000012	0.000013
Abecedarian preschool	0.150	0.054	\$10,188.09	0.000015	0.000005
Rapid assessment	0.270		\$9.45	0.028571	
(high estimates)		0.392	\$18.89		0.020752
Rapid assessment	0.175		\$9.45	0.018519	
(low estimates)		0.324	\$18.89		0.017152

Table 2. Comparison of Effect Sizes, Costs, and Effectiveness-Cost Ratios for Various Interventions to Raise Student Achievement

^a Annualized effect size. ^b Annual cost per student, adjusted for inflation to August 2006 dollars.

^c Effect size in SD units divided by annual cost per student.

concrete proposal that can be evaluated. It is currently the only proposal that adequately addresses the fact that most teachers are tenured and, therefore, cannot be replaced without adequate cause. For this reason, it is perhaps the most promising strategy for implementing a value-added teacher assessment system. However, the student achievement effect size of 0.057 SD, calculated from the data by Gordon, Kane, and Staiger (2006), is small and the annual cost per student of \$624.72 is high, implying a low costeffectiveness ratio of 0.000091 (Yeh and Ritter 2009). A version of the proposal by Gordon, Kane, and Staiger (2006) using 5th year college graduates is somewhat cheaper (\$512.27 per student) but the cost-effectiveness ratio

remains low (0.000111). Because the proposal by Gordon, Kane, and Staiger (2006) is not cost-effective, relative to the most cost-effective approaches listed in Table 1, it is likely that the value-added teacher assessment strategy is not cost-effective.

Yet another approach for raising the quality of the teaching force is to impose a requirement that all new teacher applicants achieve a minimum score of 1,000 on the SAT exam. However, low student achievement effect sizes of 0.004 *SD* in reading and 0.015 *SD* in math, coupled with high costs of raising teacher salaries to recruit the necessary pool of teachers, imply low effectiveness-cost ratios of 0.000004 in reading and 0.000017 in math (Yeh 2009b).

Table 2 lists additional ratios drawn from published cost-effectiveness studies: voucher programs, charter schools, a 10% increase in educational expenditure, and increased educational accountability (Yeh 2007); an upper-bound estimate for comprehensive school reform (Yeh 2008); an upper-bound estimate for class size reduction (Yeh 2008, 2009a); the Perry and Abecedarian preschool programs (Yeh 2008) and rapid assessment, which is a program that individualizes instruction and frequently assesses each student's reading comprehension and math problem-solving ability (Yeh 2008).

Discussion

The ratios in Table 2 suggest that the most cost-effective approach for raising student achievement is rapid assessment. Rapid assessment is approximately one magnitude (10 times) as cost-effective as comprehensive school reform, two magnitudes as cost-effective as class size reduction or a 10% increase in per pupil expenditure or the proposal by Gordon, Kane, and Staiger (2006) or increased educational accountability, three magnitudes as cost-effective as NBPTS teacher certification or imposing a minimum SAT score of 1,000 on new teacher applicants or voucher programs or high-quality preschool, and four magnitudes as cost-effective as charter schools (a magnitude of gain implies that student achievement would increase 10 times faster for every dollar invested in rapid assessment rather than the alternative).

The magnitude of the differences in cost-effectiveness suggests that improving student achievement by 0.175 *SD* in reading for every American student would take 9 months and cost \$459.0 million using rapid assessment (\$9.45 multiplied by projected 2006 PK-12 enrollment of 48,574,000 students). To achieve the same effect size through NBPTS certification would take 88 years and cost \$1.4 trillion (\$326.53 multiplied by 48,574,000 students and 88 years).

A question that arises is whether it is appropriate to compare an intervention such as NBPTS certification that is designed to improve the quality of the teaching force with an intervention such as rapid assessment, which aims to provide teachers with information to improve student performance. As noted earlier, however, the ultimate goal of improving the quality of the teaching force is to improve student achievement (if achievement does not improve, then what does it mean to assert that teacher quality has improved?). As noted earlier, a systematic review of the research literature conducted by the National Academy of Sciences concluded that there is insufficient information about the effect of the NBPTS certification process on broad outcomes including teacher recruitment and retention, the quality of school- and district-level teacher professional development, possible spillover effects on the teaching practices of NBPTS teacher colleagues or school and district educational standards, or effects of the program as a form of professional development (see Hakel, Koenig, and Elliott 2008).

The only rigorous studies of the effects of NBPTS certification involve the usefulness of certification as a signal of student performance in math and reading. Because it is well established that basic skills in math and reading, as measured by standardized tests, predict educational attainment and earnings (Currie and Thomas 2001; Murnane, Willet, and Levy 1995; Winship and Korenman 1999; Neal and Johnson 1996; O'Neill 1990), it is appropriate to conduct a cost-effectiveness study that focuses on the signaling effect of NBPTS certification with regard to student achievement in math and reading.

If the results of standardized tests of math and reading achievement matter, what do the results of the cost-effectiveness analysis tell us? Perhaps the most striking implication is that additional expenditure on systems such as rapid assessment that provide teachers with information to improve student performance is more cost-effective than expenditure on efforts to improve the quality of the teaching force through NBPTS certification, value-added assessment of teacher performance, or imposing a requirement that new teachers must achieve a minimum score of 1,000 on the SAT exam. These analyses suggest a need to reexamine the basic premise that higher standards for teacher selection are the key to improved student achievement. Instead, it may be more productive to implement classroom assessment systems designed to help teachers rapidly diagnose and address student weaknesses.

While diagnosing and addressing student weaknesses is important, a second—and perhaps more important—reason that rapid assessment may be effective is that it may directly improve student engagement. Research suggests that an effective way of engaging students and building intrinsic interest in academic work is to provide performance feedback through rapid assessment of math and reading performance (see Yeh, forthcoming, for a review).

A final concern is that a focus on diagnostic assessment systems might appear to imply a conception of teaching that is driven by narrow measures of math and reading achievement. Clearly, teachers must exercise professional judgment in using the results of diagnostic assessments in math and reading, just as physicians must exercise professional judgment in using the results of diagnostic tests that provide narrow measurements of blood pressure and heart rate. The full benefits of diagnostic assessments are likely to depend on mentoring from teachers who have successfully used those assessments, just as novice physicians learn to use diagnostic results under the guidance of more experienced physicians. However, when the results of diagnostic assessments are used properly, the results of the cost-effectiveness analysis suggest that these assessments are a more cost-effective use of society's resources than NBPTS certification for the purpose of improving student achievement in math and reading.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

Funding

The author(s) received no financial support for the research and/or authorship of this article.

Notes

1. The best available estimate of the career duration of the average teacher was derived using proportional hazards modeling, which accounts for the difficulty of estimating career duration when some members of the research sample have not exited the teaching profession by the end of the research study period (Murnane, Singer, and Willett 1988). Proportional hazards modeling incorporates information about the pattern of teacher attrition during the study period to predict the median length of each spell of teaching. Using data from Michigan covering a 12-year time period, Murnane, Singer, and Willett (1988) provided separate estimates, for six subject area specialties, of the duration of the average

teacher's first two spells of teaching. The authors reported the percentage distribution of teachers across the six subject area specialties, as well as the percentage of teachers in each of the six subject areas who returned to teaching after a career interruption. I used this information to calculate the average career duration for an average teacher, weighted by the percentage distribution of teachers across the six subject area specialties and including the expected length of a second spell of teaching based on the probability of a second spell. This average (9.11 years) may be conceptualized as a weighted combination of two averages: (a) the average career duration for teachers who exit the teaching profession before they can be certified (3 years, based on a survival function that falls to zero by the end of Year 5, because National Board certification is only available to teachers who have 3 or more years of teaching experience [Cavalluzzo 2006], and because the certification process takes approximately 2 years after teachers become eligible and apply for certification [Sanders, Ashton, and Wright 2005; Goldhaber and Anthony 2006], it probably takes a minimum of 5 years from the date a teacher begins teaching to the date of certification) and (b) the average career duration for teachers who stay long enough to be certified (that is, all teachers who stay longer than 5 years). Based on the survival functions and proportional weights provided by Murnane, Singer, and Willett (1988), the average career duration of teachers who stay for at least 5 years is 12.86 years, suggesting that the average NBPTS-certified teacher could be expected to teach for 7.86 years (12.86 minus 5) after certification by NBPTS. Thus, the costs of the certification process may be amortized over 7.86 years, the average period over which the benefits would be realized. Although Murnane, Singer, and Willett (1988) offer the most sophisticated estimate of the average career duration for all teachers, an alternative estimate may be derived from NRC data regarding the career duration of elementary and secondary teachers in Michigan and North Carolina (Boe and Gilford 1992). Based on the NRC data, the average career duration for elementary teachers is 8.45 years, while the average duration for secondary teachers is 5.85 years. Weighting by the total number of elementary (1,363,937.3) and secondary (1,033,065.5) teachers (U.S. Department of Education 2007) suggests an average teaching career of 7.33 years. This estimate is shorter than the estimate of 9.11 years used in the current analysis. An analysis using the shorter figure would effectively increase the annual cost of NBPTS certification by reducing the period over which the costs of certification would be amortized.

2. The benefits of certified teachers would be gained very slowly. Students can only gain the benefits of certified teachers when building principals replace teachers who are performing below NBPTS standards with certified teachers (there is no benefit to students if existing teachers become certified). However, lack of NBPTS certification would not be grounds for firing existing teachers, so

dismissal is not an option. Furthermore, releasing probationary teachers who are uncertified (but would otherwise be rehired) would not only create a glut of uncertified teachers but would also choke off the future supply of certified teachers, because NBPTS certification is not available to teachers until they have accumulated a minimum of 3 years of teaching experience. Thus, the primary process through which a building principal could replace teachers who are performing below NBPTS standards with NBPTS-certified teachers is through normal turnover of existing staff. While annual teacher turnover is 13.2%, approximately half of the replacements are new teachers (Ingersoll 2001) who cannot be NBPTS certified, because certification only occurs after 3 years of teaching plus an application process that takes an average of 2 years (see paragraph above), suggesting that building principals could replace a maximum of 6.6% of their teaching staffs every year with NBPTS-certified teachers (assuming an adequate supply of certified teachers). If "x" is the size of the teaching staff, then .066x teachers may be replaced in the 1st year, and .066 (x - x).066x) teachers may be replaced in the 2nd year with NBPTS-certified teachers. and so forth. The cumulative 5-year replacement rate is given by the following function: .066x + .066(x - .066x) + .066[x - .066x - .066(x - .066x)] +.066 [x - .066x - .066(x - .066x) - .066 [x - .066x - .066 (x - .066x)]] +.066[x - .066x - .066(x - .066x) - .066[x - .066x - .066(x - .066x)] -.066 [x - .066x - .066 (x - .066x) - .066 [x - .066x - .066 (x - .066x)]]= 5(.066)x - 10 (.066)2x + 10 (.066)3x - 5 (.066)4x + (.066)5x =.28922134x Thus, the cumulative replacement function reaches 28.92% after 5 years. In practical terms, for a building with 25 teachers, the principal may replace (on average) 1.65 teachers in the 1st year, and a total of 7.23 teachers after 5 years, with NBPTS-certified teachers. The annual replacement rate declines every year because the principal has replaced a growing proportion of teachers with certified replacements. Thus, a building principal can replace at most 28.92% of the teaching staff (7.23 teachers in a building with 25 teachers) with NBPTS-certified teachers within 5 years, given the current rate at which experienced teachers turn over, and one additional NBPTS-certified teacher can be added in Year 6. If normal turnover means that a constant 21.3% of teachers have less than 5 years of teaching experience, 78.7% of the teaching staff, or 19.68 experienced teachers in a building of 25 teachers, could potentially be replaced with NBPTS-certified teachers. However, based on an average NBPTS passing rate of 59.7%, 11.75 of these teachers are already performing at the NBPTS level. The principal can improve student achievement by replacing the remaining 7.93 experienced teachers who are performing below the NBPTS standard with Board-certified teachers. The cumulative replacement function suggests that a building principal could not achieve the objective of replacing the

7.93 teachers any earlier than 6 years, even under the optimistic assumption that all of the teachers who leave voluntarily are the teachers who are performing below the NBPTS level. If high-performing teachers are just as likely as lowperforming teachers to leave the profession, the rate at which a building principal can replace low-performing teachers with Board-certified teachers is halved, and the process takes twice as long. Thus, any benefits of Board certification would be realized very slowly.

3. A study finding positive effects of teacher certification (Darling-Hammond et al. 2005) did not control for student or school fixed effects (Podgursky 2006).

References

- Angrist, J. D., and J. Guryan. 2004. Teacher testing, teacher education, and teacher characteristics. Paper read at One Hundred Sixteenth Annual Meeting of the American Economic Association, January 3–5, May, San Diego, CA.
- Boe, E. E., and D. M. Gilford, eds. 1992. *Teacher supply, demand, and quality: Policy issues, models, and data bases*. Washington, DC: National Academies Press.
- Bond, L., T. Smith, W. K. Baker, and J. A. Hattie. 2000. The certification system of the National Board for Professional Teaching Standards: A construct and consequential validity study. Greensboro, NC: Center for Educational Research and Evaluation, The University of North Carolina at Greensboro.
- Cavalluzzo, L. 2006. Is National Board certification an effective signal of teacher quality? Alexandria, VA: The CNA Corporation. Retrieved October 29, 2006, from http://www.cna.org/documents/CavaluzzoStudy.pdf
- Clotfelter, C. T., H. F. Ladd, and J. L. Vigdor. 2006. Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources* 41:778-820.
 ——. 2007. *How and why do teacher credentials matter for student achievement?* No city: National Center for Analysis of Longitudinal Data in Education Research.
- Cunningham, G. K., and J. E. Stone. 2005. Value-added assessment of teacher quality as an alternative to the National Board for Professional Teaching Standards: What recent studies say. In *Value added models in education: Theory* and applications, ed. R. Lissitz. (pp. 209-232). Maple Grove MN: JAM Press.
- Currie, J., and D. Thomas. 2001. Early test scores, socioeconomic status, school quality and future outcomes. *Research in Labor Economics* 20:103-32.
- Darling-Hammond, L., D. J. Holtzman, S. J. Gatlin, and J. V. Heilig. 2005. Does teacher preparation matter? Evidence about teacher certification, Teach for America, and teacher effectiveness. *Education Policy Analysis Archives*, 13(42). Retrieved April 17, 2010 from http://epaa.asu.edu/ojs/article/viewFile/147/273.
- Education Commission of the States. 2006. Synthesis of reviews of "The valueadded achievement gains of NBPTS-certified teachers in Tennessee: A brief

report." Denver, CO: Author. Retrieved October 30, 2006, from http://www.ecs.org/html/special/nbpts/PanelReport.htm

- Educational Testing Service. 2004. Where we stand on teacher quality. Princeton, NJ. Retrieved October 29, 2006, from http://ftp.ets.org/pub/corp/position_paper.pdf
- Ferguson, R. 1998. Teachers' perceptions and expectations and the black-white test score gap. In *The black-white test score gap*, ed. C. Jencks and M. Phillips, 273-317. Washington, DC: Brookings Institution Press.
- Finn, J. D., S. B. Gerber, C. M. Achilles, and J. Boyd-Zaharias. 2001. The enduring effects of small classes. *Teachers College Record* 193 (2):145-183.
- Goldhaber, D. 2002. The mystery of good teaching: Surveying the evidence on student achievement and teachers' characteristics. *Education Next* 2:50-5.
- Goldhaber, D., and E. Anthony. 2006. Can teacher quality be effectively assessed? National Board certification as a signal of effective teaching. Retrieved October 29, 2006, from http://www.urban.org/UploadedPDF/411271_teacher_quality.pdf
- ———. 2007. Can teacher quality be effectively assessed? National Board certification as a signal of effective teaching. *Review of Economics and Statistics* 89:134-50.
- Goldhaber, D., D. J. Brewer, and D. Anderson. 1999. A three-way error components analysis of educational productivity. *Education Economics* 7:199-208.
- Goldhaber, D., D. Perry, and E. Anthony. 2003. *NBPTS certification: Who applies and what factors are associated with success?* n.c.: University of Washington and The Urban Institute.
- Gordon, R., T. J. Kane, and D. O. Staiger. 2006. *Identifying effective teachers using performance on the job* (Discussion Paper 2006-01) Washington, DC: The Brookings Institution.
- Greenwald, R., L. V. Hedges, and R. D. Laine. 1996. The effect of school resources on student achievement. *Review of Educational Research* 66:361-96.
- Grissmer, D. W., A. Flanagan, J. Kawata, and S. Williamson. 2000. *Improving student achievement: What state NAEP test scores tell us*. Santa Monica, CA: RAND Corporation.
- Hakel, M. D., J. A. Koenig, and S. W. Elliott, eds. 2008. Assessing accomplished teaching: Advanced-level certification programs. Washington, DC: The National Academies Press.
- Hanushek, E. A. 1989. The impact of differential expenditures on school performance. *Educational Researcher* 18:45-50.
 - . 1992. The trade-off between child quantity and quality. *Journal of Political Economy* 100:84-117.

—. 1996. A more complete picture of school resource policies. *Review of Educational Research* 66:397-409.

- Hanushek, E. A., J. F. Kain, D. M. O'Brien, and S. G. Rivkin. 2005. *The market for teacher quality*. Cambridge, MA: National Bureau of Economic Research.
- Hanushek, E. A., J. F. Kain, and S. G. Rivkin. 1999. Do higher salaries buy better teachers? National Bureau of Economic Research Working Paper No. 7082.
- Harris, D. N., and T. R. Sass. 2007. The effects of NBPTS-certified teachers on student achievement. Madison, WI, and Tallahassee, FL: University of Wisconsin and Florida State University.
- Hedges, L. V., R. D. Laine, and R. Greenwald. 1994. Money does matter somewhere: A reply to Hanushek. *Educational Researcher* 23:9-10.
- Ingersoll, R. M. 2001. Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal* 38:499-534.
- Ladd, H. F., T. R. Sass, and D. N. Harris. 2007. The impact of National Board certified teachers on student achievement in Florida and North Carolina: A summary of the evidence prepared for the National Academies Committee on the evaluation of the impact of teacher certification by NBPTS.
- Levin, H. M. 1988. Cost-effectiveness and educational policy. *Educational Evalua*tion and Policy Analysis 10:51-69.
- Levin, H. M., G. V. Glass, and G. Meister. 1987. A cost-effectiveness analysis of computer-assisted instruction. *Evaluation Review* 11:50-72.
- Levin, H. M., and P. J. McEwan. 2001. Cost effectiveness analysis: Methods and applications. Thousand Oaks, CA: SAGE.
 - ——, eds. 2002. Cost-effectiveness and educational policy. Larchmont, NY: Eye On Education.
- McCaffrey, D. F., and S. G. Rivkin. 2007. Empirical investigations of the effects of National Board of Professional Teacher Standards certified teachers on student outcomes: A report prepared for the National Research Council.
- McColskey, W., J. H. Stronge, T. J. Ward, P. D. Tucker, B. Howard, K. Lewis, and J. L. Hindman. 2006. Teacher effectiveness, student achievement, and National Board certified teachers: A comparison of National Board certified teachers and non-National Board certified teachers: Is there a difference in teacher effectiveness and student achievement?: National Board for Professional Teaching Standards. Retrieved October 30, 2006, from http://www.nbpts.org/UserFiles/File/Teacher_Effectiveness_Student_Achievement_and_National_Board_Certified_Teachers_D_-_McColskey.pdf
- Murnane, R. J., J. D. Singer, and J. B. Willett. 1988. The career paths of teachers: Implications for teacher supply and methodological lessons for research. *Educational Researcher* 17:22-30.
- Murnane, R. J., J. Willet, and F. Levy. 1995. The growing importance of cognitive skills in wage determination. *Review of Economics and Statistics* 77:251-66.

- National Board for Professional Teaching Standards. 2006a. *About us*. Retrieved October 29, 2006, from http://www.nbpts.org/about_us
- ———. 2006b. The certification process. Retrieved October 26, 2006, from http:// www.nbpts.org/become_a_candidate/assessment_process
- ——. 2006c. National Board certification: Eligibility verification forms and instructions. Retrieved October 26, 2006, from http://www.nbpts.org/ UserFiles/File/2006_Eligibility_Verif.PDF
- Neal, D., and W. R. Johnson. 1996. The role of pre-market factors in black-white wage differences. *Journal of Political Economy* 104:869-95.
- Nye, B, L. V. Hedges, and S. Konstantopoulos. 2001. Are effects of small classes cumulative? Evidence from a Tennessee experiment. *Journal of Educational Research* 94 (6):336-345.
- Nye, B., S. Konstantopoulos, and L. V. Hedges. 2004. How large are teacher effects? *Educational Evaluation and Policy Analysis* 26:237-57.
- O'Neill, J. 1990. The role of human capital in earnings differences between black and white men. *Journal of Economic Perspectives* 4:25-46.
- Podgursky, M. 2006. Response to "Does teacher preparation matter?" Retrieved October 30, 2006, from http://www.nctq.org/nctq/research/1114011044390.pdf
- Rice, J. K., and L. J. Hall. 2008. National Board certification for teachers: What does it cost and how does it compare? *Education Finance and Policy* 3:339-73.
- Sanders, W. L., J. J. Ashton, and S. P. Wright. 2005. Comparison of the effects of NBPTS certified teachers with other teachers on the rate of student academic progress. Cary, NC: SAS Institute, Inc.
- Stancavage, F. B., J. H. Mitchell, V. Bandeira de Mello, F. E. Gaertner, A. K. Spain, M. L. Rahal, and T. Rahman. 2006. National Indian education study part II: The educational experiences of fourth- and eighth-grade American Indian and Alaska Native students (NCES 2007-454) Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.
- Stone, J. E. 2006. The value-added achievement gains of NBPTS-certified teachers in Tennessee: A brief report. Retrieved October 30, 2006, from http://www. education-consumers.com/briefs/stoneNBPTS.shtm
- The National Academies. 2008. National Board certification identifies strong teachers, but many school systems are not using Board-certified teachers' expertise. Washington, DC: The National Academies. Retrieved July 12, 2009, from http://www8.nationalacademies.org/onpinews/ newsitem.aspx?RecordID=12224
- U.S. Department of Education. 2007. *Common core of data: State nonfiscal survey*. Washington, DC: National Center for Education Statistics. Retrieved August 12, 2007, from http://nces.ed.gov/ccd/bat/

- Vandevoort, L. G., A. Amrein-Beardsley, and D. C. Berliner. 2004. National Board certified teachers and their students' achievement. *Education Policy Analysis Archives* 12:1-44.
- Wenglinsky, H. 2001. Teacher classroom practices and student performance: How schools can make a difference (Research Report RR-01-19) Princeton, NJ: Educational Testing Service.
- Wilson, S. M., and R. E. Floden. 2003. Creating effective teachers: Concise answers for hard questions. An addendum to the report teacher preparation research: Current knowledge, gaps and recommendations. Denver, CO: Education Commission of the States; American Association of Colleges for Teacher Education.
- Winship, C., and S. D. Korenman. 1999. Economic success and the evolution of schooling and mental ability. In *Earning and learning: How schools matter*, ed. S. E. Mayer and P. E. Peterson, 49-78. Washington, DC: Brookings Institution Press.
- Wright, P., S. Horn, and W. Sanders. 1997. Teachers and classroom heterogeneity: Their effects on educational outcomes. *Journal of Personnel Evaluation in Education* 11:57-67.
- Yeh, S. S. 2006. *Raising student achievement through rapid assessment and test reform*. New York: Teachers College Press.
- ——. 2007. The cost-effectiveness of five policies for improving student achievement. American Journal of Evaluation 28:416-36.
- ——. 2008. The cost-effectiveness of comprehensive school reform and rapid assessment. *Education Policy Analysis Archives* 13. Retrieved from http:// epaa.asu.edu/epaa/v16n13/
- 2009a. Class size reduction or rapid formative assessment? A comparison of cost-effectiveness. *Educational Research Review* 4:7-15.
- —. 2009b. The cost-effectiveness of raising teacher quality. *Educational Research Review* 4:220-32.
- . Forthcoming. Understanding and addressing the achievement gap through individualized instruction and formative assessment. *Assessment in Education*.
- Yeh, S. S., and J. Ritter. 2009. The cost-effectiveness of replacing the bottom quartile of novice teachers through value-added teacher assessment. *Journal of Education Finance* 34:426-51.

Bio

Stuart S. Yeh is an associate professor and coordinator of the Evaluation Studies Program in the department of organizational leadership, policy, and development at the University of Minnesota. He recently completed cost-effectiveness analyses of 22 approaches for raising student achievement.