

The Confidence Men

Selling adequacy, making millions

Checked:

Picus and Associates. 2006.

An Evidence-Based Approach to School Finance Adequacy in Washington.

Checked by Eric A. Hanushek

Lawsuits aimed at compelling legislatures to increase school funding have been filed in some 42 states. Courts have found for the plaintiffs in more than half of the cases on the grounds that schools are not “adequately” funded (see Figure 1). These decisions have, in effect, changed the way education appropriations are made, moving decisionmaking from legislatures to the courts. Instead of flowing from the political process, determinations of adequate appropriations come from judges who are informed by paid consultants. Recently, adequacy plaintiffs have suffered some serious setbacks (see *legal beat*, page 11). Undaunted, they soldier on.

In the state of Washington, adequacy plaintiffs filed a new lawsuit in early 2007 that is expected to rely heavily on a report prepared at the request of a gubernatorial-appointed commission, Washington Learns. This report, “An Evidence-Based Approach to School Finance Adequacy in Washington,” claims to present scientific evidence of exactly what needs to be done to bring every child to proficiency as defined under state and federal law. The advance, if true, would go far beyond this specific court case and could revolutionize American education. For if, indeed, we now know how to create an effective educational system, and only the funds



ILLUSTRATION / DANIEL BAXTER

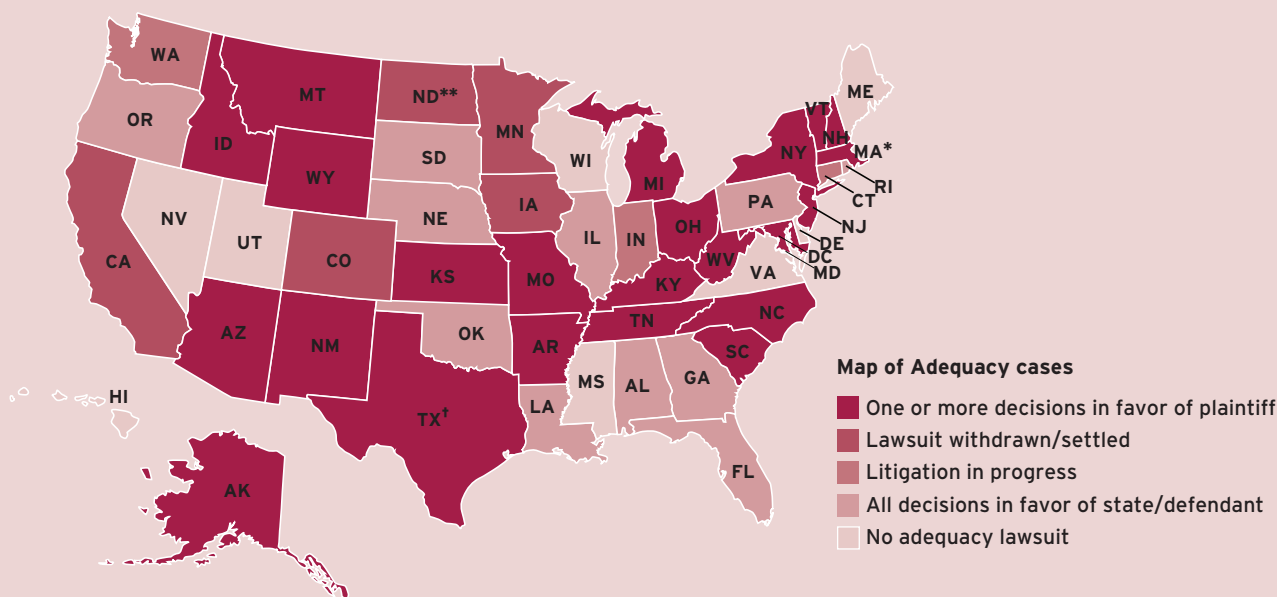
are lacking, then the country’s education problems can be solved.

The analysts who purport to have assembled this knowledge are led by two professors, Lawrence Picus of the University of Southern California and Allan Odden of the University of Wisconsin. The two formed a consulting

group known as Picus and Associates and have become increasingly popular among groups seeking to expand school spending, be they plaintiffs in funding lawsuits, teachers unions, or state departments of education. The Washington Learns commission asked Picus and Associates to recommend

The Adequacy Landscape (Figure 1)

Adequacy lawsuits have been filed in 42 states since the early 1980s, and plaintiffs have won in over half of these states.



* Defendant win in *Hancock v. Driscoll* (2005)

** On stay pending review of governor's plan to increase spending

† Defendant win in *West Orange-Cove Consolidated ISD v. Neeley* (2005)

SOURCES: Martin R. West and Paul E. Peterson, eds., *School Money Trials: The Legal Pursuit of Educational Adequacy* (Brookings, 2007); National Access Network, Teachers College, Columbia University, <http://www.schoolfunding.info/>

policy changes that will place the state's education system on a sound footing. Specifically, Picus and Odden answer the question, "What are the high-impact education programs and strategies that will allow every school to provide each Washington student with the opportunity to learn at or above proficiency on state standards as measured by the Washington Assessment of Student Learning, with proficiency standards calibrated over time to those of NAEP [National Assessment of Educational Progress], or even the performance of students in other countries?"

Even if only the state of Washington were getting precise, scientific answers to such critical questions, the work of the Picus-Odden team would command the attention of national policy-makers. But the consulting group has

already established a national reputation for its ability to ascertain, scientifically, what needs to be done in education—and precisely how much it costs to do it—through prior studies along much the same lines prepared for policy-makers in Kentucky, Arkansas, Arizona, and Wyoming.

Of course, the evidence base does not change very rapidly, as is evident from the various reports, which were carried out between 2003 and 2006. The 2006 study conducted for Washington Learns has an extensive bibliography, some 260 entries. But, since the production of the cost study for Kentucky in 2003, only 30 new references were added (including the obligatory reference to Thomas Friedman's *The World Is Flat*). So similar are the studies that at times it seems the copy

function of the Microsoft word processor deserves to be listed among the authors.

The ease with which one report can build on another does not seem to translate into efficiencies in the consulting group's operations, at least as reflected in the fees charged. According to available records, the Kentucky study, conducted in 2003, was executed for \$349,000. Arkansas's original study, conducted the same year, cost about the same initially but rose to over twice that amount (\$800,000) when the authors accepted a commission to ascertain whether districts used their extra money in a way consistent with the consultants' evidence-based policies. Wyoming, a small but rich state, was asked to pay \$1,260,000 in 2005 for a calibration of

check the facts

CONSULTANTS HANUSHEK

its finance formula along evidence-based lines and a subsequent implementation study. Washington, in 2006, managed to squeeze the price back down to the total Arkansas figure, although Washington could get only the original evidence-based analysis without the follow-up.

Even the Wyoming deal is a bargain, however, if the study can answer the question posed by the Washington Learns commission. After all, we spend some \$500 billion nationally on K–12 education, and even small improvements applied to the nation's schools could quickly cover the study costs.

The Picus-Odden Miracle

The frequency with which education policy initiatives of the past, though based on high hopes, have yielded disappointing results when implemented in the field has led to rather low expectations. As a general rule, in education discussions a policy is considered successful if an evaluation has shown it to have a statistically significant positive effect on student outcomes. Translated, there must be a high degree of certainty that positive results were not simply the result of chance. But just finding that some policy is likely to improve student outcomes does not mean that the improvement will reach the high levels sought by Washington Learns, or by others with similar views about what students should know. The research would have to provide evidence about the magnitude of improvements in achievement that can be expected, and these improvements would have to be large.

Such evidence is precisely what Picus and Odden purport

to provide for their fees. They have combed the research evidence to provide rather precise, and remarkable, predictions about the achievement effects of programs whose power has apparently escaped the attention of almost all other researchers.

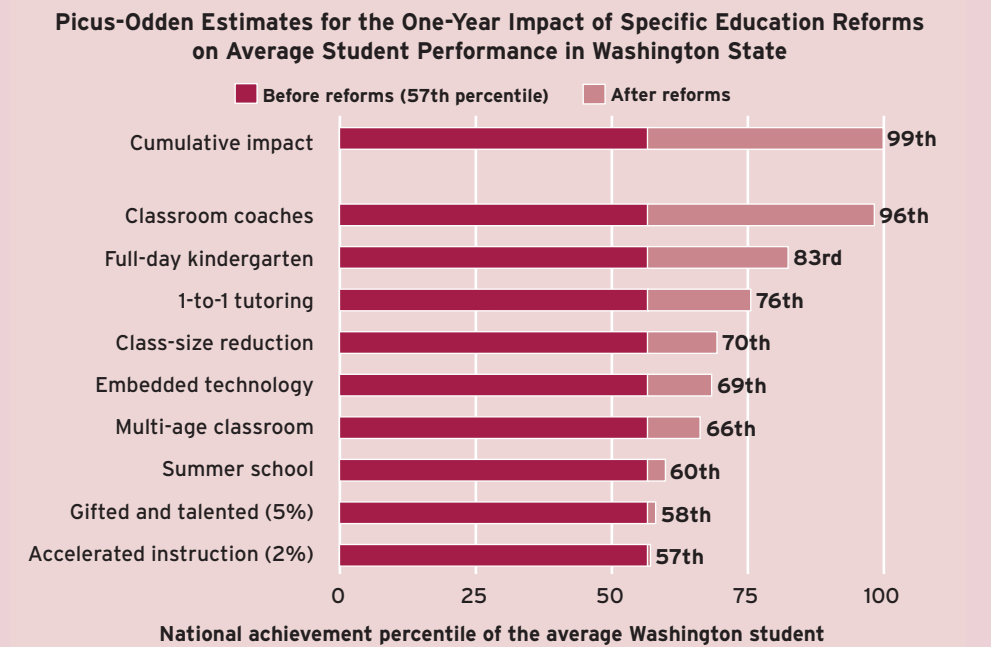
Picus and Odden convey the magnitude of achievement gains that can be expected from their evidence-based policies through a unit of measurement known as effect size. Effect size is the change in standard deviations of achievement that can be expected, according to the research, from the introduction of a given policy. In itself,

that step is perfectly acceptable, as the unit is widely used in education research.

Discussion of effect sizes and standard deviations is something most policymakers, even when introduced to the concepts in their undergraduate statistics course, would rather avoid. But some heuristics will help to understand the essence of effect sizes and make clear the import of the Picus and Odden evidence. The National Assessment of Educational Progress (NAEP) measures achievement in different grades and attempts to put it on a common scale. One full standard deviation (an effect size of 1.0) is roughly

Science Fiction (Figure 2)

Picus and Odden's estimates of program effects yield implausible results. If Washington State were to pursue all their recommended strategies, the average student in the state, now at the 57th percentile of all U.S. students, supposedly would be performing at the 99th percentile within a year.



Note: The impact of the above interventions accounts for the fact that some programs would affect only a subset of all Washington students (e.g., gifted or disadvantaged). Where Picus-Odden reported ranges of possible impacts, the average of the maximum and minimum was used as a point estimate. The pre-reform performance percentile is calculated from the performance of 4th- and 8th-grade Washington students on the math and reading portions of the 2005 National Assessment of Educational Progress.

SOURCES: Author's calculations from Odden, Picus, Goetz, and Feranich (2006); National Center for Education Statistics



Most Americans would be extraordinarily satisfied with average gains of one full standard deviation for a school or district. Picus and Odden claim to be able to do that three or possibly even six times over for all students in Washington.

equal to the average difference in test score performance between a 4th grader and an 8th grader. In other words, it is a big effect, as the typical 8th grader has learned quite a bit since 4th grade.

By this perspective, any education strategy that in a single year can raise average achievement of a large aggregate of students by one full standard deviation must be taken very seriously. Pursued systematically, it could eliminate the persistent ethnic test-score gap (which is about one full standard deviation) or could vault the math and science performance of U.S. students beyond counterparts in Korea, Singapore, and Japan (who are about one-half of a standard deviation ahead now).

Picus and Odden identify strategies they claim can do that, and much more. They provide “scientific evidence” to support the claim that a specific set of policies can shift average student performance upward by *three to six standard deviations*, an extraordinary gain. The policies they identify include providing a year of full-day kindergarten, reducing class size to 15 students through grade 3, using multi-age classrooms, hiring classroom coaches, employing one-to-one tutoring for disadvantaged students, getting half of the students eligible for free and reduced-price lunch to attend summer school, embedding technology

within the classroom, creating a gifted and talented program for the top 5 percent of all students, and accelerating instruction for the 2 percent of students capable of benefiting from it (see Figure 2). The range in claimed impact reflects the fact that they sometimes admit to uncertainty about the exact effect size from a specific program.

Most Americans would be extraordinarily satisfied with average gains of one full standard deviation for a school or district. Picus and Odden claim to be able to do that three or possibly even six times over for all students in Washington. After their policies are fully implemented in Washington, Albert Einstein, were he not participating in these programs, would find himself achieving at or below the state average.

This can all happen within one year of application of these policies, the consultants say. But they would not give these programs just a single year. They would apply them, where appropriate, across all years of schooling. (Full-day kindergarten, for example, happens just once for each student.) If one then assumes a cumulative impact from giving students not just a single application but continuing treatment through grade 12, the gains reach astronomical proportions, somewhere in the range of 23 to 57 standard deviations.

The Truth behind the Numbers

This, of course, is the stuff of science fiction novels, not research-based school policies. How does a well-funded study, conducted by scholars of national reputation, reach such startling conclusions? The procedure is roughly as follows:

- 1) Find a study, preferably one that has some surface credibility, that shows that a particular intervention had a certain effect on a particular group of students.
- 2) Ignore all the studies of that intervention that show a smaller effect or no effect at all.
- 3) Interpret the study as identifying a true causal relationship, not just a correlation or association.
- 4) Finally, assume that the conditions that produced the very large effect can be perfectly replicated throughout the state of Washington.

Take full-day kindergarten, for example, which Picus-Odden estimate to have by itself an impact of 0.77 standard deviations on student achievement for advantaged and disadvantaged students alike. (In NAEP terms, this by itself would be equivalent to three full years of later schooling.) Picus and Odden cite a 1997 meta-analysis by John Fusaro that shows such an impact. But they disregard Fusaro’s own strong warning: “A seductive conclusion from these results is that attendance at full-day kindergartens causes students to achieve at a higher level than attendance at half-day kindergartens. It is imperative, however, that we strenuously resist succumbing to such a seduction.” Meanwhile, Picus and Odden ignore a large body of literature that shows little impact on advantaged students and smaller impacts on disadvantaged ones, to say nothing of the empirical reality that the 56 percent of students currently attending schools that have full-day kindergarten do not surpass the remaining 44 percent attending schools without full-day kindergarten by anything like a 0.77 margin. Note, for example, that black

check the facts

CONSULTANTS HANUSHEK

students and disadvantaged students are currently more likely to attend schools with full-day kindergarten than more advantaged students.

Or take summer school, which Picus and Odden estimate would have an effect size of 0.45 standard deviations. This policy recommendation is apparently based on a single study in 2000 of the Voyager summer learning program, although they note that a major meta-analysis suggests widely varying effect sizes from the evaluations of different studies. Note also that in Odden's peer review in 2004 of William Driscoll's and Howard Fleeter's Ohio study of the costs of bringing all students to proficiency in math and reading in order to comply with NCLB, he castigates the study's authors, who called for expanded summer school, because they "reference no research to support this assertion, when in fact most research shows that summer school as typically administered has little if any impact on learning."

And so on. There is little reason to believe that the effect sizes identified in their work indicate what can be expected from implementing any policy on a broad scale.

The approach of Picus and Odden to policies is simple: if a program shows a large positive effect in one study, it should immediately be implemented across the state. Indeed, they assert in public hearings that adopting anything less than the complete set of recommended programs would constitute an inadequate program, and that they would testify to the inadequacy in court.

Are Costs Important?

The primary purpose of reviewing the evidence on programs is to establish the cost of providing a new and improved (adequate) education. The various programs suggested by Picus and Odden have very different price tags associated with them. They make it hard to tell from their report what prices might go

expenditure of \$7,800 per pupil and average teacher compensation (salary plus benefits) of \$60,000 for the state of Washington. We can first translate these into the cost per recipient for each program based on resource demands and then take into account the proportion of all students who receive the program. The results show wide variations in costs. For example, full-day kindergarten would increase average spending in the state by \$154 to \$300 per student, while the K-3 class size reduction would increase average spending by \$410 to \$800 per student. Some programs have no obvious costs. For example, multi-age classrooms might reasonably be taken as free. Similarly, changes in curriculum do not in general have significant added costs (past, say, an initial teacher-training period). Other programs, such as skipping grades, would actually save money, since students would spend 12 rather than 13 years in the system.



The approach of Picus and Odden to policies is simple: if a program shows a large positive effect in one study, it should immediately be implemented across the state. Indeed, they assert in public hearings that adopting anything less than the complete set of recommended programs would constitute an inadequate program.

These patterns are repeated when one goes to the other "evidence-based" recommendations of Picus and Odden, including class size reduction and professional development. Their estimate of the benefits of professional development comes directly from the professional association representing those who supply professional development.

with each of the programs, because they bury the costs within the staffing of each prototypical school. It is, nonetheless, relatively easy to obtain reasonable cost estimates for each program.

The basic building blocks for calculating the cost per pupil of the various policies Picus and Odden propose are the approximate average

Once program costs are separated, one can immediately see the variation that exists and can make judgments about where money is better (more efficiently) spent. A simple cost calculation gives the improvements in student achievement (measured again in standard deviations) that could, by the Picus and Odden estimates of benefits, be

expected for a \$100 addition to spending per pupil from each of the separate programs. By their low-end estimates of benefits (which total to just three standard deviations), each \$100 spent on classroom coaches would be expected to yield at least a 0.25 standard deviations gain in achievement, very similar to the expected gain for full-day kindergarten. Their class-size reduction proposal would yield only *one-sixth* that gain, or 0.04 standard deviations, an effect very similar to that for one-to-one tutoring.

Using the upper range of their effect size estimates, \$100 spent on classroom coaches would yield a gain of over one-half standard deviations in student achievement, and one-to-one tutoring would yield a one-quarter standard deviations improvement. According to their estimates, some of their favored programs (such as classroom coaches) are more than 10 times as cost efficient as others, such as class size reduction for K–3.

Picus and Odden contend that all programs, regardless of cost, must be simultaneously undertaken. But it is clear that the programs they identify have very different expected returns on spending. Their method of distributing costs through their prototypical schools provides no information on the relative efficiency of investing in the various components. Nor does it say anything about the costs of improving outcomes if done efficiently. Unless there are unlimited funds to spend on educational programs, it would not make sense to put the money into all the programs without regard to cost.

What Are States Paying For?

Cost estimates are an important component in the politics of court and legislative deliberations on schools. The adequacy debates are typically motivated by obvious and real shortfalls in the achievement of a state's students, but a combination



Why do different states and organizations pay ever-increasing amounts to

see this research review when Google would bring up the most recent version immediately and without expense?

Clients want a bottom-line statement about how much spending would provide an adequate education, and they want this cost estimate attached to their specific state.

of naive concerned citizens and self-interested parties invariably pushes to translate these debates into a simple dollar figure. Such translation is salient for courts and legislatures and both simplifies and focuses the issue for the media.

What Picus and Odden provide in their reports is essentially a selective review of the published literature on program effects. Why do different states and organizations pay ever-increasing amounts to see this research review when Google would bring up the most recent version immediately and without expense? The answer is simple. Clients want a bottom-line statement about how much spending would provide an adequate education, and they want this cost estimate attached to their specific state. Few people care about the “studies” on which consultants base their reports, or even their validity, because nobody really expects schools to implement these specific programs if given extra funding. Clients simply want a requisite amount of scientific

aura around the number that will become the rallying flag for political and legal actions.

Summing the added cost of the separate programs suggested by Picus and Odden, I estimate that the overall plan, if fully applied, would increase average spending in Washington by \$1,760 to \$2,760 per student, or 23 to 35 percent. This estimate of the increased spending necessary to achieve “adequacy” is very similar to the percentage increases they have recommended to other states, and numbers like these will presumably become part of the headlines surrounding the new court case.

But pity the poor states that actually implement the Picus and Odden plan. They are sure to be disappointed by the results, and most taxpayers (those who do not work for the schools) will be noticeably poorer.

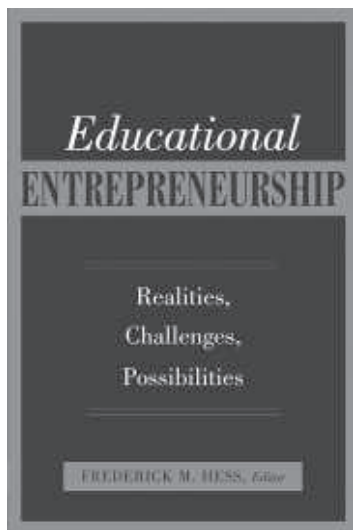
Eric A. Hanushek is a senior fellow at the Hoover Institution, Stanford University, and a member of its Koret Task Force on K–12 Education.

NEW from **Harvard Education Press**

EDUCATIONAL ENTREPRENEURSHIP

REALITIES, CHALLENGES, POSSIBILITIES

EDITED BY FREDERICK M. HESS

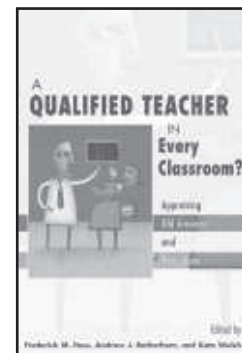
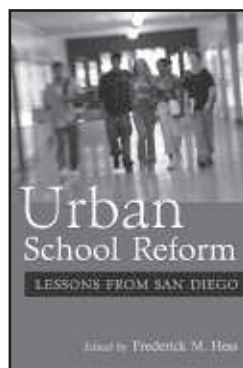
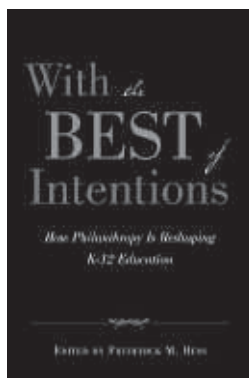


“Educational entrepreneurs are proving that the calcified delivery system of public schooling can be shaken up and retooled for the twenty-first century. Hess and his colleagues look at the phenomenon from every angle in this rich assortment of essays. This is a book that provides plenty of fuel for discussions about where school reform is headed.”

—NELSON SMITH, PRESIDENT, NATIONAL ALLIANCE FOR PUBLIC CHARTER SCHOOLS

2006, 312 pages
Paperback \$29.95 / ISBN 1-891792-25-3
Library \$59.95 / ISBN 1-891792-26-1

ALSO FROM FREDERICK M. HESS



To order, call 1-888-437-1437 or visit us online at www.harvardeducationpress.org