

Data Vacuum

Only larger voucher experiments will yield answers

School Vouchers: Examining the Evidence

By Martin Carnoy

Economic Policy Institute, 2002.

Rhetoric Versus Reality: What We Know and What We Need to Know About Vouchers and Charter Schools

By Brian Gill, P. Michael Timpone, Karen Ross, and Dominic Brewer

RAND Corporation, 2001.

School Vouchers: Publicly Funded Programs in Cleveland and Milwaukee

General Accounting Office, 2001.

"The Hidden Research Consensus for School Choice"

By Jay P. Greene, in *Charters, Vouchers, and Public Education*,

edited by Paul Peterson and David Campbell

Brookings Institution Press, 2001.

"Market-Based Reforms in Urban Education"

By Helen F. Ladd

Paper presented at the Seminar on Creating Change in Urban Public Education, December 7-8, 2001, Cambridge, Massachusetts.

"Vouchers for Elementary and Secondary Education"

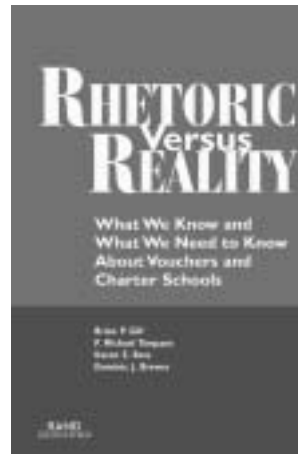
By Isabel Sawhill and Shannon Smith, in *Vouchers and the Provision of Public Services*, edited by Eugene Steuerle, Van Doorn Ooms, George Peterson, and Robert Reischauer

Brookings Institution Press, 2000.

"What Research Can Tell Policymakers about School Choice"

By Paul Teske and Mark Schneider

Journal of Policy Analysis & Management, 2001.



In their sheer volume, reviews of the small body of empirical work on school vouchers are beginning to eclipse the research literature itself. It is not often that so much is written about so little, but scholars are jostling to have the final word on what we know about vouchers as the Supreme Court prepares to rule on their constitutionality.

All of the reviewers decry the contentious political and ideological haranguing that surrounds the public debate over vouchers, and many blame the media for conspiring in the frenzy. As a corrective, each promises a clear-headed, impartial account of the empirical findings on school choice. Given the expressed motives of the reviews and the modest size of their subject matter, one would hope that a consensus would emerge.

Would that it were so. The authors find little about which to agree. Only in the face of overwhelming and unambiguous evidence does any consensus emerge. When scholars survey the nascent empirical literature on school choice, they see very different things and discern very different lessons. And

they will continue to do so until more, and better, data are collected from larger, better-financed voucher programs.

The Reviews

Stanford University education professor Martin Carnoy declares his intent to interject a "balanced perspective" in an empirical literature written "mainly by researchers who openly and actively support vouchers" and media that "report results from these analyses without necessary caveats and alternative views." Carnoy is deeply skeptical about the positive test-score impacts observed for African-Americans in the randomized field trials conducted in New York City; Dayton, Ohio; and Washington, D.C. He concludes: "The question to ask is not whether these latest Peterson-group reports overestimate private school effects, but by how much." (Full disclosure: Paul Peterson, editor-in-chief of *Education Next*, and I were among the scholars conducting these evaluations.) Carnoy is no less critical of Jay Greene's recent analysis of the Florida voucher program—or what Carnoy calls "the latest round of voucher advocacy research." (See Letters, p. 7, for the debate between Greene and Carnoy. See "The Looming Shadow," *Research*, Winter 2001, for Greene's study of the Florida A+ program.)

Like Carnoy, Helen Ladd argues that the evidence on achievement is at best "preliminary," and if it supports the claims of voucher advocates, it does so only under very restrictive conditions and for a very small subset of the urban poor. However, the Duke University economist is less distracted than Carnoy by the supposed misrepresentations of voucher advocates posing as researchers and more impressed by an empirical lit-

erature that consistently turns up negative findings. Ladd argues, for instance: “The evidence simply does not support the claims of those who argue in favor of more parental choice and competition on the instrumental grounds that it will make an education system significantly more productive than it would otherwise be.” Though the only data on large-scale voucher programs she finds convincing come from her own study of the New Zealand public-choice program and Martin Carnoy’s work in Chile, Ladd confidently asserts that vouchers will not help those public schools that are having the hardest time of it. “The bottom line is clear: Large-scale expansion of parental choice and competition will not, by itself, solve the problems of the most distressed urban schools.” If anything, the effects of vouchers on the schools that are “left behind” are surely “adverse.”

Much like Carnoy’s and Ladd’s reviews, RAND’s book-length summary of the literature emphasizes the limitations of the existing empirical work on school choice. Its general assessment of the findings, however, is considerably more positive. According to Brian Gill, P. Michael Timpane, Karen Ross, and Dominic Brewer, the book’s authors, all of the empirical evidence supports the contention that vouchers improve parental satisfaction. The authors say that the existing research on student achievement, while still preliminary, bodes well for vouchers. “Small-scale, experimental privately funded voucher programs targeted to low-income students suggest a possible (but as yet uncertain) modest achievement benefit for African-American students after one or two years in voucher schools.” Contrary to the claims that voucher programs “cream” the best students from public schools, RAND’s scholars find that voucher programs have successfully placed “low-income, low-achieving, and minority students in voucher schools.” They raise concerns, however, about the underrepresentation of students with disabilities and poorly educated parents in private schools. They

also find that while targeted voucher programs may “modestly” alleviate racial segregation patterns in urban cities, the evidence from large-scale, unregulated choice programs suggests that vouchers may increase levels of social stratification.

In August 2001, the General Accounting Office issued a tepid report on the publicly funded school voucher programs in Cleveland and Milwaukee. (For reasons that are never fully explained, the report ignores all evidence from privately financed programs in the United States as well as all choice

**When scholars survey
the nascent empirical
literature on school
choice, they see
very different things
and discern very
different lessons.**

initiatives conducted abroad.) Both the Cleveland and Milwaukee programs successfully targeted poor, lower performing, and predominantly minority students in public schools. On average, voucher students were placed in smaller classes in smaller schools with less experienced teachers. The GAO report describes the research on student achievement as, at best, mixed and preliminary. While some “contract researchers found little or no significant improvement in voucher students’ achievement,” other investigators “found some positive effects.” The report’s conclusion laments the generally poor quality of the existing studies and advocates state funding for the collection of more achievement data and the conduct of studies that yield higher response rates.

Brookings Institution scholars Isabel Sawhill and Shannon Smith give the existing literature on school vouchers a cautiously positive review. Virtually all

studies on school vouchers, Sawhill and Smith write, show that private school parents are markedly more satisfied with their children’s school than public school parents. But after reviewing the reports issued by a trio of research teams studying the Milwaukee program—one headed by the University of Wisconsin’s John Witte, another by Harvard University’s Paul Peterson, and a third by Princeton University’s Cecilia Rouse—Sawhill and Smith claim that “it is simply not possible at the current time to render a clear verdict on the outcome of the experiment.” Finally, Sawhill and Smith anticipate that vouchers may lead to “a greater segmenting of recipients by race, ethnicity, income, or ability,” and hence recommend that further research be conducted on the subject.

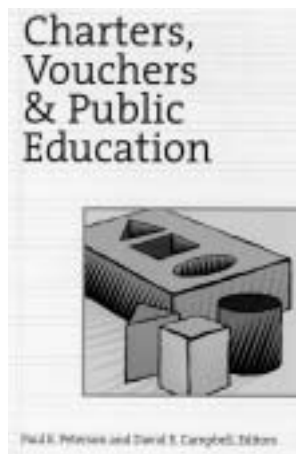
The reading of the literature by Paul Teske and Mark Schneider, political scientists at the State University of New York at Stony Brook, is more optimistic still. Teske and Schneider note that the existing empirical work on school vouchers is quite positive on a variety of issues: academic considerations appear paramount when parents choose schools; voucher recipients are more satisfied with their schools than their peers within public schools; and vouchers lead to “clear performance gains for some groups of students using the vouchers, particularly blacks, compared with the control group.” Teske and Schneider caution policymakers, however, that a rising tide of evidence suggests that school vouchers may lead to greater social stratification and racial segregation.

Finally, Jay Greene sees “uniformly positive” findings in the school voucher studies. While “one would never know it from the media coverage,” there exists a “hidden consensus” among scholars that vouchers work. Students who use vouchers score significantly higher on test scores than their public school peers—just as they are more tolerant and their parents are more satisfied. Though enterprising parents among the disadvantaged are more able to avail themselves of choice than others, Greene

observes that this is no less true of “food stamps, Temporary Assistance for Needy Families, and virtually all other antipoverty programs.” He notes that, although few studies have examined the impact of choice on public school students, most every finding to date suggests that vouchers, rather than adversely affecting students who are “left behind” in public schools, actually lead to gains for public and private school students. Compared with other education interventions, Greene concludes, “school choice has been thoroughly and carefully studied” and has generated “consistently positive results.”

This sweep, from Carnoy to Greene, is hardly exhaustive. Indeed, the reviews keep coming. Paul Hill and Tom Lovell, both senior scholars at the Brookings Institution, recently acquired a million-dollar grant from the Bill and Melinda Gates Foundation to assemble a panel of experts on school choice, with the goal of writing the definitive statement on what we know and what requires additional study. The National Working Commission on Choice in K–12 Education, as it is called, intends to trump all of the reviewers listed above. Stay tuned: two years from now, the commission will release its conclusions.

For now, though, let’s look at what these reviews add up to. On one issue there is general agreement: parents who use vouchers are more satisfied with their private schools than are parents who apply to voucher programs but remain in public schools. Everywhere else, however, the reviews turn up inconsistencies and disagreements. While RAND, Teske, Schneider, and Greene suggest that the best evidence on student achievement comes from the new randomized field trials, what some call the “gold standard” of social-scientific research, Carnoy and Ladd remain fixated on two sources of potential contamination—attrition rates and response bias—and are inclined to dismiss (or at least heavily discount) the findings. While Teske and Schneider argue that



parents are principally motivated by academic concerns when choosing schools, Ladd sees considerable evidence that parents select schools on the basis of their ethnic composition. Though RAND concludes that “nothing is yet known empirically about the civic socialization effects of voucher and charter schools,” Greene lists a batch of studies (some of which were published after the RAND report went to press) that consistently show that private schools are doing a better job of teaching tolerance. While many of the reviews anticipate the possibility of social stratification in large-scale voucher programs, Ladd sees it as a foregone conclusion.

The reviewers disagree about the content of the empirical findings, their implications for different kinds of choice initiatives, and the quality of alternative research designs. They disagree, if only implicitly, about the relevance of choice experiments conducted abroad. They disagree about which issues require further study and which have been settled conclusively. In searching for consensus, these reviews only underscore how much disagreement lingers.

The Chimera of “Objectivity”

Lawrence Hedges of the University of Chicago, in a recent report on education research, identified the basic problem facing policymakers who are considering

choice initiatives. “To put it bluntly,” Hedges wrote, “in the face of diverse and apparently conflicting information about the potential effects of a policy, it is difficult to know which studies to believe.” Each of the reviews summarized above tried to rectify this situation. Unfortunately, disagreements among the reviewers only replicate controversies among primary researchers, and the clouds have yet to part.

All of the reviewers recognize a need to inject a measure of reason into the politically charged debate over school vouchers. RAND puts it this way: if policymakers are to make an informed judgment on vouchers, they need “a thorough and objective empirical assessment” on the literature. RAND promises “accurate data,” “careful objective analysis,” and “a clear picture of the choices [we] face in educating America’s citizens.” RAND is quite sure that it has it right. The trouble, though, is that all of the other reviewers think they have it right as well. Given the manifest contradictions across the reviews, the defense of impartiality begins to break down.

Rhetorically, it is perfectly obvious why all of these reviews stake a claim to objectivity. All hope to distance themselves from the ideological warring between die-hard advocates and committed critics of school choice. All want to gain traction in a debate that has become increasingly politicized, and the obvious way to do so is to appeal to reason; to plead with the reader that if he would just set aside his ideological precommitments, he would see exactly what the reviewer sees. Then, with the reviewer’s credentials in order, and the reader’s trust secured, genuine progress can be made.

Such claims to objectivity, however, are misplaced. In their attempts to provide neutral accounts of the empirical literature on school choice, all of these reviewers necessarily fail. They fail not because of their individual flaws—though some reviews certainly are more defensible than others. They fail because

the act of assigning meaning to facts is unavoidably subjective, especially when the facts come from such a small empirical literature. While the laws of science guide the practice of research, they cannot say what constitutes “persuasive” findings. They cannot specify the threshold of evidence required to “prove” a hypothesis. They cannot determine which facts have important policy implications and which are less consequential. As the education philosopher Francis Schrag notes, when evaluating (as opposed to doing) empirical research, “normative considerations appear to be decisive in determining the admissibility of facts and in providing a lens through which these facts are filtered.”

Each reviewer is correct to note that the others offer selective reviews of the evidence, adjust the standards of evidence according to the substance of the findings, and find guidance in their ideologies. Each, however, is wrong to exempt him or herself from such charges. All have strong normative commitments—about the purposes of public education, the promise and failings of markets, the appropriate level of government regulations—that inform their assessments of the modest empirical literature on school choice. To make sense of the findings within this literature, the reviewers invariably draw on these normative commitments. As sociologist Joel Best writes in *Damned Lies and Statistics*, “We may think of statistics as facts, but people make facts meaningful, and analysts’ ideologies shape the meanings they assign to social statistics.”

Scholars’ reassurances that they have their intellectual houses in order and can clearly see the empirical evidence on school choice as it is, and for what it is, amount to very little. Indeed, there is a certain irony that in attempting to distinguish truth from fiction and to distill the values and speculations and ideologies that apparently muddy the public debate on school vouchers, these reviews, when read collectively, only make matters worse.

Their subjectivity, of course, does not disqualify any of these scholars from continuing to do primary research on education. Should scientific progress depend on scientists’ hearts being pure, progress most assuredly would be illusory. As long as scholars’ findings are subject to the scrutiny of others—what the philosopher Karl Popper identifies as the requirements of “intersubjective testability”—scientific inquiry can proceed amid deeply held ideological commitments.

Advancing Knowledge

There is nothing wrong with hearing scholars’ views on the state of the empirical literature on school choice. I myself found them helpful in trying to organize my thinking about the existing body of research. Knowing their views, however, should not be confused with knowing more about the likely consequences of various kinds of school voucher programs. If we want to advance knowledge, researchers need to confront two challenges.

First, we simply need more data. Until the early 1990s, the only available evidence on school choice came from programs that limited alternatives to schools within the public sector. In the mid-1990s, a highly contentious debate erupted over whether a couple of hundred voucher students in several private schools in Milwaukee performed higher on standardized tests than their peers in local public schools. Without any solid data to anchor it, the public debate on vouchers floated adrift.

To be sure, the state of research has improved. New evidence has been collected from publicly funded voucher programs in Florida and Cleveland; new findings from randomized field trials conducted in New York City, Dayton, Charlotte, and Washington, D.C., are now available; and a growing body of evidence collected abroad provides a comparative perspective on the probable effects of large-scale choice initiatives.

Still, we lack the data required to

address a host of critical questions about school choice. RAND correctly notes that “the list of unknowns remains substantially longer than the list of knowns.” We know very little about the long-term consequences of vouchers on student achievement. Little or no data have been collected on the effect of school choice on graduation rates, incarceration rates, the probability that students will end up on welfare, the chances that they will be employed full time—all outcomes that deserve careful scrutiny. For the simple reason that all of the existing domestic choice initiatives are quite small, we still know very little about the systemic effects of vouchers on existing public and private schools.

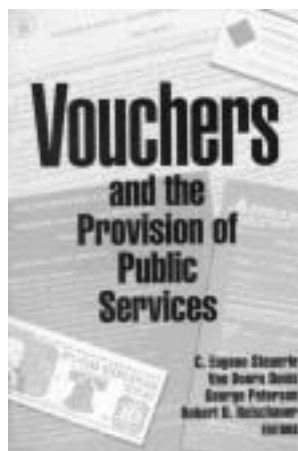
The only way to answer these questions is to begin collecting data from larger programs and investing more money in evaluations so that students can be tracked for longer periods of time. The need for richer data constitutes a primary reason why fellow researchers and I support the creation of a citywide program that offers vouchers set at the amount of per-pupil funding in area public schools. Until quality data are collected from larger voucher programs, ideology will continue to serve as a cheap substitute for hard evidence on the likely effects of school vouchers.

Some will argue that supporting additional research in itself represents a form of advocacy and cannot be justified on purely objective grounds. At one level, this undoubtedly is true. Calling for more research ultimately rests on an intuition that vouchers may achieve important public objectives and that studying them will reap valuable information about the comparative strengths of public and private education systems. Given that no study has demonstrated that targeted urban voucher programs hurt students, and several studies have shown that they are especially beneficial to low-income African Americans, I am perfectly comfortable making the call. It is worth noting, too, that insisting on an end to research also is a form of advocacy. There

is nothing neutral about arguments that we should suspend efforts to study the effects of school choice initiatives.

The second challenge researchers face concerns the statistical techniques and measurement tools currently available to evaluate educational policy. Consider, for example, the problem of self-selection bias that plagued the literature of school sector effects. While impressive data sets were assembled on the achievement levels of thousands of students in public and private schools, statisticians could not be sure whether observed differences reflected the quality of the schools or the students who self-selected into them. It was not until the 1990s that education researchers began to employ a research design that effectively mitigated the problem of self-selection. By using a lottery to assign vouchers, researchers assembled treatment and control groups that, at baseline, were indistinguishable from one another. Subsequent differences observed between the two groups could then be attributed to the intervention, without the need for estimating complex statistical models. The introduction of randomized field trials to education research is as much a boon to knowledge as the results from any single study on school choice.

Unfortunately, plenty of statistical problems besides selection biases continue to wreak havoc on education research. A series of excellent papers by economists Thomas Kane, Douglas Staiger, and Dale Ballou (see “Randomly Accountable,” *Education Next*, Spring 2002, and “Sizing Up Value-Added Assessment,” this issue) scrutinize the error built into value-added test-score measures, many of which are used in state accountability systems. The imprecision of statistical models that estimate year-to-year changes in student test scores to evaluate the quality of individual schools and teachers is sufficiently large that accountability systems frequently sanction success and reward failure. When virtually all education inter-



ventions yield rather modest test-score changes from year to year, it becomes extremely difficult to detect effects given the amount of statistical noise in our instruments. The problem of measurement error is quite real. Until we find ways of mitigating its influence, our ability to judge the effectiveness of education policies reliably is limited.

If scholars are genuinely interested in making advances, they ought to worry less about assembling panels of experts with an appropriate balance of critics and advocates and focus instead on research design issues and, whenever possible, support the collection of more and better data.

Conclusion

The evidence collected from a long-term, well-financed voucher initiative may not change many minds, least of all those already committed to one side of the debate. The findings, however, will help set boundaries on what constitutes a reasonable argument either for or against choice. The forecasts of widespread innovation brought on by competition and of cultural balkanization by privatization will have to accommodate an empirical reality that is considerably more nuanced than vouchers' most ardent advocates and critics are currently willing to admit.

Our reviewers generally agree about

one issue (the greater levels of satisfaction among voucher recipients) not because they agree ideologically or because their assessments of the research on this topic are principled and objective. Instead, they agree because they have no other choice. The existing evidence on the topic overwhelms any possible normative objection. Scholars continue to disagree about other topics (such as the likely impact of a large-scale voucher program on the education of students who remain in public schools) because the evidence assembled to date is more provisional.

The proper antidote to ideology is not an insistence of objectivity, but more and better social-science research. Pious sentiments do not make for objectivity; careful research does. In medical research, many scientists believe in the product they are studying—why else would they work on it? Yet the findings from randomized field trials discipline their thinking and restrict the conclusions they can credibly make. The once dominant views that nicotine is not addictive, or that sodium intake is universally detrimental to one's health, or that cholesterol buildup is the primary cause of heart disease have all been rebuffed, not for lack of powerful interests working on their behalf, but because a mountain of empirical evidence discredits them.

None of the current participants in the school voucher debate will cast the decisive vote on the meaning of the existing research. The next generation of researchers will, and currently they are enrolled in the private and public elementary schools being studied. The best way to help them is to develop new methods for evaluating education outcomes and to assemble high-quality evidence on the long-term programmatic effects of different programs, serving different populations, in different geographic settings.

—William G. Howell is an assistant professor of political science at the University of Wisconsin–Madison.