

Book Review
The Education Gap. Vouchers and Urban Public Schools
by William G. Howell and Paul Peterson

This book is an important contribution to our understanding of how education vouchers might work. *The Education Gap* (Washington DC: Brookings Institute) addresses an important social concern – inequities and inequalities within the US education system – with novel research evidence and a wide-ranging, smooth-toned discussion. The major contribution of the book is to present new evidence from a high-quality research effort to determine the effectiveness of education vouchers.

The book begins with a deliberation over ‘School Choice and American Democracy’, and sets out a series of key questions about school choice. Local funding, exclusive school territories, and a preference for low commuting times, all lead to school choice which is closely bound in with residential choice. The dual choice of school and residence is most restrictive on the options available to low income families. Mobile or high-income families already have school choice. And, Howell and Peterson argue, when racial discrimination is factored in, it is specifically African Americans who have the fewest school choices. They are held to benefit the most from greater choice.

Chapter 2, ‘Evaluating Voucher Programs’, begins with a review of the direct programs in the US. The main focus is on three randomized field trials of vouchers: in New York City, in Dayton, and in Washington, DC. In each city, around 500–1200 students (from low-income families) were offered a voucher of up to \$1,700 to attend a private school of their choice. Vouchers were allocated at random to the eligible applicants, and so the likelihood of confounding causes and effects is much reduced. The authors also include evidence on the nationwide voucher program of the Children’s Scholarship Fund (CSF) (again, with random assignment to students); and the Edgewood School District Program. Collected together, there is a wealth of high-quality data and novel findings on the impact of education vouchers. These findings – too numerous to cover properly in this review – are set out in Chapters 3 through 7.

Chapter 3 looks at students’ selections of schools, and schools’ selections of students. Overall, the authors find little evidence of the types of skimming, sorting, and screening evident in other studies on school choice. But, many individual findings are of interest, such as the take-up rate for private schooling. Given a voucher for \$1,200 (roughly a 50 percent discount on religious private schooling), around one-fifth to one-third of the families chose to remain in public schooling (with the CSF, the take-up rate is even lower). Is this take-up rate high or low? It’s low in that these families were already sufficiently motivated to apply to the program and so presumably were seeking options outside public schooling. But it’s high in that an exodus of this proportion of public school students would increase the size of the private sector by a factor of six.

Chapters 4 and 5 report on what types of schools the voucher students chose, and what the social consequences of these choices were. Both chapters are somewhat flat. The chosen schools do not appear especially distinctive in terms of class or school size, facilities, homework-setting, or school climate. The social consequences are similarly muted: across the experimental and control groups, parental involvement is no higher, nor is student self-confidence or behavior.

Perhaps of most interest, though, are the results in Chapter 6, ‘The Urban Test Score Gap’. Notwithstanding the caveats of using test scores (which the authors note), such scores are held as the ‘education gap’ this book wants to close. The headline results – given in Tables 6-1 and 6-2 – show “no overall private school impact of switching to a private school on student test scores in the three cities” (p. 145). But, the effect size gains for African American students were

0.18–0.30. These gains, worthwhile in themselves, would also serve to reduce the black-white test score gap.

Chapter 7 focuses on parental satisfaction with schools. A consistent finding in favor of most choice programs is that parents who get more choices are more satisfied. Freedom of choice – emphasized by voucher advocates – is an important concern, yet these correlations cannot be pushed too far. Higher satisfaction may be a result of winning the random assignment. Unavoidably, field trials cannot assign children blindly into the voucher program, so more satisfaction may plausibly be attributed to the receipt of \$1,400 to spend on private schooling. Without blind assignment, it is difficult to ascertain – from self-reports of participants in the trial – what effect more choice has on parental satisfaction. Even where satisfaction is higher, the magnitude is unknown. Greater satisfaction is hard to price under many circumstances, but there are three specific concerns here. First, satisfaction levels across the experimental and the control groups appear to converge over the three years, suggesting a temporary fillip from the voucher. Second, Howell and Peterson report satisfaction levels by school type (public/private), not by voucher receipt, and this tends to overstate the satisfaction premium from a voucher program. Third, parents may have contributed additional own funds toward private schooling fees; higher satisfaction with schooling may therefore mean lower satisfaction from other purchases. Nevertheless, if the value of satisfaction can be estimated, and it proves non-trivial, then the extra freedom of choice may be an important advantage.

The final Chapter, ‘Vouchers and Urban Schools’, offers both a summary and a projection for voucher programs. Here, the authors address broad issues, co-mingling a discussion of external validity with proposals for voucher reforms. The agenda is bold in some respects, narrow in others. So, the authors discuss the constitutionality of voucher programs, arguing that, even though “vouchers alter students’ religious practices”, “for those members of the [Supreme] Court who believe that vouchers, in practice, must be *primarily educational*, the evidence now available may be dispositive” (p.191). Next, the authors acknowledge the limited impact that current choice programs (including charter schools) have to “enrich the existing educational market” (p.194); the programs are too small, too short, and too dogged by political revisions. Moreover, only some parents – those in the inner-city – need extra choices. Hence, a new agenda emerges: full-value vouchers (equal to public school per-pupil expenditures) offered to all inner-city parents. In the short-term, Howell and Peterson predict “a large-scale, adequately financed urban voucher program may initially attract between 10 and 15 percent of public school students” (p.200). Targeted, not wholesale, voucher reform is the new approach.

It is debatable how many voucher skeptics this approach will win over. Targeted voucher reform may be the making of a new consensus on the utility of vouchers: many of the criticisms of voucher programs as inequitable and socially divisive apply to large-scale voucher schemes, applied across all income groups. However, targeted programs are unlikely to make any dint in the ‘public school monopoly’ that has been the voucher advocate’s *bête noire*. And, such proposals are unlikely to be welcomed by fiscal conservatives, or home-schoolers, who seek alternative privatization reforms. These issues are likely to influence the politics of vouchers in the future.

In summary, *The Education Gap* makes for intelligent debate, with a wealth of persuasive evidence. To conclude, though, three interlinked concerns are noted; these do not affect the main research conclusions, but are important nonetheless.

The first concern relates to the description of the experiment. Fundamentally, RFTs analyze ‘interventions’, i.e. discrete, fixed, or controllable acts that the experimental group is treated with. If the RFT is performed correctly, and the experimental group outcomes differ from the control group outcomes, then a highly plausible inference is that the intervention *caused* the difference. Therefore, it is imperative to specify – in as much detail as possible – what the intervention was; and this imperative is all the more pressing when the intervention is as complicated and as amorphous as a voucher. Yet, the authors do this only partially. A \$1,000

voucher in New York is very different from a \$1,000 voucher in Dayton. A time-limited voucher is different from an open-ended voucher. A student-eligible voucher is different from a family-eligible voucher. And these differences gain further salience if the experimental group interprets the voucher in various ways. The stipulations on the voucher need to be specified precisely; these stipulations compose the *cause* that should bring about the effect. In fact, the terms to describe a voucher are well-known, having been set out several years ago (Levin, 1992). To aid the readers' interpretation and assessment of external validity, explicit and transparent statement of the voucher stipulations are needed.

The second concern relates to the description of the research method. There is little argument that RFTs – in general – have strong internal validity; but only when they are undertaken properly. Of course, no research protocol that is poorly applied will have high internal validity, but there is a reasonable case that RFTs are relatively difficult to implement in practice (relative to other methods, but also relative to randomized clinical trials). If so, different judgment criteria should be used to evaluate RFTs, to include more explicit consideration of practical implementation problems. This is not a case of holding RFTs to a higher standard, but of recognizing that ease of data collection varies across methods. The assertion that experimental research is a “gold standard” has been considered by Cook and Payne (2002, 174), who assert, “It is clearly not that [gold standard] in educational contexts, given the difficulties with implementing and maintaining randomly created groups, with the sometimes incomplete implementation of treatment particulars, with the borrowing of some treatment particulars by control group units, and with the limitations to external validity that often follow from how random assignment is achieved.” These important cavils need to be addressed; it cannot be simply assumed that research findings from RFTs are compelling. Moreover, the findings from random assignment do not automatically trump findings from other methods (as the authors implicitly acknowledge by including evidence from the nationwide voucher program and that in San Antonio). A hypothesis that has been repeatedly tested may be as plausible as one that has been well tested once. As a final aside on this issue, the authors make much of the effect sizes for African American students, even though students were not randomized by race; and the effect sizes in Washington (where almost all the students in the experiment were African American) were not consistently significant.

The final concern relates to the authors' assumptions that private schools are more efficient than public schools. The effectiveness of private schools has been investigated in depth, but their efficiency has not. There exists no reliable, robust data on the actual expenditures of private schools. Fees are not a good indicator, because these are only one source of revenue for private schools. It is therefore misleading to equate effectiveness with efficiency, even though it is the latter that it is critical in evaluating the achievements of private schools compared to public schools.

Clive R. Belfield

cb2001@columbia.edu

National Center for the Study of Privatization in Education, Teachers College, www.ncspe.org.