Title: Distributional Effects of a School Voucher Program: Evidence from New York City

Authors and Affiliations:
Marianne Bitler*, Department of Economics, University of California, Irvine & NBER
Thurston Domina, Department of Education, University of California, Irvine
Emily K. Penner Department of Education, University of California, Irvine

*Corresponding author. Please direct correspondence to: mbitler@uci.edu

Background/context

Excellence and equity goals motivate much of American educational policy. These two goals are not always mutually reinforcing. Some policies boost average academic achievement even as they broaden educational inequalities. Others depress academic achievement even as they narrow inequalities. The twin goals of excellence and equity should lead policy-makers to be interested in both the average effects of educational policies and their distributional consequences. But although developmental science suggests that many interventions should have heterogeneous effects, most educational evaluation research focuses on the estimation of mean treatment effects either for the population at large or for particular subgroups of interest. In this paper, by contrast, we use distributional approaches to estimate the effects of a school voucher experiment for low income elementary school students in New York City on the distribution of academic skill. The distributional approaches will test competing hypotheses that school vouchers boost achievement either primarily at the bottom of the distribution (thus mitigating inequality) or primarily at the top of the distribution (thus exacerbating inequality).

Research Question

Arguing that traditional public schools are monopolistic and inefficient, school voucher proponents aim to create more vibrant educational marketplaces. By broadening the educational choices available to parents and students and creating incentives for schools to improve, voucher programs promise to boost educational outcomes for students who might otherwise have no choice but to enroll in low-quality public schools (Chubb & Moe 1990; Friedman & Friedman 1980).

School reformers have launched a handful of voucher programs across the US over the past two decades in an attempt to demonstrate the effectiveness of this approach. In 1997, the School Choice Scholarships Foundation initiated one such program in New York City, offering three-year scholarships worth $1,400 a year to a randomly selected group of low income children in grades K–4. This program's random assignment design makes it possible to distinguish the effects of voucher availability from the potentially confounding characteristics of families who self-select into voucher programs. Mathematica Policy Research (MPR) and the Harvard University Program on Education Policy collected enrollment and achievement data from students in the treatment and control groups.

Analyses of the New York City voucher experimental data clearly indicate that vouchers influence school choice. Students randomly selected to receive a voucher were several times more likely than their peers in the control group to attend private schools (Mayer et al. 2002). It is much less clear, however, whether the voucher offer had an effect on student achievement. Several studies find that the voucher offer had a small positive effect on the academic achievement of African-American recipients (Barnard et al. 2003; Howell et al. 2002; Peterson & Howell 2004). However, for all other racial and ethnic groups the voucher program had no effect. Furthermore, subsequent analyses suggest that the observed effects for African Americans are sensitive to the definition of racial and ethnic categories and hold only when controlling for students' initial characteristics (Krueger & Zhu 2004a, 2004b). Possibly the weak average effects of

vouchers disguise larger (and possibly contradictory) voucher program effects for high or low achieving students. The theoretical literature surrounding school choice suggests two competing hypotheses regarding the effects of voucher programs on student achievement.

The "common school hypothesis" holds that voucher programs benefit low-performing students disproportionately. This hypothesis is grounded in the literature on the effects of Catholic schools. Since Catholic schools are typically smaller than public schools and their curricula are often relatively undifferentiated, low-performing students tend to benefit disproportionately from enrolling in Catholic schools (e.g., Evans & Schwab 1995; Morgan 2001). By providing a mechanism for students to opt out of neighborhood public schools and into Catholic and other private schools, voucher experiments attempt to make the positive achievement effects associated with Catholic schools more broadly available. Assuming that the Catholic school effects from the literature generalize to the schools that voucher recipients chose, the "common school hypothesis" suggests that voucher school programs have positive effects on students at the bottom end of the academic achievement distribution.

However, the "stratifying hypothesis" suggests that voucher programs magnify educational inequalities. Voucher program advocates take it for granted that parents use school choice to maximize their children's educational success. In practice, however, many parents make school choice decisions based on the convenience of the school's location, its disciplinary style, and its religious affiliation (e.g., Elacqua, Schneider, & Buckley 2006). Hastings, Kane, & Staiger (2005) hypothesize the effects of voucher programs are contingent on the quality of the school choices that families make. For students whose families make school choices on the basis of academic quality, voucher programs may have positive effects. But for students whose families make school choices based on other factors, vouchers may have zero or negative effects. If these educational preferences vary with student academic achievement, voucher programs may provide a boost for students who are high achievers pre-voucher, even as they negatively affect test scores for low achievers.

These competing hypotheses have not been thoroughly investigated in an explicitly distributional fashion. Using QTE methods to analyze this experimental data will add to our understanding of what parts of the distribution benefit from access to vouchers. We can test systematically whether the variation in effects for various subgroups found in previous literature captures important variation across the distribution of test scores.

Setting/Population/Participants/Subjects

The New York City School Choice Scholarship Program (NYCSCSP) was a three year private school choice randomized experiment. As noted above, low income students (students qualified for free school lunch) in grades K–4 were eligible to apply for vouchers of $1,400 to be used towards private school tuition. While not representative of all students, these students are most similar to those in districts facing challenges to improve. We will use NYCSCSP data collected by Mathematica Policy Research, obtained after completing an application process, to evaluate distributional effects of the voucher experiment in New York.

Program/Intervention

The NYCSCSP program was a randomized control trial of the effects of an offer of a scholarship of $1,400 on children's outcomes. After applying, low income students in grades K-4 were assigned to a treatment group – who could obtain a 3-year scholarship, or to a control group, who received no such offer. Around 2,600 students out of 10,000 eligible applicants were selected. Except for the kindergarten students, students were administered a pre-program test and post-random assignment achievement tests. Parents were also surveyed.

Research Design

The potential outcomes model provides a framework for estimation of the effects of a treatment. Each individual i has two potential outcomes, Y_{1i} and Y_{0i} (for our purposes, a test score or index of student behavior or motivation). Person i has outcome Y_{1i} if assigned to the treatment group and outcome Y_{0i} if assigned to the control group. D(i) denotes the group that i is assigned to in a randomized experiment. If person i is assigned to the treatment group, then D(i) = 1, and if person i is assigned to the control group, D(i) = 0; the treatment effect on person i is defined as $d_i=Y_{1i} – Y_{0i}$.

Quantiles, Average Treatment Effects, and Quantile Treatment Effects

Let Y be a random variable with a cumulative distribution function (CDF) F(y), where $F(y) = Pr[Y \leq y]$. Then, the qth quantile of the distribution F(y) is defined as the smallest value y_q such that $F(y_q)$ is at least as large as q (e.g., $y_{0.5}$ is the median). Now consider two (marginal) distributions F_1 (the CDF for the potential outcomes if D = 1), and F_0 (the CDF for the potential outcomes if D = 0). We define the difference between the qth quantiles of these two distributions as $y_q = y_{q1} – y_{q0}$, where y_{qd} is the qth quantile of distribution F_d.

The joint distribution of (Y_{0i},Y_{1i}) is not identified without assumptions. However, if program assignment is independent of the potential outcomes, the difference in means, or average treatment effect, $d = E[d_i] = E[Y_1] - E[Y_0]$, is identified because each expectation requires only one of the two marginal distributions. Similarly, identification of the marginal distributions implies identification of the quantiles y_{qd}, and thus identification of the differences in their quantiles, $y_q = y_{q1} – y_{q0}$. In fact, in the experimental settings of Specific Aims 1 and 2, the quantile treatment effect (QTE) is the estimator of this difference in the quantiles of the two marginal distributions. For example, in these experimental studies, we consistently estimate the QTE at the 0.50 quantile by subtracting the control group's sample median from the treatment group's sample median. Graphically, QTE estimates are the differences in the inverse CDFs of the outcome for the treatment and control groups. For an example, see our preliminary results reported in Appendix B. of the effects of being assigned to get an offer of a private school voucher (Figure 1).

Testing whether Mean Effects in Subgroups Explain QTEs on the Full Distributions

A common strategy in program evaluation for examining heterogeneous effects is to compute mean treatment effects separately for specific subgroups. We will use the methods from Bitler, Gelbach, & Hoynes (2011) to test whether mean effects within such subgroups (or others that theory suggests are good predictors of being someone who gains or loses from the interventions) can explain the heterogeneity uncovered by QTE. We first will construct QTE within subgroups of interest. We then examine whether

members of these subgroups are concentrated in various points of the outcome distribution for the treatment and control groups. For example, this can be done by seeing whether the SES composition within each percentile of the treatment or control distribution varies across the distribution.

Data Collection and Analysis

After checking the treatment and control groups are balanced, we will estimate QTE for the Iowa Test of Basic Skills Math and Reading composite scores and for two school characteristics: class size and an index of parental satisfaction with the child's school. We will explore the extent to which the overall heterogeneity can be replicated by subgroup means, paying attention to the characteristics suggested by the evaluators to define groups who had different gains from assignment to the treatment group as well as suggestions of theory.

Findings/Results

Figure B1 in Appendix B shows our preliminary findings for the effects of voucher receipt on student test scores measured one year after the voucher lottery occurred. The NYCSCSP had no mean effect on student math achievement (Howell et al. 2002; Krueger & Zhu 2004a, 2004b). Our preliminary findings suggest that this non-finding may not hold across the distribution. By contrast, vouchers seem like they may have had a positive effect on low-performing students and a negative effect on high performers.

Conclusions

Our preliminary findings suggest that vouchers may have different effects at different points in the distribution of student achievement. Further work will reveal how systematic this finding is, and explore the extent to which effects within subgroups are similar to those for the whole distribution.

Appendices

Appendix A: References

Barnard, J., Frangakis, C.E., Hill, J.L., & Rubin, D.B., (2003), 'Principal stratification approach to broken randomized experiments: A case study of school choice vouchers in New York City,' Journal of the American Statistical Association 98:462, 299–311.

Bitler, M., Gelbach, J., and Hoynes, H. (2006). 'What mean impacts miss: Distributional effects of welfare reform experiments,' American Economic Review 96:4, 988–1012.

Bitler, M.P., Gelbach, J.B., & Hoynes, H.W. (2011), 'Can variation in subgroups' average treatment effects explain treatment effect heterogeneity? Evidence from a social experiment,' unpublished.

Chubb, J.E. & Moe, T.M. (1990). Politics, Markets, and America's Schools. Brookings Institute Press: Washington, D.C.

Elacqua, G., Schneider, M. & Buckley, J. (2006.) 'School Choice in Chile: Is It Class or the Classroom?" Journal of Policy Analysis and Management 25:3, 577–601.

Evans, W. & Schwab, R. (1995) 'Finishing high school and starting college: Do Catholic schools make a difference?' Quarterly Journal of Economics 110: 941–974.

Friedman, M. & Friedman, R. (1980). Free to Choose. New York: Harcourt Brace Jovanovich.

Hastings, J., Kane, T., & Staiger, D. (2005), 'Parental preferences and school competition: evidence from a public school choice program.' National Bureau of Economic Research Working Paper No.11805.

Howell, W.G., Wolf, P.J., Campbell, D.E., & Peterson, P.E., (2002), 'School vouchers and academic performance: Results from three randomized field trials,' Journal of Policy Analysis and Management 21:2, 191–217.

Krueger, A. & Zhu, P. (2004a), 'Another look at the New York City school voucher experiment,' American Behavioral Scientist 47:5, 658–698.

Krueger, A. & Zhu, P. (2004b), 'Inefficiency, subsample selection bias, and nonrobustness,' American Behavioral Scientist 47:5, 718–728.

Mayer, D.P., Peterson P.E., Myers, D.E., Tuttle, C.C., & Howell, W.G. (2002), 'School choice in New York City after three years: An evaluation of the school choice scholarship program.' Mathematica Policy Research, Inc.: Washington, D.C.

Morgan, S.L., (2001), 'Counterfactuals, causal effect hetereogeneity, and the Catholic school effect on learning,' Sociology of Education 74, 341–374.

Peterson, P.E., & Howell, W.G., (2004), 'Efficiency, bias, and classification schemes: A response to Alan B. Krueger and Pei Zhu,' American Behavioral Scientist 47:5, 699–717.

Appendix B: Figure

Figure B1: QTE estimates of NYCSCSP effects on first-year math test scores