Title: The Participant Effects of Private School Vouchers across the Globe: A Meta-Analytic and Systematic Review

Authors and Affiliations:
M. Danish Shakeel (mdshakee@uark.edu), Dept. of Education Reform, University of Arkansas
Kaitlin P. Anderson (kaitlina@uark.edu), Dept. of Education Reform, University of Arkansas
Patrick J. Wolf (pwolf@uark.edu), Dept. of Education Reform, University of Arkansas

Abstract Body

Background / Context:

School choice has emerged as a key demand side intervention in school reform globally. School vouchers act as a market based reform by allowing parents to choose any school for their children. Both government and privately sponsored voucher programs exist. The effectiveness of voucher programs is fiercely disputed in both academic and policy circles. Most reviews of the school voucher literature have been selective, not systematic. Prior research by Rouse & Barrow (2009); Anderson, Guzman, & Ringquist (2013); and Epple, Romano & Urquiola (2015), either did not systematically search for all the empirical evaluations of school voucher participant effects or relied heavily on non-experimental findings even when a large number of more rigorous studies were available. A thorough meta-analysis informed by a true systematic review of all the available randomized controlled trial (RCT or "experimental") studies would provide the foundation for a greater scholarly consensus regarding the ability of school vouchers to improve outcomes for students.

Purpose / Objective / Research Question / Focus of Study:

The objective of this meta-analysis is to rigorously assess the participant effects of private school vouchers, or in other words, to estimate the average academic impacts that the offer (or use) of a voucher has on a student. This review will add to the literature by being the first to systematically review all Randomized Control Trials (RCTs) in an international context. Our meta-analytic results will focus on the RCTs because these are the "gold standard" of program evaluation in terms of assessing causal relationships. RCTs essentially compare a treatment group (those receiving the offer of a voucher) relative to a control group (those who did not receive the offer of a voucher). In RCTs the assignment of a voucher is random, and therefore the issue of selection bias is resolved in expectation.

The majority of RCTs studying the participant effects of school vouchers have been conducted in the United States. While voucher systems exist in many parts of the world, only a small number of voucher RCTs have been conducted outside the US. Therefore, we will present three meta-analytic estimates of the impacts of school vouchers: (1) just in the U.S.; (2) just outside the U.S.; and (3) globally including the U.S. and all other countries.

Our initial search was guided by the following research question: *What is the impact of private school vouchers globally on the student achievement of those students offered the vouchers?*

Our focus throughout this study will be to see what impact, if any, school voucher programs, in the United States and throughout the world, have had on student test scores. If the findings are mixed, we shall try to determine unique patterns that are driven either by geography or relevant program design components. We will also compare overall outcomes for reading and math scores for programs within the US vs. outside the US and publically funded vs. privately funded programs. This can be helpful for policymakers designing future private school voucher programs. Reading assessments will only be included if they were in English, regardless of the language of the country in which they were administered. We do this to ensure commonality in the international reading assessments and also because the international voucher evaluations in

the meta-analysis come from developing countries where English language skills are highly valued.

Setting:

The RCTs included in our analysis were located in four countries: the United States of America, Kenya, Colombia and India. Although this study will be representative of four continents: North America, South America, Africa, and Asia, the majority of RTCs were administered within the United States. The U.S. studies covered programs in Charlotte, NC; Dayton, OH; Milwaukee, WI; New York City; Toledo, OH; and Washington, DC.

Population / Participants / Subjects:

The participants in the RCTs were children who attended private schools through a school voucher. The grades analyzed ranged from K to 12, although most individual RCTs included a shorter grade range in their analysis. The sample sizes for treatment and control groups as well as the overall sample sizes will be reported in our study and informed our meta-analysis calculations.

Intervention / Program / Practice:

The programs evaluated were publically or privately funded school voucher or K-12 "scholarship" programs. Voucher programs provide tuition scholarships to eligible students that enable them to attend their choice of any participating private school. Most of the private schools that participate in voucher programs in the U.S. and other countries are relatively low-cost schools with per-student costs below the average amount spent in area public schools. The duration of studies analyzed ranged from one year to six years. The earliest program evaluated was administered in 1990 in Milwaukee, WI, and the latest program evaluated was administered in 2011 in Delhi, India.

Research Design:

The research design of the studies that inform the meta-analysis was random assignment of children to treatment and control groups. Most studies had a one-stage randomization through administration of a lottery while one study in Andhra Pradesh, India (Muralidharan & Sundararaman, 2015) was based on a two-stage randomization (randomly assign students within randomly assigned villages). We combine the results of the experimental studies systematically, using the impact estimates and variances reported in the actual studies, to generate overall measures of average voucher impact (Cohen's *d*) along with 95% confidence intervals around those estimates.

Data Collection and Analysis:

For this meta-analysis, we identified publications from computer and networked searches through a variety of sources. The first search was through EBSCO, JSTOR and ProQuest databases available at the library at the University of Arkansas (2,737 articles). The second search utilized the Google Scholar search engine (6,570 articles). Additionally, we searched various internet sources including but not limited to the websites of the National Bureau of Economic Research (NBER), University of Chile, Uppsala University in Sweden, and Poverty Action Lab at MIT. No new RCT studies were found through this search. Last, we conducted a network search based on matching our list of potential sources with earlier publications on

school vouchers internationally and review by Dr. Patrick J. Wolf, an author of this study (no additional articles found). From our two primary searches (library and Google Scholar), we found 9,307 articles in total.

Following the search stage, the team members excluded duplicates studies (543), studies whose titles and/or abstracts were not relevant to school vouchers or were reporting results not relevant to participant effects (8,488). Then, the team members located the full articles of all the remaining studies in the list and read them one-by-one. At this stage, 255 studies not relevant to our analysis or having non-experimental research designs were excluded. See Appendix B for details on the studies eliminated at each stage.

The remaining twenty-one articles were coded in MS Excel for details on author, publication year, location, funding type (public/private), years of evaluation, grades analyzed, outcome (reading(English)/math), size of treatment and control group and overall sample size. Finally, some studies had multiple evaluation years for the same program. We keep only the studies reporting the last year for which the program evaluation results were available. The entire search process was performed separately by at least two team members so they could match their results and minimize human error. An additional six studies were excluded at the coding stage due to repeat coverage or insufficient information.

As some studies did not report their findings in detail, we made necessary assumptions to derive accurate sample sizes for treatment and control groups. For the meta-analysis, we weighed each study by the inverse variance and calculated the pooled standard deviation and effect size using Cohen's d. We also calculated the unbiased d and the standard error for the effect size. When required, the effect sizes were also calculated by correlation and t-ratio. Lastly, the grand effect size and lower and upper bound of the overall 95% confidence interval were also calculated.

Findings / Results:
We report effect sizes for fifteen studies for math scores and fourteen studies for reading scores. For math scores, the effect sizes are positive for twelve studies except for Rouse (1998), Bettinger (2003), and Muralidharan and Sundararaman (2015). Although the math effects are positive for most studies, we fail to reject the null hypothesis for all but two studies: Howell, Wolf, Campbell, and Peterson's DC results (2002) and Greene, Peterson, and Du (1999). For reading studies, the effect sizes are positive for thirteen studies except for Krueger and Zhu (2004). Although the effects are positive for most studies, we fail to reject the null effect for all but four studies: Cowen (2008); Muralidharan and Sundararaman (2015); Wolf (2015); Howell, Wolf, Campbell, and Peterson's DC results (2002). Overall global effect size from meta-analysis indicates null impacts on math scores [95% CI: -0.003 to 0.057 standard deviations] and positive, but small impacts on reading scores [95% CI: 0.066 to 0.127 standard deviations].

For math outcome measures, the effect size is null overall. Studies in the US ($d = 0.031$) [95% CI: -0.008 to 0.069 standard deviations] tended to be slightly more positive than studies outside the US ($d = 0.021$) [95% CI: -0.027 to 0.069 standard deviations]. Moreover, there is a significant positive effect size for privately funded vouchers ($d = 0.037$) [95% CI: 0.001 to 0.074 standard deviations] and a null effect of government funded vouchers ($d = 0.005$) [95% CI: -

0.050 to 0.059 standard deviations]. We fail to reject the null math effects for publically funded vouchers as well as for vouchers in US and non-US regions considered separately. Privately funded programs seem to produce large gains in math scores while government funded programs have an overall null effect. However, this does not imply causality that the private funding is causing the differential effect.

For reading outcome measures, the effect size is positive overall for studies in and outside of the US. The effect size is much greater for non-US regions ($d = 0.136$) [95% CI: 0.087 to 0.185 standard deviations] than for US ($d = 0.071$) [95% CI: 0.032 to 0.111 standard deviations]. Additionally, the effect size for privately funded vouchers ($d = 0.102$) [95% CI: 0.064 to 0.139 standard deviations] is more than that for publically funded vouchers (0.087) [95% CI: 0.033 to 0.142 standard deviations]. Thus, for reading scores, private- as well as government-funded programs produced positive effects, but again privately funded programs seem to produce larger effects than publicly funded programs. However, we cannot be confident that the funding structures are causing the differential. Last, for reading scores, none of the confidence intervals for effect sizes (public funding, private funding, within US, non-US) contains zero, indicating that across all these different comparisons, we find significantly positive impacts of school vouchers on reading scores globally.

Conclusions:
This meta-analysis contributes to the field by combining and systematically evaluating rigorous evidence from RCT studies. This review provides a broader overview of all the rigorous experimental findings and will have important policy implications about the effectiveness of voucher programs generally. While voucher programs are growing across the globe, a meta-analysis of the participant effect of vouchers internationally was lacking. As the first meta-analysis of its type, it will help establish the baseline for future studies.

We should emphasize that for evaluations of the same program done over multiple years, we chose to analyze the results for the latest year available. Voucher programs appear to be having overall positive effects in reading, but more RCTs outside of the US are needed to reflect the presence of these programs around the world.

In terms of recommending policy, there are a couple different conclusions we can draw from these results. We found that in general, privately funded programs show more positive effects, but this could be the result of several different things. For example, it could be that private donors may have better planning, implementation, and oversight than government forms of funding. In addition, it could be that privately funded programs are more likely to have financial support for RCT studies when they are thought to be succeeding, and that these types of studies are more prevalent in the literature.

Appendix A. References

Anderson, M. R., Guzman, T., & Ringquist, E. J. (2013). Evaluating the effectiveness of educational vouchers. In E. J. Ringquist (Ed.), Meta-analysis for public management and policy (pp. 311-351), San Francisco, CA: Jossey-Bass.

Angrist, J., Bettinger, E., Bloom, E., King, E., & Kremer, M. (2001). *Vouchers for private schooling in Colombia: Evidence from a randomized natural experiment* (No. w8343). National Bureau of Economic Research.

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),* 299-323.

Bettinger, E. (2005, October). Private School Vouchers in Colombia. *In Draft paper prepared for the conference, Mobilizing the Private Sector for Public Education, World Bank and Kennedy School of Government, Harvard University, Cambridge, MA*.

Bettinger, E., & Slonim, R. (2003). *The effect of educational vouchers on academic and non-academic outcomes: Using experimental economic methods to study a randomized natural experiment*. Mimeo, Case Western Reserve University.

Carlson, D. E., Cowen, J. M., & Fleming, D. J. (2014). Third-Party Governance and Performance Measurement: A Case Study of Publicly Funded Private School Vouchers. *Journal of Public Administration Research & Theory, 25*, 897-922.

Centre for Civil Society. (2014). *School vouchers for girls*. Paper presented at the 3rd annual International School Choice and Reform Academic Conference, Fort Lauderdale, FL, January 2014.

Correa, J. A., Parro, F., & Reyes, L. (2014). The Effects of Vouchers on School Results: Evidence from Chile's Targeted Voucher Program. *Journal of Human Capital, 8(4),* 351-398.

Cowen, J. M. (2008). School choice as a latent variable: Estimating the "complier average causal effect" of vouchers in Charlotte. *Policy Studies Journal*, 36(2), 301-315.

Dixon, Paulin. (2013). *International Aid and Private Schools for the Poor: Smiles, Miracles and Markets*, Northampton, MA: Edward Elgar.

Epple, D., Romano, R. E., & Urquiola, M. (2015). *School Vouchers: A Survey of the Economics Literature* (No. w21523). National Bureau of Economic Research.

Forster, G. (2013) A Win-Win Solution: The Empirical Evidence on School Vouchers. Indianapolis: *Friedman Foundation for Educational Choice*.

Green, J. P. (2000). The Effect of School Choice: An Evaluation of the Charlotte Children's Scholarship Fund Program. *Manhattan Institute for Policy Research Civic Report,* (12).

Greene, J. P., Peterson, P. E., & Du, J. (1999). Effectiveness of school choice: The Milwaukee experiment. *Education and Urban Society, 31(2),* 190-213.

Howell, W. G., & Peterson, P. E. (2000). School choice in Dayton, Ohio: An evaluation after one year. *Cambridge, MA: Program on Education Policy and Governance, Harvard University.*

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.

Jin, H., Barnard, J., & Rubin, D. B. (2010). A modified general location model for noncompliance with missing data revisiting the New York City School Choice Scholarship Program using principal stratification. *Journal of Educational and Behavioral Statistics, 35(2),* 154-173.

Lamarche, C. (2008). Private school vouchers and student achievement: A fixed effects quantile regression evaluation. *Labour Economics, 15(4),* 575-590.

Lara, B., Mizala, A., & Repetto, A. (2011). The Effectiveness of Private Voucher Education: Evidence From Structural School Switches. *Educational Evaluation and Policy Analysis, 33(2),* 119-137.

Kremer, M., Miguel, E., & Thornton, R. (2009). Incentives to learn. *The Review of Economics and Statistics, 91(3)*, 437-456.

Krueger, A. B., & Zhu, P. (2004). Another look at the New York City school voucher experiment. *American Behavioral Scientist, 47*(5), 658-698.

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 Scholarships Program. Final Report.

Morgan, C., Petrosino, A., & Fronius, T. (2013). *A Systematic Review of the Evidence of the Impact of School Voucher Programmes in Developing Countries*. EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Muralidharan, K., & Sundararaman. V. (2015). The aggregate effect of school choice evidence from a two-stage experiment in India. *The Quarterly Journal of Economics, 130(1).*

Peterson, P. E., & Howell, H. G. (2003). Latest results from the New York City voucher experiment. *Cambridge, MA: Program on Education Policy and Governance, Harvard University.*

Rouse, C. E. (1998). Private School Vouchers And Student Achievement: An Evaluation Of The Milwaukee Parental Choice Program. *Quarterly Journal of Economics, 113(2),* 553-602.

Rouse, C. E., & Barrow, L. (2009). School vouchers and student achievement: Recent evidence and remaining questions. *Annual Review of Economics, Vol. 1, no. 1,* pp. 17–42.

Tooley, J. (2009). *The Beautiful Tree: a personal journey into how the world's poorest people are educating themselves*, Washington, D. C.: Cato Institute.

Witte, J. F., Carlson, D., Cowen, J. M., Fleming, D. J., & Wolf, P. J. (2012). *MPCP* Longitudinal Educational Growth Study: Fifth Year Report. SCDP Milwaukee Evaluation Report #29. *School Choice Demonstration Project, Department of Education Reform, University of Arkansas.*

Wolf, P. J., Egalite, A. J., & Dixon, P. (2015). Private School Choice in Developing Countries: Experimental Results from Delhi, India. *Handbook of International Development and Education,* 456.

Wolf, P. J., Kisida, B., Gutmann, B., Puma, M., Eissa, N., & Rizzo, L. (2013). School vouchers and student outcomes: Experimental evidence from Washington, DC. *Journal of Policy Analysis and Management, 32(2)*, 246-270.

Wolf, P. J. & McShane, M. (2013). Is the juice worth the squeeze? A benefit/cost analysis of the District of Columbia Opportunity Scholarship Program. *Education Finance and Policy, 8 (1) (2013),* pp. 74–99.

Appendix B. Details on Search and Exclusion Process

	Number of Articles
Search 1 (University of Arkansas Library)	
Three library sources (EBSCO, JSTOR, ProQuest)	2,737
Duplicates Removed	-534
Unique articles (EBSCO, JSTOR, ProQuest)	2,203
Excluded Based on Title and/or Abstract	-2,075
Remaining Articles (EBSCO, JSTOR, ProQuest)	**128**
Search 2 (Google Scholar)	
Number of Google Scholar Sources Initially Found	6,570
Excluded Based on Title and Abstract	-6,413
Remaining Google Articles	157
Duplicates Removed	-9
Remaining Articles (Google Scholar)	**148**
Sum of Remaining Articles (Both Searches)	**276**
Excluded Based on Full Article	-255
Excluded at Coding Stage (due to repeat coverage or insufficient information)	-6
Total search results (RCTs)	**15**

Appendix C. Tables and Figures

Not included in page count.

Table 1: Study Characteristics for Math Outcomes

S. No.	Authors	Pub Year	Program Evaluated	Location	Funding	Duration of Study	Grades	Sample Size	Outcome Measure	Results	Comments
1	Rouse	1998	Milwaukee Parental Choice Program (MPCP)	Milwaukee, WI (USA)	Public	1990-1994 (5 years)	K to 8	N=3177 (T=1589, C=1588)	Math	Positive	
2	Greene, Peterson & Du	1999	Milwaukee Parental Choice Program (MPCP)	Milwaukee, WI (USA)	Public	1990-1994 (5 years)	K to 8	N= 317 (T=237, C = 80)	Math	Positive	3- and 4- year outcomes combined. Calculated treatment and control using Table 1 (p. 199).
3	Greene	2000	Charlotte Children's Scholarship Fund	Charlotte	Private	1999-2000 (1 year)	2 to 8	N=357 (T = 223, C=134)	Math	Null	Used the "Instrumental w/ background controls" as this was the most rigorous measure (Table 3). Used the p-value to calculate a t-statistic in order to find the standard error of the effect size. The sample size was split based off of the total 357, split by the treatment/control ratio in Table 1.
4	Angrist, Bettinger, Bloom, King & Kremer	2002	Programa de Ampliacion de Cobertura de la Educacion Secundaria (PACES)	Colombia	Public (partly funded by World Bank)	1995-1999 (4 years)	6 to 10	N=282 (T=157, T=125)	Math	Null	
5a	Howell, Wolf, Campbell & Peterson	2002	School Choice Scholarships Foundation	NYC	Private	1997-1999 (2 years)	1 to 4	N=1199 (T= 600, C=599)	Math	Null	Year 2 results. Assumed 50/50 treatment and control split. Combined math and reading score, so the effect size was assumed the same for both subjects.
5b	Howell, Wolf, Campbell & Peterson	2002	Parents Advancing Choice in Education	Dayton, OH	Private	1998-2000 (2 years)	K to 12	N=382 (T= 191, C = 191)	Math	Null	Year 2 results. Assumed 50/50 treatment and control split. Combined math and reading score, so the effect size was assumed the same for both subjects.
5c	Howell, Wolf, Campbell & Peterson	2002	Washington Scholarship Fund	Washington, DC	Private	1998-2000 (2 years)	K to 8	N= 725 (T=363, C=362)	Math	Positive	Year 2 results. Assumed 50/50 treatment and control split. Combined math and reading score, so the effect size was assumed the same for both subjects.
6	Bettinger & Slonim	2003	Children's Scholarship Fund	Toledo, OH (USA)	Private	1998-2001 (4 years)	K to 8	N=360 (T=118, C=242)	Math	Null	
7	Krueger & Zhu	2004	New York City School Choice Program	NYC	Private	1997-2000 (4 years)	K to 4	N=1801 (T=901, C=900)	Math	Null	Assumed 50/50 split of T and C due to lack of detailed data.
8	Bettinger	2005	Programa de Ampliacion de Cobertura de la Educacion Secundaria (PACES)	Colombia	Public (partly funded by World Bank)	1995-1999 (4 years)	6 to 10	N=282 (T=141, C=141)	Math	Null	Assumed 50/50 split of T and C due to lack of detailed data.
9	Cowen	2008	Charlotte Children's Scholarship Fund	Charlotte	Private	1999-2000 (1 year)	2 to 8	N=694 (T=347, C=347)	Math	Null	Assumed 50/50 split of T and C due to lack of detailed data.
10	Kremer, Miguel & Thornton	2009	Girls' Scholarship Program	Kenya	Private	2001 -2003 (3 years)	6 to 8	N=3602 (T=970, C=2632)	Math (Girls)	Null	This was a merit-based voucher for girls only (but still assigned randomly within the group of eligible students)
11	Kisida, Gutmann, Puma, Eissa & Rizzo	2013	District of Columbia OpportunityScholarship Program (OSP)	DC	Public	2004-2009 (6 years)	K to 12	N= 1330 (T=516, C=814)	Math	Null	
12	Muralidharan & Sundararaman	2015	Andhra Pradesh (AP) School Choice Experiment	Andhra Pradesh, India	Private	2008-2012 (4 years)	1 to 5	N=4385 (T=1675, C=2710)	Math	Null	
13	Wolf, Egalite & Dixon	2015	Ensure Access to Better Learning Experiences (ENABLE)	Delhi, India	Private	2011-2013 (2 years)	K to 2	N=1618 (T=835, C=783)	Math	Null	Information provided by Dr. Wolf and Anna Egalite on request

Table 2: Study Characteristics for Reading Outcomes

S. No.	Authors	Pub Year	Program Evaluated	Location	Funding	Duration of Study	Grades	Sample Size	Outcome Measure	Results	Comments/ Assumptions Made
1	Rouse	1998	Milwaukee Parental Choice Program (MPCP)	Milwaukee, WI (USA)	Public	1990-1994 (5 years)	K to 8	N=3163 (T=1582, C=1581)	Reading	Null	
2	Greene, Peterson & Du	1999	Milwaukee Parental Choice Program (MPCP)	Milwaukee, WI (USA)	Public	1990-1994 (5 years)	K to 8	N= 317 (T=237, C = 80)	Reading	Null	3- and 4- year outcomes combined. Calculated treatment and control using Table 1 (p. 199).
3	Greene	2000	Charlotte Children's Scholarship Fund	Charlotte	Private	1999-2000 (1 year)	2 to 8	N=357 (T = 223, C=134)	Reading	Null	Used the "Instrumental w/ background controls" as this was the most rigorous measure (Table 3). Used the p-value to calculate a t-statistic in order to find the standard error of the effect size. The sample size was split based off of the total 357, split by the treatment/control ratio in Table 1.
4	Angrist, Bettinger, Bloom, King & Kremer	2002	Programa de Ampliacion de Cobertura de la Educacion Secundaria (PACES)	Colombia	Public (partly funded by World Bank)	1995-1999 (4 years)	6 to 10	N=283 (T=157, T=126)	Reading	Null	
5a	Howell, Wolf, Campbell & Peterson	2002	School Choice Scholarships Foundation	NYC	Private	1997-1999 (2 years)	1 to 4	N=1199 (T= 600, C=599)	Reading	Null	Year 2 results. Assumed 50/50 treatment and control split. Combined math and reading score, so the effect size was assumed the same for both subjects.
5b	Howell, Wolf, Campbell & Peterson	2002	Parents Advancing Choice in Education	Dayton, OH	Private	1998-2000 (2 years)	K to 12	N=382 (T= 191, C = 191)	Reading	Null	Year 2 results. Assumed 50/50 treatment and control split. Combined math and reading score, so the effect size was assumed the same for both subjects.
5c	Howell, Wolf, Campbell & Peterson	2002	Washington Scholarship Fund	Washington, DC	Private	1998-2000 (2 years)	K to 8	N= 725 (T=363, C=362)	Reading	Positive	Year 2 results. Assumed 50/50 treatment and control split. Combined math and reading score, so the effect size was assumed the same for both subjects.
6	Krueger & Zhu	2004	New York City School Choice Program	NYC	Private	1997-2000 (4 years)	K to 4	N=1801 (T=901, C=900)	Reading	Null	Assumed 50/50 split of T and C due to lack of detailed data.
7	Bettinger	2005	Programa de Ampliacion de Cobertura de la Educacion Secundaria (PACES)	Colombia	Public (partly funded by World Bank)	1995-1999 (4 years)	6 to 10	N=283 (T=157, T=126)	Reading	Null	Assumed 50/50 split of T and C due to lack of detailed data.
8	Cowen	2008	Charlotte Children's Scholarship Fund	Charlotte	Private	1999-2000 (1 year)	2 to 8	N=694 (T=347, C=347)	Reading	Positive	Assumed 50/50 split of T and C due to lack of detailed data.
9	Kremer, Miguel & Thornton	2009	Girls' Scholarship Program	Kenya	Private	2001-2003 (3 years)	6 to 8	N=3602 (T=970, C=2632)	Reading (Girls)	Null	This was a merit-based voucher for girls only (but still assigned randomly within the group of eligible students).
10	Wolf, Kisida, Gutmann, Puma, Eissa &	2013	District of Columbia Opportunity Scholarship Program (OSP)	DC	Public	2004-2009 (6 years)	K to 12	N= 1328 (T=855, C=473)	Reading	Positive	
11	Muralidharan & Sundararaman	2015	Andhra Pradesh (AP) School Choice Experiment	Andhra Pradesh, India	Private	2008-2012 (4 years)	1 to 5	N=4217 (T=1607, C=2610)	Reading	Positive	
12	Wolf, Egalite & Dixon	2015	Ensure Access to Better Learning Experiences (ENABLE)	Delhi, India	Private	2011-2013 (2 years)	K to 2	N=1618 (T=835, C=783)	Reading (English)	Positive	Information provided by Dr. Wolf and Anna Egalite on request.

Fig 1: Effect sizes of vouchers on students' math outcomes

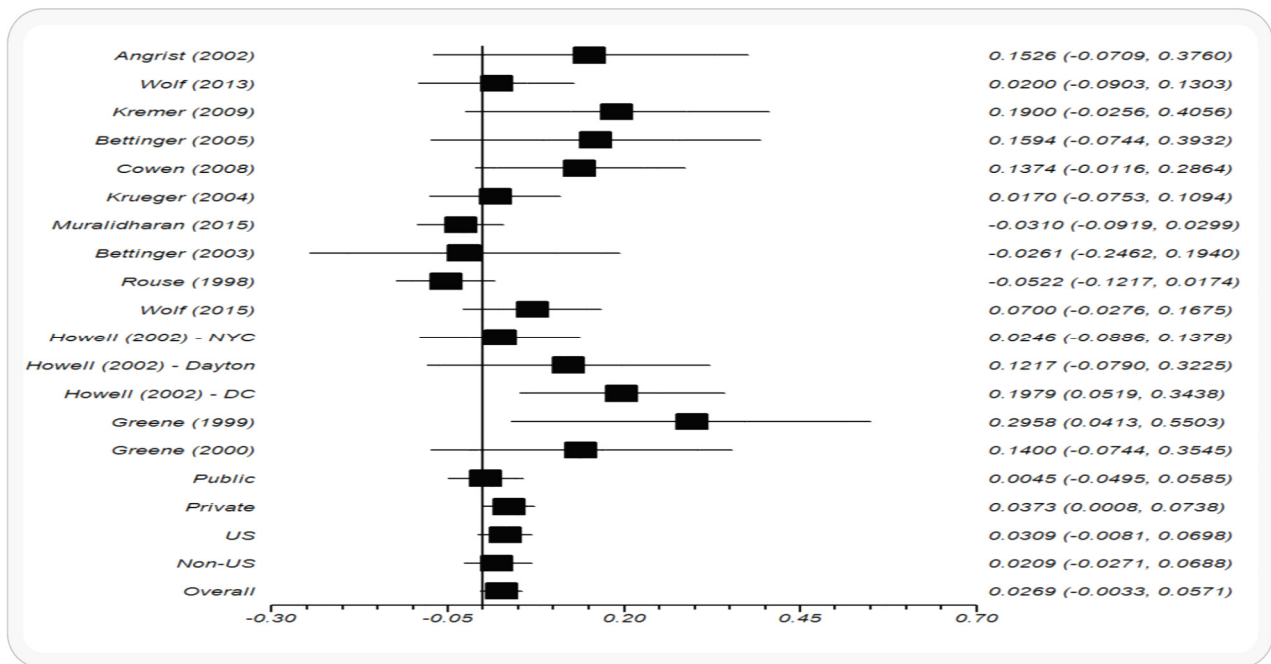|	
Angrist (2002)	0.1526 (-0.0709, 0.3760)
Wolf (2013)	0.0200 (-0.0903, 0.1303)
Kremer (2009)	0.1900 (-0.0256, 0.4056)
Bettinger (2005)	0.1594 (-0.0744, 0.3932)
Cowen (2008)	0.1374 (-0.0116, 0.2864)
Krueger (2004)	0.0170 (-0.0753, 0.1094)
Muralidharan (2015)	-0.0310 (-0.0919, 0.0299)
Bettinger (2003)	-0.0261 (-0.2462, 0.1940)
Rouse (1998)	-0.0522 (-0.1217, 0.0174)
Wolf (2015)	0.0700 (-0.0276, 0.1675)
Howell (2002) - NYC	0.0246 (-0.0886, 0.1378)
Howell (2002) - Dayton	0.1217 (-0.0790, 0.3225)
Howell (2002) - DC	0.1979 (0.0519, 0.3438)
Greene (1999)	0.2958 (0.0413, 0.5503)
Greene (2000)	0.1400 (-0.0744, 0.3545)
Public	0.0045 (-0.0495, 0.0585)
Private	0.0373 (0.0008, 0.0738)
US	0.0309 (-0.0081, 0.0698)
Non-US	0.0209 (-0.0271, 0.0688)
Overall	0.0269 (-0.0033, 0.0571)

Fig 2: Effect sizes of vouchers on students' reading outcomes

Angrist (2002)	0.202 (-0.033, 0.437)
Wolf (2013)	0.110 (-0.002, 0.222)
Kremer (2009)	0.190 (-0.026, 0.406)
Bettinger (2005)	0.212 (-0.023, 0.448)
Cowen (2008)	0.165 (0.016, 0.314)
Krueger (2004)	-0.020 (-0.112, 0.072)
Muralidharan (2015)	0.116 (0.054, 0.178)
Rouse (1998)	0.046 (-0.024, 0.116)
Wolf (2015)	0.150 (0.052, 0.248)
Howell (2002) - NYC	0.025 (-0.089, 0.138)
Howell (2002) - Dayton	0.122 (-0.079, 0.322)
Howell (2002) - DC	0.198 (0.052, 0.344)
Greene (1999)	0.243 (-0.011, 0.498)
Greene (2000)	0.180 (-0.035, 0.394)
Public	0.087 (0.033, 0.142)
Private	0.102 (0.064, 0.139)
US	0.071 (0.032, 0.111)
Non-US	0.136 (0.087, 0.185)
Overall	0.097 (0.066, 0.128)