International Public Management Journal
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/upmj20

Customer Satisfaction and Educational Outcomes: Experimental Impacts of the Market-Based Delivery of Public Education

Brian Kisida[a] & Patrick J. Wolf[a]
[a] University of Arkansas

Accepted author version posted online: 25 Mar 2015.

To cite this article: Brian Kisida & Patrick J. Wolf (2015) Customer Satisfaction and Educational Outcomes: Experimental Impacts of the Market-Based Delivery of Public Education, International Public Management Journal, 18:2, 265-285, DOI: 10.1080/10967494.2014.996629

To link to this article: http://dx.doi.org/10.1080/10967494.2014.996629

International Public Management Journal

CUSTOMER SATISFACTION AND EDUCATIONAL OUTCOMES: EXPERIMENTAL IMPACTS OF THE MARKET-BASED DELIVERY OF PUBLIC EDUCATION

BRIAN KISIDA AND PATRICK J. WOLF
UNIVERSITY OF ARKANSAS

ABSTRACT: *School choice, through private school vouchers or direct government subsidies, is a mechanism of outsourcing government services in the United States, Europe, South America, and the Pacific Rim. While extensive research exists on the effects of private school choice programs, nearly all focus on test score outcomes. Lost in the heated debates about the effectiveness of private school vouchers is substantial discussion of the effects on parental satisfaction. Drawing from a federally funded evaluation of a means-tested private school choice program in Washington, DC, we examine whether customer satisfaction is greater when education is delivered through a market-based governance structure. Because the program was oversubscribed in its early years of operation, vouchers were awarded by lottery, allowing us to experimentally determine the impacts. Our analysis reveals evidence that the program had a sustained positive impact on parental satisfaction. Moreover, positive student achievement and attainment impacts strengthen the validity of parental satisfaction as a reliable outcome measure.*

INTRODUCTION

National, state, and local governments around the globe have attempted to improve the quality of service delivery by relying on private sector entities to serve citizen clients. Outsourcing services to private providers has been especially controversial in the area of education. Although several European, South American, and Pacific Rim countries have a long history of using government funds to finance private educations, the recent implementation of private school choice programs in Sweden, England, and the United States has generated political controversy and

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/upmj

International Public Management Journal, 18(2), pages 265–285
DOI: 10.1080/10967494.2014.996629

ISSN: 1096-7494 print/1559-3169 online

academic debate. Much of that debate has focused exclusively on student test scores as an evaluative metric. Recent theoretical and empirical work in public management, however, has emphasized customer satisfaction as an important gauge of service effectiveness. In this article, we examine a school voucher field experiment in Washington, DC, to test the hypothesis that low-income parents tend to be more satisfied with their child's education when allowed to select private providers instead of a government entity. We further test whether parent self-reports of school satisfaction correlate with other established measures of student outcomes.

School vouchers are payments made by governments to families to enable them to enroll in a private school of their choosing (Wolf 2008b). Other mechanisms by which governments promote private school choice include direct government support of private schools, tax credits for contributions to private school scholarship funds, and subsidized family education savings accounts.

Government-funded private school choice programs are common throughout the world (Macedo and Wolf 2004). The Netherlands has provided government funds for students to attend private schools for a century (Vermeulen 2004). Government-sponsored private schooling is also available in Belgium, Denmark, and Sweden. In Chile, the national government has administered a private school voucher program since 1981 (Lara, Mizala, and Repetto 2011). Commonwealth countries such as Canada and New Zealand also provide government funding to students who choose to attend private schools (Campbell 2004; LaRoque 2005).

Private school vouchers are controversial in the United States and are commonly portrayed as a threat to public education. Currently, 35 voucher or voucher-like programs exist in the U. S., funded either directly by governments or indirectly through corporate or individual tax credits. In all cases, these voucher programs are only available to low-income families or students attending low-performing government schools. Nevertheless, political controversy surrounding school vouchers remains.

Nearly a dozen rigorous empirical evaluations of school voucher or voucher-type programs have been conducted in the United States (Wolf 2008a). Those studies have focused primarily on the impact of these programs on student test scores. Recently, the field of public management has considered customer satisfaction as a more complete measure of program performance, especially in the field of education (e.g., Charbonneau and Van Ryzin 2011; Favero and Meier 2013). Unlike highly specific measures like student test scores, parent satisfaction presumably incorporates more complete aspects of what we expect schools to provide, including school safety and the nurturing of character traits.

The data we use to test the impact of school choice on parental satisfaction come from the District of Columbia Opportunity Scholarship Program (OSP). The OSP is the first and only federally funded private school voucher program in the United States, established as a five-year pilot program in 2004. The Opportunity Scholarships provided by the program were vouchers that could be used at any of more than 70 participating private schools and were worth up to $7,500. In most cases, the voucher was sufficient to cover school tuition, meaning families did not have to "top it up." From 2004 to 2009, nearly 8,500 students applied for and nearly 3,000 used an Opportunity Scholarship. After a protracted political battle, the

program was reauthorized and expanded as the final element of a budget agreement between Republicans in Congress and President Obama in 2011.

Like most school voucher and tax-credit scholarship programs in the United States, the DC OSP targeted disadvantaged students (Wolf 2008a). A substantial literature documents the relationship between disadvantaged students and less-effective public schools in terms of high school completion, teacher quality, facilities, and school safety (e.g., Boyd et al. 2008; Lippman, Burns, and McArthur 1996). Additionally, survey data have shown that income is negatively correlated with school satisfaction (Charbonneau and Van Ryzin 2011). To be eligible, a student's family income had to be at or below 185% of the federal poverty level—about $36,000 for a family of four in 2004. Among the initial cohort of eligible applicants, 99% were African American or Hispanic and their annual family income averaged $18,742 (Wolf et al. 2005). Since the program was oversubscribed, vouchers were awarded by lottery, though preference in the lottery was given to public school students attending schools that had been designated "schools in need of improvement" (SINI) under the federal government's accountability system. The scholarship lottery provided a rare opportunity to evaluate a government program using a randomized field trial.

THEORY

This research represents a practical application of two general theories in public management: rational choice and organizational governance. In his classic work *Inside Bureaucracy*, Downs (1964, 2) posits that "bureaucratic officials, like all other agents in society, are significantly . . . motivated by their own self-interests." Bureaucrats tend to be rational in that their behavior is shaped by incentive systems in their organizational environment, whether pecuniary or purposive (Brehm and Gates 1999).

For organizations, governance structure refers to whom and what determines how the entity operates. Ouchi (1980) describes three archetypal governance structures for organizations: hierarchies, markets, and clans. Organizations governed by hierarchies have the benefits of "command-and-control" but the liabilities of rigidity and insularity. Organizations governed by markets have the benefits of flexibility and customer responsiveness but the liabilities of unpredictability and inequity. Organizations governed by clans or "strong culture" organizations have the benefits of a shared value system and missionary zeal but the liabilities of exclusivity and closemindedness. As Wilson (1989) has noted, a major advantage of market-based governance structures is the feedback generated when customers "vote with their feet" to receive services from one organization and not another.

Chubb and Moe (1990, 2) claim that the causes of U.S. public schools' inefficiency and unresponsiveness "are, in fact, the very institutions that are supposed to be solving the problem: the institutions of direct democratic control." They argue that the effectiveness of schools is "reflective of the institutional contexts in which they operate" (Chubb and Moe 1990, 2). In light of this, they argue that the key to solving the problem is reform of the governance structure of schools through private school choice.

Chubb and Moe's (1990) thesis is based on a perceived failure of the democratic institutions that govern schooling. In the current democratic process of school governance, they argue that the "most powerful political groups by far are those with vested interests in the current institutional system: teachers' unions and myriad associations of principals, school boards, superintendents, administrators, and professionals" (Chubb and Moe 1990, 11). This is problematic because the rational self-interests of school leaders may not be aligned with or incentivized to meet the needs of parents and students.

In light of these concerns, Chubb and Moe (1990) argue that, in a market setting, schools would more responsibly serve the needs of their clients. Because most school board elections are held separately from elections for more prominent political positions and are dominated by mobilized interest groups such as teachers unions, market mechanisms may be more responsive than the current institutionalized structure of democratic governance and regulation. Moreover, in a market-based educational setting, clients are able to sort themselves into schools that may be a better match for their particular needs. By increasing parental power in the consumer/provider relationship, and allowing parents to sort themselves into schools they choose, the market approach can potentially offer greater institutional responsiveness and greater compatibility between provider and client.

Though skeptical of the market approach to education services, Henig (1994, 5) articulates how democratic accountability could be generated from both government regulation and market pressures:

> School choice proposals would shift the focus of educational decision making from the government arena—in which elected officials, public bureaucracies, and organized interest groups are central players—to a market-based arena, in which the personal preferences of children and their families presumably will have a more prominent place. To the familiar claim that market forces are *more efficient* modes for allocating scarce resources is added the claim that they are *more responsive* to the felt needs and desires of the average citizen. In a strange twist, the shift away from democratic processes and institutions is defended by reference to values we associate with democracy. Markets, it has been argued, can be more democratic than democracy itself.

Ultimately, these theoretical claims can be informed by empirical inquiry. The extensive research examining the quality of private schooling in the United States, Europe, and South America, however, tends to suffer from two shortcomings. First, most of the evaluations of private school choice are observational and prone to selection problems (e.g., Lubienski and Lubienski 2013). Since parents who enroll their child in private school are self-selected, any subsequent differences in outcomes between private and public schools are based on some unknown combination of unobserved selection factors and actual school performance differences. Second, nearly all of the evaluations of private education focus on student test score outcomes.[1] Test score outcomes, while important, are imprecise

measures of the quality of an educational experience and fail to capture the broad range of services parents and policymakers desire from schools (Favero and Meier 2013). Parental satisfaction with schools is a critical area in need of further rigorous study.

This paper addresses these gaps, drawing from a federally funded experimental evaluation of the DC Opportunity Scholarship Program. Because demand exceeded the supply of available vouchers, a lottery awarded private school vouchers to 1,387 program applicants, while the remaining 921 applicants served as the control group. We analyze the impacts of the program on parents and students who used these scholarships to choose a private school, with a deliberate emphasis on parental satisfaction with various aspects of their child's schooling.

PRIOR RESEARCH

Prior school voucher studies have generally focused on student achievement. Though the findings have been modestly positive, they have been inconsistent in their pattern of results and have yet to produce a scholarly consensus (Wolf 2008b; Barrow and Rouse 2008; Dronkers and Avram 2010; Dronkers and Robert 2008; Angrist et al. 2002; Wolf et al. 2013).

Existing literature on private school vouchers in the United States includes observational and experimental studies. Observational studies take the populations of voucher or private school students and public school students as they have naturally occurred and attempt to control for confounding factors that introduce selection bias. Although observational studies of school vouchers can have greater external validity than experiments, they have less internal validity and are subject to more biases than experiments because unmeasured self-selection factors cannot be ruled out (e.g., Barrow and Rouse 2008, 7; Levin 1998, 374–375). Observational studies of school voucher programs in Milwaukee have reported few statistically significant achievement effects (Witte 2000; Witte et al. forthcoming). Observational analyses of Cleveland's voucher program have reported a mix of positive and null achievement results (Metcalf et al. 2003; Greene, Howell, and Peterson 1998).

Findings from evaluations of the effects of government-subsidized private schools in Europe on student achievement have varied from moderately large positive effects to small positive effects, depending on the methods and models used to control for student self-selection (Dronkers and Avram 2010; Dronkers and Robert 2008). Evaluations of Chile's universal private school voucher program have been more mixed, with many reporting modest and contingent positive test score effects alongside evidence that the program exacerbated the income-based stratification of educational outcomes (e.g., Mizala and Torche 2012; Lara, Mizala, and Repetto 2011; Sapelli 2005; McEwan, Urquoila, and Vegas 2008).

Experimental analyses take advantage of scholarship lotteries to assign eligible applicants into randomized "treatment" (offer of a voucher) and "control" (no offer) groups. Since only chance and the voucher offer distinguish the groups,

significant differences in outcomes can be attributed to the program (e.g., Cook and Payne 2002; Cook and Campbell 1979, 56). The outcomes of the control group represent what would have happened to the treatment group absent winning the lottery, making the control group the ideal counterfactual. Because of this strength, experiments have been dubbed the "gold standard" for evaluating programs (e.g., Tufte 2006, 145; Boruch, De Moya, and Snyder 2002, 74).

Prior to the DC Opportunity Scholarship Program, a total of nine analyses of achievement outcomes had been conducted using experimental data from voucher and voucher-type lotteries in the United States. Five of these analyses—two each of Charlotte and Milwaukee and one of DC—reported statistically significant achievement gains from vouchers for the overall sample of participants in at least one year of the study (Greene 2001; Cowen 2008; Greene, Peterson, and Du 1999; Rouse 1998; Howell et al. 2002). Three of the analyses—two of New York and one of Dayton, Ohio—found significant achievement gains only for African American students (Mayer et al. 2002; Barnard et al. 2003). One study reported no significant achievement gains from the New York program (Krueger and Zhu 2004). The only major random assignment evaluation of a school voucher-type program outside the United States, which we know of, is of a large pilot program in Colombia (Angrist et al. 2002). Similar to experimental studies in the United States, the study found that student achievement increased modestly due to the voucher, and the size and significance of the impacts differed across educational subject domains.

In summary, the results from previous experimental and observational studies of voucher-like policies provide modest support for Milton Friedman's (1955) claim that vouchers will improve student achievement in the form of scores on standardized tests. Still, debate surrounding expansion of school vouchers in the United States remains lively (e.g., Ravitch 2013).

Generally lost in the disputes about the achievement effects of private school vouchers is substantial discussion of the effects on parental satisfaction. Yet, focus-group research finds that parents participating in voucher programs are more likely to mention indicators of success such as school safety, attentiveness to homework, and aspirations to attend college than they are to cite test scores (Stewart and Wolf 2014). Test scores may be an available metric by which to assess voucher programs, but they fail to capture the full range of services parents desire from schools (Favero and Meier 2013).

Only a handful of prior studies of school vouchers have examined their effect on parent satisfaction (Witte 2000; Witte et al. 2008), and only one has employed experimental methods (Howell and Peterson 2006). All of the studies find that parents are more satisfied when given the opportunity to choose a private school. In these studies, however, satisfaction is treated as an impact of little consequence compared to student achievement. In part, the lack of emphasis on parental satisfaction stems from claims from choice skeptics who argue that parents, especially low-income parents, are unqualified to make informed decisions when choosing schools (Berliner and Biddle 1995). Empirical evidence, however, finds that parents charged with choosing a school gather more information about their child's school than

non-choosing parents (Kisida and Wolf 2010; Schneider et al. 2000). As such, we might expect that parents exercising choice are well-equipped to perceive school quality and rely on those perceptions when reporting school satisfaction. Moreover, previous research has demonstrated that subjective measures, like parent survey data, tend to correspond with "objective" measures of school quality, like aggregate test score levels (Charbonneau and Van Ryzin 2011).

DATA AND METHODS

Informed by theory, we bring two research hypotheses to this analysis. First, using a voucher to attend a private school will have a positive impact on satisfaction. Second, positive impacts on satisfaction will align with other important student outcomes. Thus, we expect that measures of student achievement and educational attainment will validate any positive impacts on parental satisfaction levels. If, however, we observe positive satisfaction impacts but negative effects on achievement and attainment, it would suggest that voucher users' satisfaction may be misplaced.

The foundation of our evaluation is a field experiment that compares the outcomes of eligible student applicants randomly assigned to receive an offer (treatment group) or not receive an offer (control group) of an OSP scholarship. We operationalized the lottery as a stratified random sample, also known as "block randomization" due to provisions of the program legislation that prioritized students attending a SINI school. Because certain types of students faced different probabilities of assignment to the treatment group, the outcome data for each student are weighted by the inverse of the likelihood of scholarship award (see Online Appendix for full details of the randomization).

The data for the analysis of parent satisfaction come from annual surveys. Response rates over the four years of the study ranged from 66% to 72%. Parents were asked to assign grades to their child's school on a scale ranging from A through F, which we have recoded to 0-4 for our outcome analysis. In order to get a broad picture of parental satisfaction across multiple characteristics, we created a second scale composed of the average ratings of individual school characteristics using an ordinal scale ranging from 1 to 4, representing the responses of *very dissatisfied, dissatisfied, satisfied*, and *very satisfied*. Parents were asked to rate satisfaction with their school's location, safety, class sizes, facilities, respect between teachers and students, teacher communication, parental support, discipline, racial mix of students, amount of observance of religious traditions, and services for students with special needs. The Cronbach's alpha for the scale items was .95 across all years.

A second survey measuring high school graduation focused on the 500 randomly assigned applicants scheduled to graduate high school by June 2009. Of these students, 202 were assigned to the treatment group and 298 to the control group. In the summer of 2009, their parents were contacted to determine if their child had graduated from high school, was still enrolled, or had dropped

out. Thus, the outcome measure used here is parent self-reports of students' educational attainment. A total of 63.2% of these parents responded to the survey—62.9% of treatment group parents and 63.4% of control group parents (Wolf et al. 2010, Table A-9, A-27). A non-experimental study of school vouchers and educational attainment in Milwaukee using the same parent phone survey employed here and administrative records from schools found that parent reports were consistent with school graduation records in more than 98% of cases (Cowen et al. 2013).

Measures of reading and math achievement used in this analysis come from annually administered Stanford Achievement Tests (version 9). Response rates for outcome testing across the four years ranged from 69% to 77%. Students in the sample were relatively low-performing, with average scores in the lower third of national percentile rankings.

Not all participants in experimental evaluations comply with their random assignment to the treatment. In the DC OSP experiment, 25% of students offered a voucher never used it to attend a private school. Because the outcomes of treatment no-shows count towards the average for the treatment group as a whole, these violations of random assignment have the practical effect of underestimating the impact of actually using the treatment.

To accurately estimate the impact of using a voucher to purchase private schooling, we use instrumental variables (IV) analysis. In practice, IV analysis involves running a two-stage regression to arrive at unbiased estimates of the effects of using a voucher to attend a private school on a particular outcome (Howell et al. 2006, 49–51; Angrist et al. 2002). In the first stage, the results of the treatment lottery and baseline student characteristics predict the likelihood that individual students attended a private school with the use of a scholarship.[2] In the second stage, that predicted estimate operates as the independent variable of interest and is used to estimate the effect of attending a private school on subsequent outcomes. Since the estimate of voucher use is based on the lottery outcome and baseline descriptive statistics, and it replaces the endogenous measure of actual scholarship use that is tainted by self-selection, the IV procedure generates unbiased estimates of the effect of voucher *use* for the participants. The IV estimates are unbiased because, conditional on baseline characteristics that are controlled for in the analysis, the lottery variable yields the same predicted probability of voucher use for all treatment group members, whether they actually used their voucher or not (see Online Appendix Tables A.1.1, A.2.1, and A.3.1 for intention to treat estimates of the *offer* of a scholarship on parental satisfaction).

For the impacts on parental satisfaction, the first stage of the IV analysis is used to predict scholarship usage in the corresponding survey year so that usage temporally aligns with annually reported satisfaction. For the analysis of educational attainment and achievement, we predict total years of scholarship use as a result of the voucher offer. This approach takes into account the potentially cumulative effect of higher or lower levels of treatment dosage. Finally, for the attainment analysis, we also predict scholarship ever-use in order to provide meaningful interpretations of treatment usage on graduation rates.

We provide additional consideration of the impacts on policy-relevant subgroups of participants. The subgroups include students performing relatively higher or lower at baseline,[3] girls or boys, elementary or middle and high-school students, and whether the students were attending a SINI school when they applied to the program. Interaction terms were added to regression models to derive the average program impact by subgroup.[4] In our results section, some impacts are presented as effect sizes. An effect size is a measure of magnitude obtained by rescaling the beta coefficients as a percentage of one standard deviation of the distribution of each outcome for the control group.

PARENTAL SATISFACTION RESULTS

Average School Grade

Each year, parents were asked to assign their school a grade of A–F. Across all years, the full sample of parents consistently assigned their child's school a higher grade if they attended a private school in that year (Figure 1). The effect sizes for scholarship use based upon the full sample are generally large, ranging from .34 to .53, and all are statistically significant beyond the 90% confidence level (see Online Appendix Table A.1.2 for effect sizes across all subgroups). Average grades given by treatment users were relatively stable across the four years we examined.

These results are consistent across parents of male and female students across all years (Table 1). School grades are also significantly higher for scholarship-using parents of higher- and lower-performing students, except for lower-performing

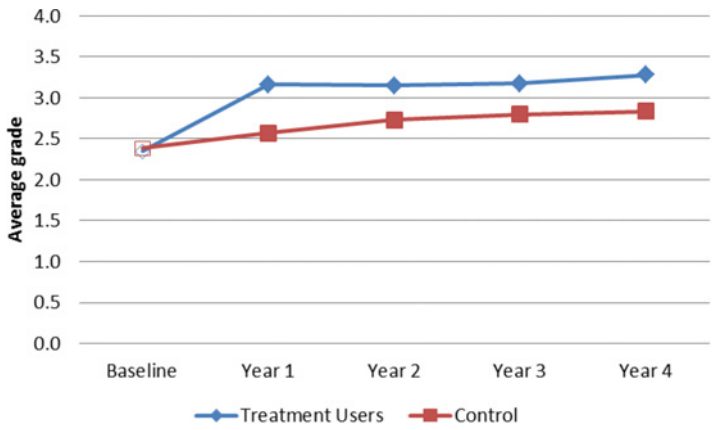

Figure 1. Average School Grade Assigned by Parents. *Notes*: Instrumental variable estimations are derived from the lottery results and a consistent set of baseline covariates. A–F scale was recoded to a numeric scale that ranges from 0–4. See Online Appendix Table A.1.2 for additional details and effect sizes. Sample weights used.

TABLE 1
Average School Grade Assigned by Parents, Full Sample and Subgroups

		Year 1	*Year 2*	*Year 3*	*Year 4*
Full sample	Treatment user mean	3.16	3.15	3.17	3.28
	Control mean	2.57	2.73	2.79	2.83
	Difference	0.59***	0.42***	0.38***	0.45***
Females	Treatment user mean	3.21	3.23	3.18	3.33
	Control mean	2.60	2.67	2.83	2.85
	Difference	0.60***	0.56***	0.35***	0.48***
Males	Treatment user mean	3.10	3.05	3.17	3.21
	Control mean	2.53	2.79	2.76	2.81
	Difference	0.57***	0.26**	0.41***	0.40*
Higher performers	Treatment user mean	3.19	3.24	3.30	3.27
	Control mean	2.66	2.77	2.85	2.91
	Difference	0.54***	0.47***	0.44***	0.36**
Lower performers	Treatment user mean	3.08	3.12	2.86	3.33
	Control mean	2.37	2.64	2.66	2.66
	Difference	0.71***	0.47*	0.20	0.67**
Grades K–5	Treatment user mean	3.36	3.30	3.29	3.35
	Control mean	2.62	2.80	2.90	2.95
	Difference	0.74***	0.50***	0.40***	0.40***
Grades mid/high	Treatment user mean	2.90	2.97	3.03	3.20
	Control mean	2.51	2.66	2.68	2.67
	Difference	0.39***	0.31**	0.35**	0.53*
SINI	Treatment user mean	2.97	3.04	3.04	2.95
	Control mean	2.45	2.56	2.76	2.86
	Difference	0.52***	0.48***	0.27*	0.09
Non-SINI	Treatment user mean	3.29	3.24	3.27	3.48
	Control mean	2.65	2.86	2.81	2.81
	Difference	0.64***	0.37***	0.45***	0.67***

Notes: Instrumental variable estimations are derived from the lottery results and a consistent set of baseline covariates. See Online Appendix Table A.1.2 for additional details and effect sizes. Sample weights used. Year 1 $N = 1{,}680$; year 2 $N = 1{,}550$; year 3 $N = 1{,}410$; year 4 $N = 1{,}230$ (*N*s rounded to the nearest 10 per IES publication policy).

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

students in the third year. The grades assigned by treatment users are statistically higher than the control group across grade level subgroups across all years, and for scholarship users from non-SINI schools. Parents of students from SINI schools were not significantly more satisfied in the fourth year.

Parental Satisfaction Scale

The results from our satisfaction scale composed of the ratings of 12 individual school characteristics are consistent with the previous measure of school grade.

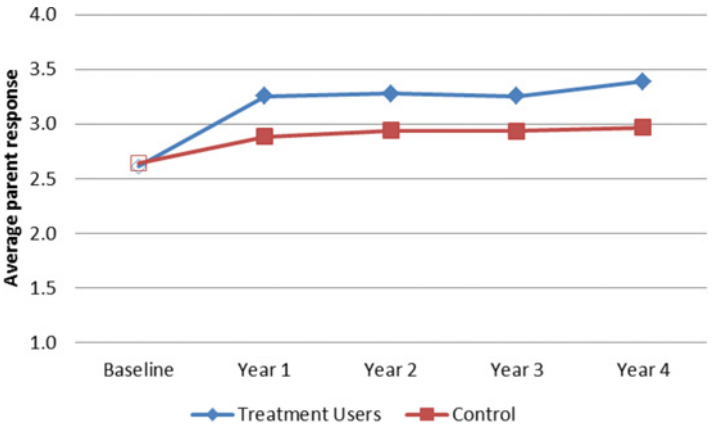

Figure 2. Parent Satisfaction Scale. *Notes*: Instrumental variable estimations are derived from the lottery results and a consistent set of baseline covariates. The survey Likert scale was recoded to a numeric scale that ranges from 1–4. See Online Appendix Table A.2.2 for additional details and effect sizes. Sample weights used.

Across the full sample, parents of treatment users consistently rate their child's school higher than control group parents (Figure 2). Effect sizes are large, ranging from .42 to .57 of a standard deviation, and all are statistically significant (see Online Appendix Table A.2.2 for effect sizes across subgroups).

The results for subgroups are consistent with the overall results. Across parents of male and female students, treatment users are more satisfied with their school's characteristics (Table 2). The same holds true for parents of students in lower and higher grade levels and parents of SINI and non-SINI students. Satisfaction with school characteristics is also significantly higher across all years for parents of higher- and lower-performing students in the treatment group, except for lower-performing students in year three.

Parent Satisfaction Impacts on Key School Characteristics

Urban private schools have a strong reputation regarding particular school characteristics, such as academic quality, teacher communication with parents, and school safety. They have a less favorable reputation regarding school facilities, class size, and services for students with special needs. To investigate possible differences subsumed by our 12-item scale, we examine these six important features in isolation.

The results are largely consistent with our two previous measures, with treatment users consistently ranking these individual characteristics significantly higher across each year of the program (Table 3). An exception is services for students with special needs. Parents generally are least satisfied with this school characteristic and the difference between treatment and control group parents is not statistically significant beyond the first year of the program (see Online Appendix Table A.3.2 for effect sizes across all individual characteristics).

TABLE 2
Parent Satisfaction Scale, Full Sample and Subgroups

		Year 1	*Year 2*	*Year 3*	*Year 4*
Full sample	Treatment user mean	3.25	3.28	3.26	3.39
	Control mean	2.88	2.94	2.93	2.97
	Difference	0.37***	0.34***	0.32***	0.42***
Females	Treatment user mean	3.30	3.25	3.27	3.38
	Control mean	2.84	2.93	2.94	2.94
	Difference	0.47***	0.32***	0.34***	0.44***
Males	Treatment user mean	3.20	3.31	3.24	3.40
	Control mean	2.94	2.95	2.93	3.00
	Difference	0.26***	0.36***	0.31***	0.40***
Higher performers	Treatment user mean	3.26	3.33	3.29	3.41
	Control mean	2.92	2.96	2.95	3.00
	Difference	0.34***	0.37***	0.34**	0.41***
Lower performers	Treatment user mean	3.25	3.27	3.17	3.36
	Control mean	2.79	2.90	2.90	2.89
	Difference	0.45***	0.37**	0.27	0.47***
Grades K–5	Treatment user mean	3.35	3.41	3.32	3.39
	Control mean	2.92	2.97	2.99	3.04
	Difference	0.43***	0.44***	0.32***	0.35***
Grades mid/high	Treatment user mean	3.13	3.12	3.18	3.43
	Control mean	2.85	2.91	2.86	2.87
	Difference	0.28***	0.21**	0.32***	0.56**
SINI	Treatment user mean	3.17	3.23	3.20	3.24
	Control mean	2.80	2.81	2.88	2.96
	Difference	0.36***	0.42***	0.32**	0.29*
Non-SINI	Treatment user mean	3.32	3.32	3.30	3.49
	Control mean	2.95	3.04	2.97	2.97
	Difference	0.37***	0.27***	0.32***	0.52***

Notes: Instrumental variable estimations are derived from the lottery results and a consistent set of baseline covariates. See Online Appendix Table A.2.2 for additional details and effect sizes. Sample weights used. Year 1 $N = 1{,}690$; year 2 $N = 1{,}570$; year 3 $N = 1{,}440$; year 4 $N = 1{,}230$. Ns rounded to the nearest 10 per IES publication policy.

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

EDUCATIONAL OUTCOMES

Educational Achievement

A persistent criticism of school choice programs is that parents may choose schools that do not produce the outcomes expected from publicly funded education. Were this the case, then we would question the value of parental satisfaction as a reliable measure of school performance. We test this by examining two key socially desirable outcomes: student achievement and educational attainment.

TABLE 3
Treatment User Effect Sizes for Individual School Characteristics

Key Satisfaction Scale Items	*Year 1*	*Year 2*	*Year 3*	*Year 4*
Academic quality	0.48***	0.40***	0.35***	0.48***
	(0.07)	(0.08)	(0.09)	(0.13)
	1,640	1,520	1,380	1,190
Teacher and parent communication	0.39***	0.42***	0.28***	0.45***
	(0.07)	(0.08)	(0.09)	(0.13)
	1,640	1,530	1,390	1,200
Safety	0.42***	0.34***	0.39***	0.51***
	(0.07)	(0.08)	(0.09)	(0.13)
	1,650	1,540	1,400	1,210
School facilities	0.48***	0.42***	0.29***	0.47***
	(0.07)	(0.08)	(0.10)	(0.13)
	1,630	1,530	1,380	1,190
Class sizes	0.55***	0.43***	0.44***	0.52***
	(0.07)	(0.08)	(0.10)	(0.13)
	1,640	1,540	1,400	1,200
Services for students with special needs	0.21**	0.18	0.19	0.13
	(0.10)	(0.11)	(0.14)	(0.19)
	1,000	930	860	730

Notes: Instrumental variable estimations are derived from the lottery results and a consistent set of baseline covariates. Robust standard errors in parentheses. Sample weights used. See Online Appendix Table A.3.2 for additional details and items. Ns rounded to the nearest 10 per IES publication policy.
***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

In terms of reading achievement, we find treatment users perform slightly better than the control group (Table 4). For each year of scholarship use, students gain 5% of a standard deviation. For subgroups, the gains are largest for female students, higher-performing students, middle/high school students, and students from non-SINI schools. The impacts for male students, lower-performing students, and SINI students are not statistically significant.

For math achievement, there is no discernible difference between treatment users and the control group for the full sample. Moreover, regardless of the subgroup examined, treatment users perform no better or worse than control group students in math achievement.

Educational Attainment

Overall, the probability of a student graduating from high school increased by 21 percentage points if they used a scholarship, from 70% for the control group to 91% for treatment users (Table 5). Estimated as a function of the number of years of scholarship use, each year of use increases the likelihood of high school graduation by an average of nine percentage points. Girls appear to especially benefit, as the likelihood of graduating high school tops out at 100% for female scholarship users

TABLE 4
Yearly Achievement Impacts for Treatment Users

	Reading		*Math*	
	Scale Score Points	*Effect Size*	*Scale Score Points*	*Effect Size*
Full sample	1.81**	0.05**	0.05	0.00
	(0.74)	(0.02)	(0.66)	(0.02)
Females	2.20**	0.06**	1.13	0.03
	(0.91)	(0.03)	(0.87)	(0.03)
Males	1.36	0.04	−1.19	−0.04
	(1.18)	(0.03)	(1.01)	(0.03)
Higher performers	1.97**	0.06**	−0.10	−0.00
	(0.86)	(0.02)	(0.79)	(0.02)
Lower performers	1.43	0.04	0.42	0.01
	(1.39)	(0.04)	(1.22)	(0.04)
Grades K–5	1.69*	0.04*	0.07	0.00
	(0.95)	(0.02)	(0.92)	(0.02)
Grades mid/high	1.98*	0.06*	0.03	0.00
	(1.15)	(0.03)	(0.88)	(0.03)
SINI	1.12	0.03	0.41	0.02
	(1.20)	(0.03)	(1.00)	(0.04)
Non-SINI	2.26**	0.06**	−0.19	−0.01
	(0.94)	(0.02)	(0.88)	(0.03)

Notes: Instrumental variable estimations are derived from the lottery results and a consistent set of baseline covariates. Robust standard errors in parentheses. Sample weights used. Total $N = 1{,}990$; including female = 1,000, male = 1,000, high performers = 1,350, low performers = 640, Grade K–5 = 1,030, Mid/High school = 960, SINI = 870, Non-SINI = 1,130. Ns rounded to the nearest 10 per IES publication policy.

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

compared to 75% for females in the control group. The impact of the program on the educational attainment of males was not statistically significant, but both lower-performing and higher-performing subgroups experienced attainment impacts. The impact on students from SINI and non-SINI schools was 21 and 20 percentage points, though the impact on students from non-SINI schools was not statistically significant.

DISCUSSION

Our first hypothesis is confirmed by the data analysis. Parents of students attending private schools through the DC Opportunity Scholarship Program were more satisfied with their child's school. Whether measured by a school GPA or by a Likert-type satisfaction measure, satisfaction was higher for parents of students using a voucher to access private schools. When parsed into smaller subgroups,

TABLE 5
Attainment Impacts for Treatment Users

	Received a High School Diploma			
	Ever Treatment User	*Control*	*Difference*	*Gain per Year of Scholarship Use*
Full sample	.91	.70	.21***	.09***
			(.08)	(.03)
Females	1.00	.75	.26***	.12***
			(.09)	(.04)
Males	.81	.66	.14	.05
			(.14)	(.05)
Higher performers	.99	.79	.20**	.08**
			(.09)	(.04)
Lower performers	.79	.49	.30*	.14*
			(.17)	(.07)
SINI schools	.87	.66	.21**	.09**
			(.09)	(.03)
Non-SINI schools	1.00	.82	.20	.09
			(.24)	(.10)

Notes: Instrumental variable estimations are derived from the lottery results and a consistent set of baseline covariates. Impact estimates are reported as marginal effects. Robust standard errors in parentheses. Sample weights used. Total $N = 320$; including female = 150, male = 170, high performers = 210, low performers = 110, SINI = 230, Non-SINI = 90. "Ever Treatment User" results derived from a two-stage least squares model using the lottery as an instrument in the first stage to predict if a student ever used a scholarship. "Gain per Year of Scholarship Use" is derived from a two-stage least squares model using the lottery as an instrument in the first stage to predict total years of scholarship use. Treatment group students in this attainment analysis, on average, used their scholarships between two and three years. *N*s rounded to the nearest 10 per IES publication policy.

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

the impacts of the program on customer satisfaction remained statistically significant for most groups in all years. Exceptions include the parents of students in the lower third of the performance distribution at baseline, for whom impacts were not statistically significant during the third year of the program. Additionally, the average school grade reported by parents of students from SINI schools was not statistically different as a result of the treatment in the fourth year. Though parents of lower performers and parents of students from SINI schools were more satisfied in all but these instances, their less consistent pattern of results suggests that finding better options through school choice may be more challenging for the most disadvantaged families of an already disadvantaged population.

Just as the voucher program's impacts on customer satisfaction did not vary much based on the type of student, it also did not vary much based on the specific schooling feature assessed. For this analysis, we selected three specific schooling features for which urban private schools have a strong reputation for excellence—academic quality, teacher communication with parents, and school safety—as well

as three features for which urban private schools have a less sterling reputation—school facilities, class size, and services for students with special needs. Across five of these features, the effect of using a voucher resulted in a significant increase in customer satisfaction compared to control group parents. This pattern of results is consistent with the idea that access to private schooling generates an overall "halo" effect, leading parents to be relatively non-differentiated regarding their evaluation of specific school features.

Is it possible that this "halo effect" is actually a "placebo effect" from having won something? Similarly, some choice skeptics have argued that parents may respond to an experimental setting by changing their short-term behavior due to a "Hawthorne effect" (Henig 1996). If greater school satisfaction from winning a voucher lottery were a placebo or Hawthorne effect, with no real substantive content, then the effect would likely fade over time as the recipient became further removed from "victory." Our data do not support such a trend. The positive impacts on parental satisfaction are remarkably stable over time both in terms of their magnitude and statistical significance.

One reason that parents tended to be more satisfied with their child's school if using a voucher could be due to improved educational outcomes their children experienced in reading achievement and high school graduation. Previous research suggests that parents view their child's educational attainment as a primary concern. The significantly higher levels of reading achievement and high school graduation as a result of using a scholarship at least partially confirm our second research hypothesis and indicate that the satisfied DC parental customers are choosing schools that offer tangible benefits related to the goals of public education. The consistent finding of no impacts of the voucher program on math achievement, however, does not support the second research hypothesis. Nevertheless, the fact that parental satisfaction increased in tandem with higher reading achievement, higher attainment, and comparable levels of math achievement when compared to the control group suggests parental satisfaction is not misplaced.

There are important limitations to this study. Like most experimental evaluations, it focused on a distinct set of study participants in a particular situation. The study was limited to parents and students in low-income families in Washington, DC, who applied to the program in 2004 and 2005. Other families who applied to the program in different years, higher-income families ineligible for the program, and low-income families in other places could experience school vouchers differently than this distinctive group.

Parents that applied were likely among the most dissatisfied public school parents. Families who did not apply may not have experienced gains in satisfaction were they offered a voucher to choose a private school because they might not have been as motivated to explore alternatives. On the other hand, applicants might be especially demanding of their child's school, resulting in heightened scrutiny of school conditions. It is possible that winning the lottery raised the expectations of parents in the treatment group. A growing body of public administration literature has explored the interplay of expectations and satisfaction through the principle that satisfaction equals performance minus expectations—formally referred to as the "expectancy disconfirmation" model (e.g., James 2009; Poister and Thomas 2011;

Van Ryzin 2004). If winning the lottery in fact led to higher expectations, the self-reported satisfaction of the treatment group may actually underestimate the increase in school performance they experienced (James 2009). At the same time, lottery losers may have been particularly dejected, which may have led to lower overall expectations in the control group. This is an important consideration for future research, especially in experimental research situations where subjects are aware of the intended effect of the intervention.

CONCLUSION

Customer satisfaction is an important measure of organizational performance. Although there can be ambiguity regarding who is the customer when government organizations deliver a service, in the area of K–12 education there is general agreement that parents and students are the primary customers of schools. Theory suggests that private schools will be more responsive to their customers than will government schools operating as organizational monopolies. Moreover, many private schools differentiate themselves by providing customized learning environments, and parents given the option to choose a private school may find a schooling environment that is more compatible with their child's particular needs. With this mix of responsiveness and compatibility, private school choice has the potential to increase customer satisfaction. Although this is only one evaluation of a particular school voucher program, the results of the experimental analysis suggest that low-income urban parents are more satisfied customers when using private education providers, and parental satisfaction seems to be well-aligned with other important indicators of successful schooling.

SUPPLEMENTAL MATERIAL

A supplemental Online Appendix for this article can be accessed on the publisher's website at http://dx.doi.org/10.1080/10967494.2014.996629

NOTES

1. For a rare exception, see Howell and Peterson 2006, Chapter 7.
2. Characteristics measured at baseline and used in our regression models include reading and math baseline test scores, age, grade, household income, number of children in household, number of months at current residence, and indicator variables for having attended a SINI school, gender, African American, in special education at baseline, mother employed full-time, and mother employed part-time.
3. Higher-performing students are defined as those in the upper two-thirds of the achievement distribution at baseline. Lower-performing students are in the lower one-third of the distribution at baseline.
4. For clarity of presentation, we only present the different effects of the treatment on subgroups. We do not report the differences in treatment effects between subgroups, which is what the interaction terms specifically estimate.

REFERENCES

Angrist, J., E. Bettinger, E. Bloom, E. King, and M. Kremer. 2002. "Vouchers for Private Schooling in Columbia: Evidence from a Randomized Natural Experiment." *The American Economic Review* 92: 1535–1558.

Barnard, J., C. E. Frangakis, J. L. Hill, and D. B. Rubin. 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.

Barrow, L. and C. E. Rouse. 2008. "School Vouchers: Recent Findings and Unanswered Questions." *Economic Perspectives* 32(3): 2–16.

Berliner, D. C. and B. J. Biddle. 1995. *The Manufactured Crisis: Myths, Fraud, and the Attack on America's Public Schools*. New York: Harper Collins.

Boruch, R., D. de Moya, and B. Snyder. 2002. "The Importance of Randomized Field Trials in Education and Related Areas." Pp. 50–79 in F. Mosteller and R. Boruch, eds., *Evidence Matters: Randomized Trials in Education Research*. Washington, DC: Brookings.

Boyd, D., H. Lankford, S. Loeb, J. Rockoff, and J. Wyckoff. 2008. "The Narrowing Gap in New York City Teacher Qualifications and Its Implications for Student Achievement in High-Poverty Schools." National Bureau of Economic Research Working Paper No. 14021.

Brehm, S. and S. Gates. 1999. *Working, Shirking, and Sabotage: Bureaucratic Response to a Democratic Public*. Ann Arbor: University of Michigan Press.

Campbell, D. E. 2004. "The Civic Implications of Canada's Education System." Pp. 187–212 in P. P. J. Wolf and S. Macedo, eds., *Educating Citizens: International Perspectives on Civic Values and School Choice*. Washington, DC: Brookings.

Charbonneau, E. and G. Van Ryzin. 2011. "Performance Measures and Parental Satisfaction with New York City Schools." *The American Review of Public Administration* 42(1): 54–65.

Chubb, J. and T. Moe. 1990. *Politics, Markets, and America's Schools*. Washington, DC: The Brookings Institution.

Cook, T. D. and D. T. Campbell. 1979. *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Boston: Houghton Mifflin.

Cook, T. D. and M. R. Payne. 2002. "Objecting to the Objections to Using Random Assignment in Educational Research." Pp. 150–178 in F. Mosteller and R. Boruch, eds., *Evidence Matters: Randomized Trials in Education Research*. Washington, DC: Brookings.

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.

Cowen, J. M., D. J. Fleming, J. F. Witte, P. J. Wolf, and B. Kisida. 2013. "School Vouchers and Student Attainment: Evidence from a State-Mandated Study of Milwaukee's Parental Choice Program." *Policy Studies Journal* 41(1): 147–167.

Downs, A. 1964. *Inside Bureaucracy*. Santa Monica, CA: RAND Corporation.

Dronkers, J. and S. Avram. 2010. "A Cross-national Analysis of the Relations of School Choice and Effectiveness Differences Between Private-Dependent and Public Schools." *Educational Research and Evaluation* 16: 151–175.

Dronkers, J. and P. Robert. 2008. "Differences in Scholastic Achievement of Public, Private Government-Dependent, and Private Independent Schools: A Cross-national Analysis." *Educational Policy* 22: 541–577.

Favero, N. and K. J. Meier. 2013. "Evaluating Urban Public Schools: Parents, Teachers, and State Assessments." *Public Administration Review* 73(3): 401–402.

Friedman, M. 1955. "The Role of Government in Education." Pp. 123–144 in R. Solo, ed., *Economics and the Public Interest*. New Brunswick, NJ: Rutgers University Press.

Greene, J. P. 2001. "Vouchers in Charlotte." *Education Matters* 1(2): 55–60.

Greene, J. P., W. G. Howell, and P. E. Peterson. 1998. "Lessons from the Cleveland Scholarship Program." Pp. 357–394 in P. E. Peterson and B. C. Hassel, eds., *Learning from School Choice*. Washington, DC: Brookings.

Greene, J. P., P. E. Peterson, and J. Du. 1999. "Effectiveness of School Choice: The Milwaukee Experiment." *Education and Urban Society* 31(2): 190–213.

Henig, J. R. 1994. *Rethinking School Choice: Limits of the Market Metaphor*. Princeton, NJ: Princeton University Press.

Howell, W. G., P. E. Peterson, P. J. Wolf, and D. E. Campbell. 2006. *The Educational Gap: Vouchers and Urban Schools* revised edition. Washington, DC: Brookings.

Howell, W. G., P. J. Wolf, D. E. Campbell, and P. E. Peterson. 2002. "School Vouchers and Academic Performance: Results from Three Randomized Field Trials." *Journal of Policy Analysis and Management* 21(2): 191–217.

James, O. 2009. "Evaluating the Expectations Disconfirmation and Expectations Anchoring Approaches to Citizen Satisfaction with Local Public Services." *Journal of Public Administration Research and Theory* 19: 107–123.

Kisida, B. and P. J. Wolf. 2010. "School Choice and Information: Does Choice Lead to Better-Informed Parents?" *American Politics Research* 38(5): 783–805.

Krueger, A. B. and P. Zhu. 2004. "Another Look at the New York City School Voucher Experiment." *American Behavioral Scientist* 47(5): 658–698.

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

LaRoque, N. 2005. "School Choice: Lessons from New Zealand." Pp. 105–132 in D. Salisbury and J. Tooley, eds., *What America Can Learn from School Choice in Other Countries*. Washington, DC: Cato.

Levin, H. M. 1998. "Educational Vouchers: Effectiveness, Choice, and Costs." *Journal of Policy Analysis and Management* 17(3): 373–392.

Lippman, L., S. Burns, and E. McArthur. 1996. *Urban Schools: The Challenge of Location and Poverty*. NCES 96–184. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

Lubienski, C. A. and S. T. Lubienski. 2013. *The Public School Advantage: Why Public Schools Outperform Private Schools*. Chicago: University of Chicago Press.

Macedo, S. and P. J. Wolf. 2004. "School Choice, Civic Values, and Problems of Policy Comparison." Pp. 91–106 in P. J. Wolf and S. Macedo, eds., *Evidence Matters: Randomized Trials in Education Research*. Washington, DC: Brookings.

Mayer, D. P., P. E. Peterson, D. E. Myers, C. C. Tuttle, and W. G. Howell. 2002. *Choice in New York City After Three Years: An Evaluation of the School Choice Scholarships Program*. MPR Reference No. 8404–045. Cambridge, MA: Mathematica Policy Research.

McEwan, P. J., M. Urquiola, and E. Vegas. 2008. "School Choice, Stratification, and Information on School Performance: Lessons from Chile." *Economía* 8(2): 1–27.

Metcalf, K. K., S. D. West, N. A. Legan, K. M. Paul, and W. J. Boone. 2003. *Evaluation of the Cleveland Scholarship and Tutoring Grant Program: Summary Report 1998–2001*. Bloomington: Indiana Center for Evaluation.

Mizala, A. and F. Torche. 2012. "Bringing the Schools Back In: The Stratification of Educational Achievement in the Chilean Voucher System." *International Journal of Educational Development* 32(1): 132–144.

Ouchi, W. G. 1980. "Markets, Bureaucracies, and Clans." *Administrative Science Quarterly* 25(1): 129–141.

Poister, T. H. and J. C. Thomas. 2011. "The Effect of Expectations and Expectancy Confirmation/Disconfirmation on Motorists Satisfaction with State Highways." *Journal of Public Administration Research and Theory* 21: 601–617.

Ravitch, D. 2013. *Reign of Error: The Hoax of the Privatization Movement and the Danger to America's Public Schools*. New York: Knopf.

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.

Sapelli, C. 2005. "The Chilean Education Voucher System." Pp. 41–62 in D. Salisbury and J. Tooley, eds., *What America Can Learn from School Choice in Other Countries*. Washington, DC: Cato.

Schneider, M., P. Teske, and M. Marschall. 2000. *Choosing Schools: Consumer Choice and the Quality of American Schools*. Princeton, NJ: Princeton University Press.

Stewart, T. and P. J. Wolf. 2014. *The School Choice Journey: School Vouchers and the Empowerment of Urban Families*. New York: Palgrave Macmillan.

Tufte, E. 2006. *Beautiful Evidence*. Cheshire, CT: Graphics Press.

Van Ryzin, G. G. 2004. "Expectations, Performance, and Citizen Satisfaction with Urban Services." *Journal of Policy Analysis and Management* 23(3): 433–448.

Vermeulen, B. P. 2004. "Regulating School Choice to Promote Civic Values: Constitutional and Political Issues in the Netherlands." Pp. 67–90 in P. J. Wolf and S. Macedo, eds., *Educating Citizens: International Perspectives on Civic Values and School Choice*. Washington, DC: Brookings.

Wilson, J. Q. 1989. *Bureaucracy: What Government Agencies Do and Why They Do It*. New York: Basic Books.

Witte, J. F. 2000. *The Market Approach to Education*. Princeton, NJ: Princeton University Press.

Witte, J. F., P. J. Wolf, J. M. Cowen, D. Carlson, and D. J. Fleming. Forthcoming. "High Stakes Choice: Achivement and Accountability in the Nation's Oldest Urban Voucher Program." *Educational Evaluation and Policy Analysis*.

Witte, J. F., P. J. Wolf, J. M. Cowen, D. J. Fleming, and J. Lucas-McLean. 2008. "MPCP Longitudinal Educational Growth Study Baseline Report." School Choice Demonstration Project Milwaukee Evaluation Report #5. http://www.uaedreform.org/downloads/2008/02/report-5-mpcp-longitudinal-educational-growth-study-baseline-report.pdf.

Wolf, P. J. 2008a. "School Voucher Programs: What the Research Says about Parental School Choice." *Brigham Young University Law Review* 2008(2): 415–446.

Wolf, P. J. 2008b. "Vouchers." Pp. 635–636 in G. McCulloch and D. Crook, eds., *The International Encyclopedia of Education*. London: Routledge.

Wolf, P., B. Gutmann, N. Eissa, M. Puma, and M. Silverberg. 2005. *Evaluation of the DC Opportunity Scholarship Program: First Year Report on Participation*. Washington, DC: U.S. Government Printing Office. http://ies.ed.gov/ncee/.

Wolf, P., B. Gutmann, M. Puma, B. Kisida, L. Rizzo, N. Eissa, M. Carr, and M. Silverberg. 2010. *Evaluation of the DC Opportunity Scholarship Program: Final Report*. NCEE 2007–4009. Washington, DC: U.S. Government Printing Office. http://ies.ed.gov/ncee/.

Wolf, P. J., B. Kisida, B. Gutmann, M. Puma, N. Eissa, and L. Rizzo. 2013. "School Vouchers and Student Outcomes: Experimental Evidence from Washington, DC." *Journal of Policy Analysis and Management* 32: 246–270.

ABOUT THE AUTHORS

Brian Kisida (bkisida@uark.edu) is a Senior Research Associate in the Department of Education Reform at the University of Arkansas. His academic work emphasizes experimental methodologies and includes studies of school governance and school choice, school integration, informal learning, cultural institutions, and art and music education.

Patrick J. Wolf (pwolf@uark.edu) is Distinguished Professor and 21st Century Endowed Chair in School Choice at the University of Arkansas in Fayetteville. His specialty is conducting rigorous empirical evaluations of policy reforms. He has authored, co-authored, or co-edited four books and over 100 journal articles, book chapters, and policy reports on school choice, civic values, public management, special education, and campaign finance. He received his PhD in Political Science from Harvard University in 1995.

ONLINE APPENDIX: TECHNICAL SPECIFICATIONS OF THE RANDOMIZATION PROCESS FOR THE DC CHOICE FIELD EXPERIMENT

for

CUSTOMER SATISFACTION AND EDUCATIONAL OUTCOMES: EXPERIMENTAL IMPACTS OF THE MARKET-BASED DELIVERY OF PUBLIC EDUCATION

BRIAN KISIDA
PATRICK J. WOLF
UNIVERSITY OF ARKANSAS

We operationalized the scholarship lottery as a stratified random sample, also known as "block randomization." Each eligible applicant student was categorized based on the student's grade level (K-5, 6-8, or 9-12) and priority status (attending a public "school in need of improvement" – SINI -- or not). The overall scholarship award probability for each grade-band was first determined by the ratio of eligible applicants to available private school slots for those grades. The subgroup of students attending a SINI school was assigned a proportionally higher award probability and the subgroup of students not attending a SINI school was assigned a proportionally lower award probability. Separate lotteries were run on the various subgroups defined by their classifications into their distinct strata (e.g. K-5 SINI, K-5 non-SINI, 6-8 SINI, 6-8 non-SINI, 9-12 SINI, 9-12 non-SINI), since the lottery award probabilities were only identical for students within each strata. This process was repeated in both 2004 and 2005. It generated a randomized impact sample for the satisfaction study of 2,308 students.[1]

Because certain types of students faced different probabilities of assignment to the treatment group, the outcome data for each student were weighted by the inverse of the likelihood of scholarship award. In other words, the data from control group students from lottery strata with high probabilities of scholarship award were proportionally up-weighted and the data from treatment group students from those same high-award-probability strata were proportionally down-weighted. The reverse adjustment was made for outcome data from strata with relatively low probabilities of scholarship award. While randomizing students within these blocks was required due to provisions of the program legislation and exigencies on the ground, it had the additional benefit of increasing the power of our statistical analysis because similar students were first matched-up into blocks and then these similar students were randomly sorted into the treatment and control groups. Since randomization took place at the student level, in many cases siblings were separated into the treatment and control groups, another practice that increased the analytic power of the experiment.

NOTE

1. For the evaluation of the outcome of educational attainment in the form of high school graduation 4 or 5 years after randomization, only students in the entire 9-12 SINI and non-SINI strata in 2005, the entire 9-12 SINI and non-SINI strata in 2004, and the baseline 8th graders from the SINI and non-SINI strata in 2004 were old enough to be included in the analysis. The eligible student attainment experimental sample totaled 500.

TABLE A1.1.
ITT Effect Sizes for Average School Grade Assigned by Parent

Average School Grade (4 point scale)	*Year 1*	*Year 2*	*Year 3*	*Year 4*
Full sample	0.40***	0.29***	0.22***	0.21***
	(0.06)	(0.06)	(0.06)	(0.06)
N	1,680	1,550	1,410	1,230
Female	0.43***	0.39***	0.22***	0.25***
	(0.07)	(0.07)	(0.08)	(0.09)
N	840	780	720	630
Male	0.37***	0.17**	0.22**	0.17*
	(0.08)	(0.08)	(0.09)	(0.09)
N	830	770	690	590
Higher Performers	0.39***	0.34***	0.28***	0.19**
	(0.06)	(0.07)	(0.07)	(0.08)
N	1,160	1,070	960	830
Lower Performers	0.44***	0.17	0.10	0.25**
	(0.10)	(0.10)	(0.11)	(0.10)
N	520	480	450	400
Grades K-5	0.58***	0.38***	0.25***	0.24***
	(0.08)	(0.07)	(0.08)	(0.08)
N	880	840	790	770
Grades Mid/High	0.23***	0.19**	0.19**	0.19*
	(0.08)	(0.08)	(0.09)	(0.10)
N	800	710	620	460
SINI	0.31***	0.30***	0.16*	0.04
	(0.09)	(0.08)	(0.10)	(0.10)
N	730	680	590	480
Non-SINI	0.48***	0.28***	0.27***	0.33***
	(0.08)	(0.07)	(0.08)	(0.08)
N	950	870	820	750

Notes: Regression estimations use a consistent set of baseline covariates. Robust standard errors in parentheses. Sample weights used. *N*s rounded to the nearest 10 per IES publication policy.

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

TABLE A1.2

Treatment User Effect Sizes for Average School Grade Assigned by Parent

Average School Grade (4 point scale)	*Year 1*	*Year 2*	*Year 3*	*Year 4*
Full sample	0.53***	0.41***	0.34***	0.44***
	(0.07)	(0.08)	(0.09)	(0.13)
N	1,680	1,550	1,410	1,230
Female	0.55***	0.54***	0.31***	0.48***
	(0.09)	(0.10)	(0.11)	(0.17)
N	840	780	720	630
Male	0.51***	0.26**	0.37***	0.39*
	(0.11)	(0.12)	(0.14)	(0.20)
N	830	770	690	590
Higher Performers	0.50***	0.47***	0.41***	0.37**
	(0.08)	(0.09)	(0.10)	(0.15)
N	1160	1070	960	830
Lower Performers	0.61***	0.26*	0.17	0.61**
	(0.14)	(0.16)	(0.19)	(0.26)
N	520	480	450	400
Grades K-5	0.69***	0.49***	0.37***	0.43***
	(0.09)	(0.09)	(0.12)	(0.14)
N	880	840	790	770
Grades Mid/High	0.34***	0.31**	0.31**	0.48*
	(0.12)	(0.13)	(0.14)	(0.26)
N	800	710	620	460
SINI	0.45***	0.35***	0.25*	0.09
	(0.12)	(0.09)	(0.15)	(0.21)
N	730	680	590	480
Non-SINI	0.60***	0.38***	0.41***	0.68***
	(0.09)	(0.10)	(0.12)	(0.16)
N	950	870	820	750

Notes: Instrumental variable estimations are derived from the lottery results and a consistent set of baseline covariates. Robust standard errors in parentheses. Sample weights used. Ns rounded to the nearest 10 per IES publication policy.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2.1
ITT Effect Sizes for School Satisfaction Scale

School Satisfaction Scale	*Year 1*	*Year 2*	*Year 3*	*Year 4*
Full sample	0.39***	0.34***	0.27***	0.27***
	(0.05)	(0.06)	(0.06)	(0.06)
N	1,690	1,570	1,440	1,230
Female	0.49***	0.33***	0.31***	0.30***
	(0.07)	(0.08)	(0.08)	(0.09)
N	840	800	730	640
Male	0.27***	0.35***	0.24***	0.24***
	(0.08)	(0.08)	(0.09)	(0.09)
N	840	780	710	590
Higher Performers	0.37***	0.40***	0.31***	0.29***
	(0.06)	(0.07)	(0.08)	(0.08)
N	1170	1080	980	820
Lower Performers	0.43***	0.22**	0.20**	0.24**
	(0.10)	(0.11)	(0.10)	(0.11)
N	520	500	460	410
Grades K-5	0.53***	0.50***	0.29***	0.27***
	(0.08)	(0.07)	(0.08)	(0.08)
N	880	850	800	700
Grades Mid/High	0.25***	0.18**	0.26***	0.29***
	(0.07)	(0.09)	(0.10)	(0.11)
N	800	720	640	540
SINI	0.36***	0.39***	0.27**	0.20*
	(0.08)	(0.09)	(0.11)	(0.11)
N	740	680	600	480
Non-SINI	0.42***	0.30***	0.30***	0.34***
	(0.07)	(0.07)	(0.08)	(0.08)
N	950	890	840	760

Notes: Regression estimations use a consistent set of baseline covariates. Robust standard errors in parentheses. Sample weights used. Ns rounded to the nearest 10 per IES publication policy.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2.2
Treatment User Effect Sizes for School Satisfaction Scale

School Satisfaction Scale	*Year 1*	*Year 2*	*Year 3*	*Year 4*
Full sample	0.51***	0.49***	0.42***	0.57***
	(0.07)	(0.08)	(0.09)	(0.13)
N	1,690	1,570	1,440	1,230
Female	0.63***	0.46***	0.44***	0.59***
	(0.09)	(0.11)	(0.11)	(0.17)
N	840	800	730	640
Male	0.38***	0.52***	0.40***	0.56***
	(0.10)	(0.12)	(0.15)	(0.21)
N	840	780	710	590
Higher Performers	0.48***	0.55***	0.35**	0.57***
	(0.08)	(0.09)	(0.17)	(0.15)
N	1170	1080	980	820
Lower Performers	0.60***	0.34**	0.17	0.60**
	(0.14)	(0.16)	(0.19)	(0.26)
N	520	500	460	410
Grades K-5	0.63***	0.63***	0.43***	0.50***
	(0.09)	(0.09)	(0.12)	(0.15)
N	880	850	800	700
Grades Mid/High	0.37***	0.30**	0.42***	0.74***
	(0.11)	(0.14)	(0.16)	(0.27)
N	800	720	640	540
SINI	0.51***	0.60***	0.42**	0.43*
	(0.12)	(0.13)	(0.17)	(0.23)
N	740	680	600	480
Non-SINI	0.52***	0.41***	0.46***	0.71***
	(0.09)	(0.10)	(0.12)	(0.17)
N	950	890	840	760

Notes: Instrumental variable estimations are derived from the lottery results and a consistent set of baseline covariates. Robust standard errors in parentheses. Sample weights used. Ns rounded to the nearest 10 per IES publication policy.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A3.1
ITT Effect Sizes for Individual School Characteristics

Satisfaction Scale Items	*Year 1*	*Year 2*	*Year 3*	*Year 4*
Location	0.20***	0.10*	0.16***	0.25***
	(0.06)	(0.06)	(0.06)	(0.06)
	1,660	1,540	1,410	1,200
Safety	0.32***	0.24***	0.25***	0.24***
	(0.05)	(0.06)	(0.06)	(0.06)
	1,650	1,540	1,400	1,210
Class sizes	0.42***	0.30***	0.28***	0.25***
	(0.05)	(0.06)	(0.06)	(0.07)
	1,640	1,540	1,400	1,200
School facilities	0.36***	0.29***	0.19***	0.23***
	(0.06)	(0.06)	(0.06)	(0.06)
	1,630	1,530	1,380	1,190
Respect between teachers and students	0.26***	0.20***	0.22***	0.20***
	(0.05)	(0.06)	(0.06)	(0.07)
	1,640	1,530	1,390	1,210
Teacher and parent communication	0.30***	0.29***	0.18***	0.22***
	(0.05)	(0.06)	(0.06)	(0.06)
	1,640	1,530	1,390	1,200
Observance of religious traditions	0.49***	0.52***	0.41***	0.36***
	(0.06)	(0.06)	(0.06)	(0.07)
	1,490	1,400	1,260	1,080
Parental support for the school	0.29***	0.27***	0.26***	0.31***
	(0.06)	(0.06)	(0.06)	(0.06)
	1,620	1,520	1,380	1,180
Discipline	0.33***	0.25***	0.25***	0.28***
	(0.06)	(0.06)	(0.06)	(0.06)
	1,630	1,520	1,390	1,200
Academic quality	0.37***	0.28***	0.22***	0.23***
	(0.06)	(0.06)	(0.06)	(0.06)
	1,640	1,520	1,380	1,190
Racial mix of students	0.28***	0.19***	0.10	0.12*
	(0.06)	(0.06)	(0.06)	(0.06)
	1,600	1,480	1,350	1,160
Services for students with special needs	0.15**	0.11	0.11	0.06
	(0.07)	(0.07)	(0.08)	(0.08)
	1,000	930	860	730

Notes: Instrumental variable estimations are derived from the lottery results and a consistent set of baseline covariates. Robust standard errors in parentheses. Sample weights used. Ns rounded to the nearest 10 per IES publication policy.

*** p <0.01, ** p <0.05, * p <0.1

Table A3.2
Treatment User Effect Sizes for Individual School Characteristics

Satisfaction Scale Items	*Year 1*	*Year 2*	*Year 3*	*Year 4*
Location	0.26***	0.15*	0.25***	0.51***
	(0.08)	(0.08)	(0.09)	(0.13)
	1,660	1,540	1,410	1,200
Safety	0.42***	0.34***	0.39***	0.51***
	(0.07)	(0.08)	(0.09)	(0.13)
	1,650	1,540	1,400	1,210
Class sizes	0.55***	0.43***	0.44***	0.52***
	(0.07)	(0.08)	(0.10)	(0.13)
	1,640	1,540	1,400	1,200
School facilities	0.48***	0.42***	0.29***	0.47***
	(0.07)	(0.08)	(0.10)	(0.13)
	1,630	1,530	1,380	1,190
Respect between teachers and students	0.34***	0.29***	0.34***	0.42***
	(0.07)	(0.08)	(0.09)	(0.13)
	1,640	1,530	1,390	1,211
Teacher and parent communication	0.39***	0.42***	0.28***	0.45***
	(0.07)	(0.08)	(0.09)	(0.13)
	1,640	1,530	1,390	1,200
Observance of religious traditions	0.63***	0.72***	0.61***	0.71***
	(0.07)	(0.08)	(0.09)	(0.13)
	1,490	1,400	1,260	1,080
Parental support for the school	0.38***	0.38***	0.40***	0.64***
	(0.07)	(0.08)	(0.10)	(0.12)
	1,620	1,520	1,380	1,180
Discipline	0.43***	0.36***	0.38***	0.58***
	(0.07)	(0.08)	(0.09)	(0.13)
	1,630	1,520	1,390	1,200
Academic quality	0.48***	0.40***	0.35***	0.48***
	(0.07)	(0.08)	(0.09)	(0.13)
	1,640	1,520	1,380	1,190
Racial mix of students	0.36***	0.27***	0.15*	0.25*
	(0.07)	(0.08)	(0.09)	(0.13)
	1,600	1,480	1,350	1,160
Services for students with special needs	0.21**	0.18	0.19	0.13
	(0.10)	(0.11)	(0.14)	(0.19)
	1,000	930	860	730

Notes: Instrumental variable estimations are derived from the lottery results and a consistent set of baseline covariates. Robust standard errors in parentheses. Sample weights used. Ns rounded to the nearest 10 per IES publication policy.

*** p <0.01, ** p <0.05, * p <0.1