Undone by the Market?
The Effects of School Vouchers on Educational Inputs*

Kevin Rinz

Department of Economics
University of Notre Dame

Job Market Paper

November 2015

Abstract

By altering the market for private schooling, large-scale school voucher programs may have effects on the educational experience of private school students beyond the effects of small-scale programs. Using eight large, state-level, voucher-style programs adopted between the late 1990s and mid-2000s and a unique dataset on school expenditures and teacher compensation, I estimate the effects of vouchers on educational inputs experienced by students in private school. Large-scale, voucher-style programs alter the inputs students experience in ways that tend to worsen the experience of black students while improving the experience of white students. These effects are driven by changes in inputs deployed at newly established schools. Back-of-the-envelope calculations indicate that the market effects of vouchers are large enough to substantially reduce the benefits of moving from public to private school for black students, reversing more than 100 percent of the gains in student-teacher ratio, 87 percent of the gain in per-teacher compensation, and 51 percent of the gain in instructional hours. My estimates suggest that extrapolation from prior studies may be inappropriate when considering how larger programs affect students.

* I would like to thank Dan Hungerman for his advice and support; Abigail Wozniak, Jim Sullivan, Bill Evans, and seminar participants at the University of Notre Dame for helpful comments; Tim Weninger for advice on crowdsourcing data collection; and the Templeton Foundation and the Institute for Scholarship in the Liberal Arts at the University of Notre Dame for financial support. Contact address: 434 Flanner Hall, Notre Dame, IN 46556. Email: krinz@nd.edu

1. Introduction

For decades, advocates have promoted private school vouchers as a method of increasing parents' power to make decisions about their children's education. While early voucher-style programs were generally geographically limited to particular cities, state-level programs began to proliferate in the late 1990s, and they have become increasingly popular in recent years.[1] Twenty-three such programs were created by states from 2011 through 2015, and total spending on voucher-style programs now exceeds $1 billion annually.

Prior research has focused mainly on the effects of city-level voucher programs and on voucher recipients. Students who have been randomized into voucher receipt are tracked over time and used to estimate effects of vouchers on educational outcomes.[2] While this type of analysis is important, it does not account for broader effects and generally treats private schools as black boxes.[3] Programs that provide subsidies for private education on a large scale, including both traditional vouchers and voucher-style programs that operate through the tax code, may affect not just students who participate in them, but all private school students.[4] For example, these programs may have larger "general equilibrium" or "market" effects on the sorting of students across schools, the price of private education, school entry and exit, or the educational inputs schools deploy. Partial equilibrium estimates may provide a misleading or incomplete picture of the effectiveness of vouchers if market effects are large in magnitude.

My prior work with Hungerman (2015) suggests that there is substantial scope for large-scale voucher-style programs to influence the market for private education. We find that such programs increase private school revenue roughly dollar-for-dollar or more with state spending on them. When programs restrict eligibility to certain groups of students, the revenue increase is driven by increased private school enrollment; when programs offer unrestricted subsidies, the overall increase is driven by an increase in per-student revenue.

Voucher-style programs rely on private schools to improve students' outcomes, both directly and indirectly. Since both student decisions about private school enrollment and school resources change in response to large-scale vouchers, it is reasonable to ask how the educational experience of students in private

---

[1] The first U.S. voucher program was created for Milwaukee, Wisconsin, in 1989. It was followed by a program serving students in Cleveland, Ohio in 1995.

[2] See Witte, Sterr, and Thorne (1995); Greene, et al. (1996); Rouse (1998); Peterson, Howell, and Greene (1999); Wolf, Peterson, and West (2001); Howell, et al. (2002); and Wolf, et al. (2010) for evaluations of programs in U.S. cities. Angrist, Bettinger, and Kremer (2006) evaluate randomly assigned vouchers in several large Colombian cities.

[3] Some theoretical papers include abstract models of how private schools work (Epple and Romano, 1998; Nechyba, 2000; Ferreyra, 2007), but they assume schools' production functions are fixed and school quality varies only with the ability of the students in the school.

[4] Some states provide tax credits to taxpayers for money spent on the private education of their own children. Others provide tax credits for donations to non-profit scholarship-granting organizations. In this paper, I will refer to these tax credit programs and traditional voucher programs collectively as voucher-style programs, or sometimes simply as vouchers. When a statement refers only to traditional vouchers, I will make that explicit.

schools might change as well. The direct impact of these voucher programs is to increase enrollment in private schools. If schools do not change how they deploy those resources in a way that corresponds to changes in enrollment, students might experience, for example, larger class sizes, which could potentially dampen any benefits of attending a private school.

These changes in educational experiences may also help determine whether vouchers achieve the often-stated goal of improving outcomes for disadvantaged students. Private schools have traditionally served a fairly small share of elementary and secondary students, and those students have been largely white and more affluent than the average public school student. However, policymakers often highlight potential benefits to disadvantaged students when advocating voucher programs. Several governors who have supported or enacted voucher programs have made appeals to fairness or emphasized benefits to minority students in promoting the cause, including Jeb Bush, Chris Christie, Mike Pence, and Scott Walker.[5] In many cases, programs are designed with such students in mind. Indeed, 11 of the 23 most recent programs include provisions explicitly restricting eligibility to students from relatively low-income families. The extent to which the support provided by programs like these has found its way to disadvantaged students is unclear, though. In five early-adopting states, non-white students increased from 23 percent of private school students in 2000 to 33 percent in 2013, but gap between the median incomes of private and public school students increased over that period.[6]

In this paper, I use large-scale programs enacted in the United States between the mid-1990s and mid-2000s to estimate the effects of voucher-style programs on the educational inputs experienced by the average private school student. There are several advantages to analyzing effects on educational inputs. Vouchers are meant to improve students' educational experiences, and inputs provide a direct, real-time measure of that experience that can help verify whether these programs have their intended effects on students. Inputs are salient to parents and correspond to schools' decisions about resource allocation. Unlike outcomes such as test scores, data on inputs are available for private schools nationwide, and they are measured consistently across states, years, and school types. Prior research, which will be discussed in Section 3.1, relates various educational inputs to school and later life outcomes, and studying them "opens the black box" that private schools are often treated as.

I obtain data on enrollment, teachers, student-teacher ratios, instructional time, college attendance, and school closures from the Private School Universe Survey (PSS). Unfortunately, this data does not contain important financial information such as school revenues, expenditures, or tuition costs. Thanks to a unique

[5] An excerpt from a 2014 op-ed by Jeb Bush promoting school choice programs is illustrative: "In Florida, I worked with state lawmakers to create scholarships for low-income students and students with learning disabilities… This strategy, combined with strong school accountability, has made Florida a national leader in learning gains by both low-income and disabled students, according to results from the National Assessment of Educational Progress."

[6] According to Census and ACS data, median income of private school students in these five states (Arizona, Florida, Illinois, Iowa, and Pennsylvania) was little changed from 2000 to 2013 (around $89,000), while median income of public school students fell from about $62,500 to $54,500.

dataset of non-profit tax return information matched to PSS records for a subsample of schools, I can supplement the PSS data and examine key outcomes such as per-student expenditures and teacher compensation.

I find that voucher-style school choice programs increase the ratio of students to teachers experienced by the average private school student by about 9 percent. They also increase per-student expenditures by about 6 percent. Changes in other inputs, like instructional hours, the share of graduating students who attend college, and per-teacher compensation are smaller in magnitude.

These results for the average student mask important heterogeneity in the response to vouchers within the private school market. Analyzing these outcomes by race reveals that the increase in student-teacher ratio is driven by the experience of black students, whose increase in student-teacher ratio is nearly 2.5 times that of white students. The small impacts on other educational inputs conceal meaningful effects on the educational experience of black students, including reductions in instructional time, the share of graduating students attending four-year colleges, and per-teacher compensation. The increase in per student expenditures is driven by the experience of white students, while black students see a smaller, insignificant decline in spending. The estimates suggest that the changes in the experience of black students are driven by changes in the educational inputs deployed by newly established schools.

Back-of-the-envelope calculations indicate that the magnitude of these changes in inputs is large enough to substantially reduce the benefits of moving from public to private schooling for black students. For example, the increase in the student-teacher ratio that black private-school students experience due to the introduction of vouchers is greater than the reduction associated with moving the average black student from public to private school. Similarly, the effects of vouchers reverse 87 percent of the gain in per-teacher compensation and 51 percent of the gain in instructional hours. Meanwhile, market effects provide small supplements to the improvement in instructional hours and per-student spending experienced by white students switching to private school. My findings suggest that extrapolation from prior studies of smaller programs may be inappropriate when considering how new, large-scale, voucher-style programs will affect students.

The rest of this paper proceeds as follows. Section 2 provides background information on voucher-style programs. Section 3 discusses the data. Section 4 describes the estimation strategy. Section 5 lays out the results. Section 6 provides discussion and concludes.

2. Background and Program Details

In 1997, Arizona enacted a program that used a tax credit to generate donations to fund scholarships for elementary and secondary school students to attend private schools.[7] The first of its kind in the nation, this program represented a new approach to subsidizing private schooling. In order to receive a credit from the state, taxpayers would make donations to "school tuition organizations," (STOs) non-profits charged with selecting scholarship recipients and distributing money to the private schools they chose to attend. In return, a donor's Arizona income tax liability would be reduced by the amount of the donation, dollar-for-dollar, up to the maximum tax credit of $500 at the time the program was created. Students would then apply to STOs for scholarships and use any money they received to attend a private school in Arizona.

Though mechanically different, Arizona's tax credit is fiscally indistinguishable from a traditional voucher program. Traditional vouchers involve the government collecting money through taxation and then distributing it to particular students for use on private schooling. Under the tax credit approach, the government chooses not to collect money from taxpayers who have provided financial support for the private education of particular students. In both cases, the government uses its taxation and spending authority to provide students with subsidies for private education. When tax credits are awarded dollar-for-dollar, as they are in Arizona's case, the roundabout funding mechanism has no cost to taxpayers, since their opportunity cost of directing financial support to STOs is zero regardless of the marginal tax rate they face.

In the years following the adoption of Arizona's tax credit program, other states began to take action to make subsidies for private schooling available to students.[8] While one state, Florida (in 1999), created a traditional voucher program to provide assistance to students with learning disabilities, most programs established during this period were based on tax credits.[9] Florida (2001) also created a program similar to Arizona's, but with tax credits made available to corporate taxpayers, as did Pennsylvania (2001).[10] Arizona (2006) would subsequently add its own corporate credit. Illinois (1999) adopted a tax credit families could receive in return for spending money on their own children's private education.[11] Iowa both substantially expanded a pre-existing program (1998) similar to the one Illinois established and several years later created a

---

[7] Details of the programs used here were compiled from reports by the Research Department of the Minnesota House of Representatives (2011), the Friedman Foundation (various years), and the Alliance for School Choice (various years), as well as various state sources.

[8] In addition to the states discussed here, the District of Columbia (2004), Ohio (2003 and 2005), Rhode Island (2006) and Utah (2005) also created voucher-style programs during this period. These programs are discussed later in this section.

[9] Note that even though state tax rates are generally relatively low compared to federal rates, state tax credits can still create powerful incentives because they reduce tax liability directly, rather than reducing taxable income, as deductions do.

[10] Under these programs, STOs award scholarships using money donated by corporations, which receive tax credits at rates approaching or equaling dollar-for-dollar with their donations. See Appendix Table 1 for more details.

[11] Expenditures eligible for the Illinois tax credit include those on certain books or lab fees, so parents of students in public schools could also claim the credit in some cases. However, the first $250 spent required to be eligible for the credit.

second tax credit (2006) that directs donations to STOs, similar to those available in Arizona, Florida, and Pennsylvania. As mentioned above, the proliferation of these voucher-style programs continued and accelerated in the late 2000s and early 2010s.

Figure 1 shows spending on and student participation in these voucher-style programs over time. Spending includes the value of tax credits awarded to families who spent money on private schooling for their own children, scholarships money awarded by STOs, and money spent by the state of Florida on the McKay voucher program. The number of students participating is the sum of STO scholarships awarded, McKay vouchers awarded, and tax returns claiming credits for own-child private school spending.[12] A clear increase in utilization and financial support is evident around 2000. By 2011, programs in these five states alone were spending over half a billion dollars to assist more than half a million students annually. Despite their growing importance, I know of no other economic research on these particular programs, with the exception of Hungerman and Rinz (2015), who estimate their effects on private school revenue and enrollment.[13] As mentioned above, the revenue increase we find due to voucher-style programs raises the question of how they might affect the students who attend private schools.

The details of these programs vary from state to state. In addition to using different financing mechanisms (i.e. traditional voucher vs. tax credit for donation to scholarship organizations vs. tax credit for spending on own child), the programs also establish different eligibility requirements for students. Basic requirements that students live and attend school in the state that is offering the subsidy and be within a few years of traditional school age are common to all of these voucher-style programs. Beyond those, some states further restrict eligibility based on a student's family/household income, prior public school attendance, or the presence of a learning or developmental disability.

Some programs (e.g. Arizona's original tax credit) use none of these additional restrictions. In other cases, multiple additional restrictions are applied to a single program. States with multiple programs may differentiate between them in part based on student eligibility. In Florida for example, the traditional voucher program limits eligibility to students with disabilities who attend public school prior to the period for which they are applying to use vouchers. Florida's corporate tax credit program also requires prior public school attendance, but eligibility for that program's scholarships is restricted based on income. Table 1 summarizes variation in timing, financing, student eligibility, and other details of these programs.[14]

Subsequently enacted voucher-style programs have many features in common with those I analyze here. For example, four of the five STO tax credit programs I consider are means-tested. Of the ten STO tax

[12] Since multiple children may benefit from a tax credit claimed on a single return, this figure may slightly understate the number of students assisted by voucher-style programs. However, the limitation of tax credit eligibility to spending on dependent children makes large discrepancies between the count of tax returns claiming the credit and the number of children actually assisted unlikely.

[13] Some state-government reports consider the financial effect of these laws on public school funding. Chan (2006) considers the effect of a similar tax-credit law in Ontario on public school performance.

[14] Appendix Table 1 provides additional details related to these programs.

credit programs enacted after those included in my baseline analysis, eight are means-tested. Many subsequent traditional voucher programs are similar to Florida's McKay program in important ways. Of the 17 traditional voucher programs created after the McKay program, nine are aimed at students with disabilities. Six of those nine programs also require recipients to have previously attended public school in at least some cases. McKay-style voucher programs account for a majority of both traditional voucher programs and programs that provide assistance to students with disabilities. Some of these subsequent programs are utilized in robustness checks, as discussed in Section 5.3.

The movement toward voucher-style programs may coincide with other efforts to reform public education. Arizona began permitting charter schools in 1994, a few years before it created its first STO tax credit program. Florida increased funding for public school construction and maintenance around the same time as it created the McKay program. Pennsylvania created a tax credit to support innovative efforts to improve public schools at the same time as it created its STO tax credit program. While these efforts to improve public schooling could conceivably influence the educational inputs experienced by private school students (e.g. by inducing improvements in private schools through increased competitive pressure), they seem unlikely to lead to the pattern of results presented below.

One might be concerned that changes in educational inputs at private schools may have made them more desirable and led to the creation of subsidy programs, or that some other factor lead to both changes in inputs and the creation of voucher-style programs. Such a story seems unlikely given the discussion above about the legislative and judicial processes that can accompany educational reform. However, I can address such concerns by including a number of controls for a state's socioeconomic circumstances and the population of school-aged children. I also discuss empirical tests of the reverse causality concern below.

One might also wonder what factors determine the timing of when states adopt voucher-style programs. In some cases the timing stems in part from small changes in political circumstances. For example, Florida's programs were created after Jeb Bush, who had narrowly lost a race for governor in 1994, was elected in 1998. In a compromise with the state legislature during state budget negotiations, Arizona governor Janet Napolitano agreed to create that state's corporate tax credit after previously vetoing four similar bills (Welner, 2008). Similarly, Tom Ridge signed Pennsylvania's tax credit law after four failed attempts to get vouchers through the legislature (Averett and Wilkerson, 2001).[15]

State judicial decisions were also instrumental in determining when and whether some of these laws went into effect. State-constitutional clauses prohibiting public financial support for religious organizations, sometimes referred to as "Blaine amendments" after the author of a failed attempt to amend the federal constitution along these lines, have served as important legal obstacles to the adoption of voucher-style

[15] Others states have at various points considered but ultimately not created voucher-style programs, including California, Michigan, Vermont, and Washington.

programs.[16] Additionally, "compelled support clauses" prevent states from forcing residents to support religious ministries. Between these two types of restrictions on state spending, all but three state constitutions include potential barriers to directing state money to private schools, many of which are operated by religious organizations.[17] Multiple programs have faced legal challenges under these clauses. The Illinois program survived two challenges, which produced six different decisions before its legality was settled (Huerta and d'Entremont, 2007). Though the Arizona legislature was not initially sure of the legality of that state's original STO tax credit (Welner, 2008), the courts ultimately affirmed that program, as well. On the other hand, state courts struck down programs in Washington (1986), Vermont (1999), Colorado (2004) and Florida (2006) on various grounds. In other cases, states have attempted to amend their constitutions to clarify the legality of voucher-style programs.[18]

3. Data

As mentioned above, prior research has considered the relationships between a variety of educational inputs and student outcomes. The selection of inputs studied here is informed by that literature, some highlights of which I describe briefly here before detailing the sources of my data.

3.1. Literatures on Educational Inputs

The relationship between educational inputs and student performance has long been a subject of debate. Early work by Hanushek (1986, 1989, 1996, 1997, 1998) suggested little systematic relationship between the inputs and student performance, though others disputed that conclusion.[19] Card and Krueger, for example, found that differences in school quality, as measured by educational inputs, contribute to differences in the returns to additional years of education (1992a), and relative improvement in these inputs at schools attended by black students helped narrow the black-white earnings gap (1992b). More recent work, including quasi-experimental and experimental studies, has identified several inputs that are estimated to influence students' outcomes, sometimes even beyond school. These relationships between inputs and

[16] In 1875, James Blaine, then a congressman from Maine, originally proposed amendment that read, "No State shall make any law respecting an establishment of religion, or prohibiting the free exercise thereof; and no money raised by taxation in any State for the support of public schools, or derived from any public fund therefor, nor any public lands devoted thereto, shall ever be under the control of any religious sect; nor shall any money so raised or lands so devoted be divided between religious sects or denominations."

[17] The three states with neither Blaine amendments nor compelled support clauses are Louisiana, Maine, and North Carolina.

[18] In 2013, North Dakota considered House Concurrent Resolution 3037, which would have removed the Blaine amendment language from the state's constitution. The measure failed in the state's House of Representatives by a vote of 47-47.

[19] For a specific example of this debate dealing with class sizes, see essays by Hanushek and Krueger in Mishel and Rothstein, 2002.

outcomes provide a reason to be interested in how educational inputs might respond to large voucher-style programs.

Evidence from Project STAR indicates that reductions in class sizes can improve students' outcomes both in and out of school. Lower class sizes can increase test scores (Krueger, 1999); make students more likely to complete high school (Finn, Gerber, and Boyd-Zaharias, 2005), take the SAT or ACT (Krueger and Whitmore, 2001), and attend college (Chetty, et al., 2011); reduce the likelihood of arrest, particularly for violent and property crimes (Schanzenbach, 2006b); and reduce teen birth rates (Schanzenbach, 2006b).

Some authors have also used variation created by Project STAR to study the effects a student's peers have on her own performance. As a byproduct of randomizing students into classrooms of different sizes, Project STAR also created a situation in which each student's classmates were also randomly assigned.[20] This allows for estimation of how the sex composition (Schanzenbach, 2006a) and academic ability (Graham, 2008) of a student's peers affect her test scores. In these cases, being assigned to a class composed of more girls or higher-ability students increased test scores. Hoxby (2000) also estimated effects of plausibly random differences in the gender and racial composition of classrooms; she, too, found that students are affected by the quality of their peers. Hanushek, et al. (2003) found that the positive relationship between peer quality and academic achievement survives the inclusion of controls for many potentially confounding factors. Peer effects have been studied in other educational settings as well, including examinations of academic and social outcomes among college freshmen with randomly assigned roommates (Sacerdote, 2001; Zimmerman, 2003; Stinebrickner and Stinebrickner, 2006); disciplinary issues and test scores in classrooms containing male students with traditionally female names (Figlio, 2007); college entrance exam scores among Chinese high school students whose schools and classmates were assigned according to rules based on prior achievement (Ding and Lehrer, 2007); and major choice among college students with partially overlapping peer groups (De Giorgi, Pellizzari, and Redaelli, 2010). These studies also identified important effects of peer behavior on students' own outcomes.

Several quasi-experimental studies have found that instructional time is also positively related to students' academic outcomes. Using variation in days of school induced by snowfall intense enough to necessitate school closures, Marcotte and Hemelt (2007) found that reductions in instructional time lead fewer students to pass state proficiency exams. Similarly, Hansen (2008) found evidence that additional instructional time improves test performance based on both weather-related school cancellations and changes in the timing of test administration. Shorter school years experienced by West German students in 1966-67 (as a result of an effort to align the beginning of the school year across states) increased primary school grade repetition and decreased enrollment in higher secondary school tracks (Pischke, 2007). Matsudaira (2008),

[20] In addition to randomly assigning classmates, Project STAR also randomly assigned teachers. Dee (2004) found that being assigned to a teacher of a student's own race significantly improved math and reading test scores for both white and black students.

using a regression discontinuity design, found that summer school attendance improves math and reading test scores. Lavy (2012) found that the increase in instructional time that arose from a school finance reform in Israel also led to higher test scores.

Like class sizes, the importance of spending per student is continuously debated both academically and politically. While early studies were inconclusive, more recent evidence suggests that increased spending does improve student outcomes. Based on a review of earlier papers and a comparison of North and South Carolina, Card and Krueger (1996) found that greater school resources improved students' educational attainment and earnings. Card and Payne (2002) found that when states narrowed the spending gap between rich and poor districts after courts ruled previous funding schemes unconstitutional, the achievement gap between students in rich and poor districts also narrowed. Gibbons, McNally, and Viarengo (2012) used discontinuities in funding across school district boundaries in England to estimate effects of spending on test scores and find that students at schools that receive more funding perform better on end-of-year national exams. Using an instrumental variables strategy based on the locations of waterfalls in Norway, Haegeland, Raaum, and Salvanes (2012) also found that higher spending (arising from higher revenue from taxes on hydropower plants) leads to higher test scores. Some evidence even points to effects on longer-run outcomes. Using a 1994 change in Michigan's school funding formula, Hyman (2013) found that higher spending per student in elementary school increased college attendance and degree completion. Using school finance reforms across the United States, Jackson, Johnson, and Persico (2014) found increased spending increases completed years of education, increases earnings, and reduces adult poverty for students raised in poor families.

Though it is somewhat different from the inputs discussed above, stability also plays a role in students' educational success. A variety of studies have found negative correlations between either school changes or geographic mobility (which often necessitates a school change) and student achievement (e.g. Ingersoll, Scamman, and Eckerling, 1989; Rumberger and Larson, 1998; Temple and Reynolds, 2000; Engec, 2006; Mehana and Reynolds also provide meta-analysis to this effect). As Hanushek, Kain, and Rivkin (2004) point out, changes in school quality as students move between schools complicate identification of the effects of mobility. I briefly discuss results related to school closures in this context below.

3.2. Data on Educational Inputs

The data on educational inputs used in this paper come from the Private School Universe Survey (PSS) and tax forms filed by subset of private schools. Conducted biennially by the National Center for Education Statistics (NCES), the PSS serves as the foundation of my analysis sample. The survey is designed to capture all schools in the United States that 1) are not primarily supported by public money, 2) offer classroom instruction in at least one grade from kindergarten through 12th grade, and 3) have at least one

teacher. Where financial information is available from tax forms, it is merged onto PSS records, as will be described below.

The sample frame for the PSS is constructed in two parts. The list frame, with produces the vast majority of schools in the overall sample frame, is constructed each year by collecting lists of private schools from outside organizations that interact with them, such as state boards of education, private school associations, religious denominations, or accreditation organizations. Regional census employees construct the second part, the area frame, manually, using local resources to search for private schools in 124 Primary Sampling Units across the country.[21] The sets of schools identified through these two methods supplement the sample frame from the prior round of the PSS.

Schools receive access to the PSS online, and paper copies of the form are delivered by mail, in October at the beginning of the reference school year. Reminders and additional copies of the form are sent out beginning in late October through early December. From January through April, schools that have not yet responded receive telephone and in-person follow-up. For the 2011-12 school year, the response rate for the PSS was 92 percent, which is in line with the survey's historical performance.

The PSS collects information on enrollment, teachers, hours in the school day, days in the school year, and college attendance. Dividing total enrollment in all grades by the number of full-time equivalent teachers gives the student-teacher ratio for each school.[22] Multiplying hours in the school day by days in the school year gives hours in the school year. In select years, schools report the share of their graduates who attend college the fall after graduating, and I utilize this outcome in the results below.[23] The availability of information on enrollment by race beginning in 1993 allows me to calculate the average levels of these educational inputs experienced by members of different races. For this reason, my analysis begins in the 1993-94 school year.

Entry and exit from the PSS can also be used to identify schools that have opened or closed since the last round of the survey. Using all available years of the PSS, I identify the first year a school responds to the survey. Schools are then classified as new in that year. Since the PSS is conducted every other year, this

---

[21] The Census Bureau divides the United States into 2,062 Primary Sampling Units (PSUs), each of which consists of a single county, an independent city, or a group of contiguous counties. The area frame of the PSS includes with certainty the eight PSUs with the highest private school enrollment according to the American Community Survey. The other 116 PSUs are randomly selected.

[22] The 1993 PSS did not ask schools about pre-kindergarten enrollment. Where I can determine that a school did not offer pre-kindergarten instruction, I set pre-K enrollment to zero. In other cases, I impute 1993 pre-kindergarten enrollment at the school level using primarily within-school information. Where pre-K enrollment is available for both 1991 and 1995, I linearly interpolate the ratio of PK-12 enrollment to K-12 enrollment for 1993 and use it to impute 1993 PK-12 enrollment from 1993 K-12 enrollment. When either 1991 or 1995 data are unavailable, I use the average ratio in available years. These two approaches cover the vast majority of cases. Others are imputed using the national average ratio.

[23] Schools were asked about college enrollment among their graduates in 1999, 2001, 2003, 2007, 2009, and 2011 rounds of the PSS. Unfortunately, the survey does not ask schools what share of their students graduated. However, private schools have very high graduation rates on average. According to the 2011-12 Schools and Staffing Survey, 92.4 percent of 12th graders in private schools graduated in 2010-11.

approach identifies schools in their first two years of existence. I identify closed schools as those that do not appear in consecutive rounds of the PSS. In the last year that these schools do appear, they are classified as schools that will close within two years.[24]

The PSS does not contain any financial information about private schools. For that, I turn to data from non-profit organizations' tax returns. Although non-profits do not pay taxes, many are required to provide information about their finances via these forms on an annual basis. The Internal Revenue Service (IRS) uses Forms 990 and 990-EZ to collect that information from 501(c)(3) public charities that normally have gross receipts in excess of $50,000 and are not exempt from filing for other reasons, such as religious affiliation (this exemption in particular will be discussed more later).[25] The information provided includes not only total amounts of money received and spent, but also detailed accountings of each organization's sources of revenue, as well as extensive categorization expenditures.

I am particularly interested in two expenditure measures: total expenditures, which includes all spending an organization does in a year; and compensation expenditures, which includes spending on salaries, benefits, and payroll taxes for all employees. The revenue measure I use is program service revenue, which is revenue an organization receives for providing the service that is the basis for its tax exemption. For schools, this measure corresponds approximately to tuition (indeed, private school tuition is a canonical example of program service revenue), though some other sources of revenue, such as from food service or the sale of school supplies, may also be included. As I will discuss below, schools with unusually high levels of program service revenue per student will be excluded from my analysis of financial inputs.

I obtain data collected from tax forms from three sources. First, the National Center for Charitable Statistics (NCCS) provides access to extracts created annually by the IRS using the most recent available form from each non-profit organization. These files, which NCCS refers to as its "Core Data," are currently available for each year from 1989 through 2013 and serve as the foundation of my financial data, providing names and addresses for all non-profit organizations in all years, as well as measures of total expenditures and program service revenue.

Second, NCCS also provides a set of files containing more detailed financial variables for a limited number of years. These files, which they refer to as "Digitized Data," were created in cooperation with GuideStar, a repository of information on non-profit organizations. They provide information from nearly

[24] This approach to identifying new and closing schools may incorrectly classify schools if they exist for several years prior to being captured by the PSS sample frame or if they do not fill out the questionnaire for an extended period despite still existing.

[25] For most of my analysis period, the gross receipts reporting threshold was $25,000; it increased in 2010. Whether an organization that is required to file submits Form 990 or Form 990-EZ also depends on gross receipts, with those above $250,000 (formerly $100,000, with a higher, transitional threshold in between) filing Form 990. My analysis is unlikely to be affected by these changes in these filing rules, since the vast majority of schools that are in-scope for the PSS should easily exceed the gross receipts threshold in all years, and the financial information I use is available on both the full and EZ versions of Form 990.

every cell of Form 990 for all non-profit organizations for fiscal years 1998 through 2003. From these files, I use information on compensation expenditures.

Finally, since the NCCS Digitized Data files are only available through 2003, I collect information on compensation expenditures for more recent years myself. I obtain PDF copies of schools' tax returns from GuideStar's website and record the relevant information from them using the crowdsourcing website Crowdflower.com.[26] GuideStar provides tax forms from as early as 1998 up through the most recent form each organization has filed, so in combination the NCCS file, I have data on compensation expenditures from 1998 on. Since non-profit tax returns typically were not digitized in full before 1998, it is extremely unlikely that compensation expenditure data from prior to that year can be obtained from any source. Data on total expenditures are available from 1989 on.[27]

Since all types of non-profit organizations, not just schools, file Form 990, I identify schools in the tax data by matching the names and addresses that appear in every year of those data to those reported in the PSS beginning in 2005-06. This matching procedure, the details of which are discussed in the appendix, creates a crosswalk between the Employer Identification Number (EIN) the IRS uses to track non-profit organizations (and, more generally, all other employers) and the internal identifier variable in the PSS. I use that crosswalk to "follow" matched schools back in time from the period in which the match variables are available in both the tax data and the PSS to the beginning of the sample.

Once financial information from the tax data has been matched to PSS data, I create the two financial input measures I analyze here, per-student expenditure and per-teacher compensation.[28] In each case, I divide spending from the tax data by the relevant enrollment or teacher count from the PSS.[29] While the interpretation of per-student expenditure is fairly straightforward, the per-teacher compensation measure may not represent exactly what one might expect at first glance. The numerator consists of spending on employee compensation. This figure includes employees of all kinds, from teachers to administrators to janitorial or maintenance staff.[30] The denominator, however, is a count of full-time-equivalent teachers, a smaller group

---

[26] Details of the data processing on Crowdflower.com are provided in the appendix.

[27] There are several concepts of "year" that are relevant here. As discussed in the appendix, the NCCS uses several notions of year when compiling the data; my use of the term reflects the fiscal year for which a charity files its tax return. The PSS timing is to October of a given year; I thus match NCCS and PSS data together based on the October in which the relevant fiscal year falls. Since most fiscal years are either calendar years or July-to-June fiscal years, this matching procedure should work well.

[28] Some evidence indicates that compensation and teacher quality are positively correlated (Figlio, 1997). Moreover, schools can, when unconstrained by unions (as private schools often are), raise pay to attract more qualified teachers (Figlio, 2002).

[29] In some cases, otherwise acceptable school observations in the tax data do not report a value for compensation expenditures. In these cases, I impute compensation expenditure using the total expenditure value reported in that year and the average of compensation expenditures as a share of total expenditures from the same school in other years. Results that exclude these observations instead of imputing compensation expenditures are similar to those presented here.

[30] It does not necessarily include spending on contract workers, since those workers are not employees of the organization filing the tax return.

than that to which the compensation in the numerator is paid. One might, therefore, think of per-teacher compensation as average teacher compensation measured with error that tends to overstate it. Alternatively, one could treat this measure as the average cost of the combination of a teacher and his share of the bundle of complementary services provided by other school staff members. I prefer the later interpretation.

This combination of tax data with information from the PSS has both limitations and advantages. One important limitation is that not all private schools are required to file Form 990. In particular, schools that are affiliated with churches are not required to file. Catholic schools are especially likely to be associated with churches, so the number of Catholic schools that I am able to match to tax data is small. This may also be a concern for other religious denominations, though I am able to match a fairly large number of non-Catholic, religiously affiliated schools to tax data; about half of matched schools report a religious affiliation. Appendix Table 2 provides more details on the number and types of schools matched in various years. In an average year, my matched sample contains about 4,200 schools and 956,000 students (or, about 14 percent of private schools and 18 percent of private-school students).[31] An additional limitation arises from the fact that none of the variables used to match PSS data with tax data are available in the PSS prior to 2005-06. As such, I cannot match any school that closed prior to that year to tax data, and the construction of the matched sample is biased toward schools that remained open at least until the match variables became available.

These limitations may lead to concern about the representativeness of the matched sample and the generalizability of results based on it. In the context of estimated the effects of voucher-style programs on private school revenue, Hungerman and Rinz (2015) perform several tests to address these issues in a similarly constructed sample from 1991 through 2009, all of which suggest that the matched sample is fairly representative of private schools overall. They find that 70 percent of matched schools can be traced back to 1995, and over half can be traced back to 1991, indicating that attrition is not overwhelming as the sample is constructed back in time. Moreover, the results they present for school revenue are robust to an alternative method of identifying schools in the tax data that does not require matching with the PSS. Additionally, they find that the share of students enrolled at schools that file Form 990 is not responsive to the introduction of voucher-style programs. Although the average matched school has greater enrollment (about 231 students) and more teachers (about 22 full-time equivalent teachers) than the average private school overall (172 students and 12 teachers), trends in both of those measures are similar. See Hungerman and Rinz (2015) for more details of these tests.

Matching financial information from tax returns to data from the PSS also has an important advantage. I know of no other large-scale, annual dataset that provides this type of financial information in as much detail for US private schools as the approach employed here does. The only survey-based dataset I am aware of that contains any financial information on US private schools is the Schools and Staffing Survey.

[31] In an average year, a little over two-thirds of matched observations satisfy the sample restrictions used in my baseline analysis.

However, that survey is conducted only once every four years, only limited financial information is collected, and its sample size for private schools is smaller than what is achievable through matching with tax data.

4. Estimation

The main analysis presented here is conducted at the state level, which allows me to capture effects of voucher-style programs on educational inputs that arise from both within-school changes in resource allocation and market entry or exit by private schools.[32] I create the average level of each input I consider experienced by students in each state-year by taking enrollment-weighted averages of school-level measures of inputs.[33] To create measures of the average inputs experienced by students of different races, I use enrollment by race as the weight in averaging across schools.[34] I use this approach to calculate state-year averages of student-teacher ratio, instructional hours, share of graduating students attending four-year colleges, per-student expenditure, and per-teacher compensation, as well as the share of students attending new schools and schools that will close within two years. Means of these measures are reported in Table 2.

For inputs measures based entirely on PSS data, I use all available schools to calculate state averages. I calculate these averages only for years in which the PSS was actually conducted (i.e. every other year; I do not interpolate off years). Where tax data are required, i.e. for per-student expenditures and per-teacher compensation, I exclude schools at which program service revenue per student (an approximation of tuition) exceeds $25,000 (in 2014 dollars).[35] Some private schools may legitimately receive program service revenue per in excess of this level, but I expect that most observations excluded by this rule were either matched erroneously, or may reflect the returns of parent organizations, such as a hospital or university, that offer education or daycare to children. For these financial inputs, I use data from all years, including those in which no PSS was conducted, in order to take advantage of all available variation in the tax data. I use linearly interpolated enrollment and teacher counts to create per-student expenditure and per-teacher compensation measures in PSS off years.

Alaska and Hawaii are not included in my analysis. For states that adopted voucher-style programs I do not analyze (either because their scope was limited or they first appear late in the period covered by my

[32] School-level regressions with school fixed effects would miss potentially important changes in inputs due to market entry/exit by identifying effects using only within-school variation.
[33] In order to also incorporate PSS sample weights, the enrollment measure used to create these weighted averages is the product of enrollment in all grades (pre-kindergarten through 12) and the PSS sample weight.
[34] To provide a specific example, to create my measure of the student-teacher ratio experienced by the average student overall, I calculate the student-teacher ratio at each school, then take the weighted average across schools within each state-year, using total enrollment as each school's weight. When I create my measure of the student-teacher ratio experienced by the average black student, each school's weight is instead the number of black students enrolled. Since the PSS only reports enrollment by race for kindergarten through 12th grade, I use shares of K-12 enrollment by race to impute total PK-12 enrollment by race. In doing so, I implicitly assume that each school's pre-kindergarten grades have the same racial composition as its K-12 grades.
[35] I also exclude schools reporting per-student program service revenue less than or equal to zero.

data), I begin excluding their data in the year their first programs become available.[36] I also exclude from my analysis Ohio and Wisconsin, the states that were home to early city-level traditional voucher programs.[37] Washington, DC is excluded both to avoid comparing a city with states and because it, too, was home to a city-level traditional voucher program, created in 2004. These restrictions leave me with 46 states in my analysis sample. I use data from 1993-2011.

My baseline estimating equation is

$$y_{sy} = index_{sy}\alpha + X_{sy}\beta + \theta_s + \vartheta_y + T_{sy}\tau_s + \varepsilon_{sy} \quad (1)$$

where y_{sy} is the average educational input experienced by students in state *s* in year *y*, $index_{sy}$ counts the number of voucher-style programs in effect in each state-year (more details below), and X_{sy} is a matrix of controls for state characteristics, including median income, the unemployment rate, population density, share of residents who are non-citizens, share of residents born abroad, and shares of the population in several race/ethnicity by age bins.[38] State controls are obtained from the Bureau of Labor Statistics, the Census, and the American Community Survey (see the appendix for more details). State fixed effects are represented by θ_s, year fixed effects by ϑ_y, and state-specific linear time trends by $T_{sy}\tau_s$, where T_{sy} contains time variables for each state and τ_s contains their coefficients.[39] The ε_{sy} term is noise.

The index variable provides a simple count of the number of distinct voucher-style programs in effect in each state-year. Where no such program exists, its value is zero. When a state's first program becomes available, the value of *index* increases to one. If a state creates a second, distinct voucher-style program (e.g. creating an STO-style tax credit program in addition to a program that provides tax credits for a household's spending on its own children, as Iowa did), its value of *index* increases to two. In practice, *index* is equal to two in Arizona beginning in 2006, Florida beginning in 2001, and Iowa beginning in 2006.[40] While this approach is somewhat blunt (it does not account for differences in program size across states or over

[36] State with data excluded for these reasons include Utah (beginning in 2005), Georgia (2007), Rhode Island (2007), Louisiana (2008), Indiana (2009), Oklahoma (2010), and Colorado (2011)

[37] I sacrifice some policy variation by making this exclusion. Ohio created state-level, voucher-style programs in 2003 and 2005. Wisconsin expanded the Milwaukee program to Racine in 2011 before taking it statewide in 2013, too late for inclusion in the data used here.

[38] The race/ethnicity groups used to construct these bins are white, black, Hispanic, and other. The age groups are 6 to 10 years, 11 to 13 years, and 14 to 17 years. The share of each state's total population that falls into each of the twelve groups created by crossing these race/ethnicity and age categories with each other is included as a regressor.

[39] Quadratic time trends are included as a robustness check below.

[40] In 2009, Arizona passed "Lexie's Law," which used a corporate tax credit to provide scholarships specifically for students with disabilities through STOs. Since this program built on Arizona's existing corporate tax credit and STO infrastructure, includes a student eligibility restriction, and involved a relatively small increase in state spending on tax credits (up to $5 million annually), I treat it as growth in the existing program, and I do not increase the value of *index* when it is created. In 2011, the last year covered by my data, Arizona created a substantively different type of voucher-style program, which gives recipient families who withdraw their students from public school access to money in "Empowerment Scholarship Accounts," which they can spend on a variety of educational services, including private schooling, online education, and tutoring. Because this program differs substantially from the prior Arizona programs, I increase the value of *index* to three in Arizona in 2011, and I set the dummy variable that indicates the first year of a new program to one in that state-year. Estimates are little changed if I instead omit this state year from my analysis.

time, for example), it represents a sensible way to begin analyzing these programs, particularly during this period in which the relatively small number of available policy changes limits the extent to which the importance of various program details can be determined. Below, I discuss an alternative approach to estimation based on the amounts of money states spend on their voucher-style programs, which does account for differences in program size. Estimates from that approach are largely consistent with those based on the *index* specification, suggesting that this baselines analysis does not obscure too much by its simplicity.

In some cases, there is some ambiguity about when a new voucher-style program would be expected to first have an impact. Programs may be limited in scope at their outset. STOs may need time to accumulate donations before awarding scholarships. Tax credits for spending on a household's own children operate on a calendar year basis, the academic year ending during the first calendar year in which these credits are available may or may not be affected. With a few exceptions, *index* is coded as increasing in value in the year in which new voucher-style programs are enacted. To account for the uncertainty just described, I also include a dummy variable that is equal to one in state-years in which new programs are adopted and equal to zero otherwise.

There are three cases in which *index* is not coded as described above. Iowa's own-child tax credit was initially created in 1987, before my data begin, and its original parameters limited both the value of the credit and eligibility for it. Subsequent changes made the program more generous on both of those fronts; spending on it increased by a factor of three to $9 million from 1997 to 1998.[41] For this analysis, I code *index* as increasing to one in Iowa in 1998. In Arizona, the original tax credit program was enacted in 1997, but donations were not eligible for tax credits until 1998. Therefore, I code that program as beginning in 1998. Finally, the McKay voucher program was initially created as a pilot program in 1999, and use was very limited in its early years. Use expanded dramatically in 2001, so I code the McKay program as beginning in that year.[42] Altering the timing of these laws to their earlier enactments, or simply dropping the year of enactment and the year following enactment from the regressions for all states, yields similar estimates to those shown here.

This specification can be interpreted as a "difference-in-difference" estimate. However, the relatively small number of treatment states (five) presents a situation in which estimates may be subject to finite-sample bias, potentially complicating inference. Conley and Taber (2011) propose a bootstrap technique to address this issue. In the results that follow, I both report conventional state-clustered standard errors, as well as 95-percent confidence intervals, using an approach based on Conley and Taber's method. I discuss this approach briefly here.

I begin by using the full sample to estimate, via OLS, the equation

[41] See Rogers and Gullickson's 2012 evaluation of the program for the Iowa Department of Revenue.
[42] Florida's corporate tax credit for donations to STOs was also created in 2001, so Florida's value of *index* increases from zero in 2000 to two in 2001.

$$y_{sy} = X_{sy}\beta + \theta_s + \vartheta_y + T_{sy}\tau_s + \varepsilon_{sy} \quad (2)$$

which is identical to the main estimating equation except that $index_{sy}$ has been omitted. From this estimation, I retain the residuals $\tilde{\varepsilon}_{sy}$. Then, I randomly reassign the treatment states' index values to control states, with assignment done at the state level (e.g., the index values for all years in Iowa are assigned to Idaho). After this reassignment, I regress the new-states' residuals $\tilde{\varepsilon}_{sy}$ on the state-index profiles, and I record the coefficient. I repeat this process 5,000 times to estimate a distribution of coefficients that "should" be zero, but may be nonzero or non-symmetric due to finite-sample bias. Under the null hypothesis that α, the coefficient of interest in the actual estimating equation, is zero, I can "invert" this distribution of coefficients from the placebo regressions to construct a 95-percent confidence interval that takes the finite sample bias into account.[43] Many of the key results presented below are similarly precise either using standard errors or the Conley-Taber confidence intervals.

5. Results

As a preliminary matter, Table 3 reports the change in private school enrollment due to voucher-style programs. Although a net change in enrollment is not strictly necessary in order for vouchers to affect the educational inputs students experience, evidence that vouchers do influence an outcome like enrollment indicates that it's reasonable to think that they might have broader effects. In this setting, voucher increase enrollment in private schools by about 7,500 students (or a little over 6 percent of enrollment) in the average state. Enrollment of both white and black students increases, with white students accounting for about half of the overall increase. The change in black enrollment is estimated precisely using both clustered standard errors and Conley-Taber confidence intervals. Since white students make up about 75 percent of private school enrollment overall, vouchers reduce the share of private school students who are white slightly.

5.1. Main Results

Effects of vouchers on the average private school student's educational inputs are reported in Table 4. The largest change due to vouchers is seen in the student-teacher ratio, which increases by about 1.9 students per teacher (about 9 percent of the overall average). The average student also experiences an increase in per-student expenditure of about $750 (about 6 percent of the mean). The changes in the share of the average private school student's graduating peers who attend four-year colleges and per-teacher expenditure

[43] Conley and Taber (2011) propose two asymptotically equivalent approaches to constructing confidence intervals; one includes truly treated states in the pool to which the treatment profiles may be reassigned, and the other excludes them. It is not clear from their paper whether the repeated regressions of the first stage residuals on the reassigned treatment profiles should include state fixed effects, which makes the resulting confidence intervals wider. To be conservative, I do include state fixed effects in the replications I use to produce confidence intervals reported here. I also include truly treated states in the pool to which treatment profiles may be reassigned. Excluding those states from that pool does not change the statistical significance determination for any of the results reported here.

are statistically significant, but smaller in magnitude, and the change in instructional hours is small and not statistically significant. Taken together, these changes do not indicate a clear change in the quality of the average student's educational experience. Research associates larger class sizes with worse academic performance, but also higher per-student spending with better performance, and the changes in those two outcomes are of comparable magnitudes.

Estimating effects on educational inputs by race reveals important heterogeneity that evaluating the experience of the average private school student overall masks. Table 5 re-estimates the effects of voucher-style programs for white and black students separately. This analysis reveals that black students largely experience changes that tend to worsen the private school experience, while white students experience changes that tend to improve it.

Although the student-teacher ratio experienced by white students does increase when voucher-style programs are introduced, it increases only by about 1.0 student per teacher. Black students, on the other hand, see the student-teacher ratio they experience increase by more than 2.5 students per teacher. The difference between these two estimates is not statistically significant, but their relative magnitudes suggest that the change in the experience of the average student overall with respect to the student-teacher ratio is driven by the experience of black students.

Changes in per-student expenditures more clearly suggest differential changes in experience across races in response to voucher-style programs. For white students, spending increases by $723 per student. The change for black students is not statistically significant, but the point estimate is negative, suggesting that spending falls by $521 per student. The Conley-Taber confidence intervals for these estimates do not overlap, suggesting differential responses across races, though a Wald test based on the less-precise clustered standard errors of the two estimates fails to reject the null hypothesis that they are equal.

The small effects on other inputs estimated for the average private school student overall conceal larger, detrimental effects on black students. Black students received 26.5 fewer hours of instruction and saw the share of their peers who go on to attend four-year colleges decline by 6.2 percentage points after the introduction of voucher-style programs. The average compensation of their teachers also declined, by nearly $11,000. As shares of the average level of these inputs experienced by black students across states, these changes range from 2 percent (instructional hours) to 9 percent (peer college-going and teacher compensation), and each is statistically significant. The changes in these inputs experienced by white students are also statistically significant, but they are much smaller in magnitude and tend to be in the direction that tends to benefit students.[44] Per-teacher compensation and peer college-going see small but statistically significant increases for the average white student, while the change in instructional hours is positive but not statistically significant.

[44] For instructional hours, peer college-going, and per-teacher compensation, Wald tests indicate that the differences between the effects experienced by white and black students are statistically significant.

5.2. Margins of Adjustment to Voucher-Style Programs

There are several possible margins on which schools or students might adjust in response to vouchers. Schools' decisions about the ways that they allocate their resources, the distribution of students across schools, or the rates at which schools enter or exit the market for private education could all change. To assess the importance of these various margins of adjustment, I combine the state-level analysis just discussed with school level analysis to introduce them one at a time.

This decomposition analysis begins with school-level regressions that include the same state-level controls as the baseline analysis described above, except that they include school fixed effects instead of state fixed effects.

$$y_{isy} = index_{sy}\alpha + X_{sy}\beta + \varphi_i + \vartheta_y + T_{sy}\tau_s + \varepsilon_{isy} \quad (3)$$

These regressions are weighted using only PSS sample weights. To the extent that these regressions identify effects of voucher-style programs on educational inputs, those effects are due to within-school changes in resources.

When that same set of school-level regressions is instead weighted by school enrollment (again, using enrollment specific to the group of students being analyzed), they capture effects due to the combination of within-school decisions about resources and the distribution of students across schools. Changes in resources at schools with more students have larger effects on the levels of educational inputs students experience on average. Also, voucher-style programs may alter the inputs students experience by changing the distribution of students across schools. Moreover, to the extent that students at private schools that close subsequently attend other, pre-existing private schools, these regressions will also capture effects on inputs due to school closure. Weighting by enrollment incorporates these factors.

Finally, the state-level analysis captures the same changes as the enrollment-weighted school-level regressions, but also incorporates results from changes in the composition of schools due to school entry and exit. Comparing estimates of the effects of vouchers on educational inputs across these three sets of regressions provides an indication of which margins of adjustment are the most important.

Beginning again with the average private school student overall, shown in Figure 2, this analysis shows an important role for the market entry margin. Consider the student-teacher ratio. The left bar in the figure, which is based on a regression that includes school fixed effects and employs PSS sample weights, shows a reduction in the student-ratio of 0.372 students per teacher due to vouchers. In other words, the average pre-existing school reduced its student-teacher ratio slightly in response to vouchers. When schools in that regression are instead weighted by enrollment, as shown in the middle bar, the student-teacher ratio declines by more in response to vouchers, falling by 0.964 students per teacher. Although the difference between the left and middle bars is not statistically significant, it is instructive to state in concrete terms how switching to enrollment weights might affect the estimated effects of voucher-style programs: vouchers may

have increased enrollment at pre-existing schools with lower student-teacher ratios, or the schools that reduced their student-teacher ratios in response to vouchers may have had larger enrollments on average than those that did not.

However, the bar on the right, which represents the estimate from the regression that adds in changes in the student-teacher ratio due to school entry, shows the student-teacher ratio increasing by 1.929 students per teacher in response to voucher-style programs. That estimate is opposite in sign to, double the magnitude of, and statistically significantly different from the effect shown in the middle bar, which comes from a regression that utilizes within-school variation in resource allocation and enrollment, as well as differences in enrollment across schools. Since the estimate shown in the right bar is based on not only that same variation and but also variation due to school entry, the difference between the middle and right bars suggests that students attending schools that enter the market as a result of vouchers experience substantially higher student-teacher ratios than those at pre-existing schools.

The entry margin also contributes substantially to the overall effects of vouchers on the per-student expenditure and per-teacher compensation experienced by the average private school student, though within-school changes already lead to improvements in these inputs once enrollment weights are applied. The change in peer college-going is due largely to within-school, enrollment-weighted changes, with little contribution from the entry margin.

Considering the average private school student does not conceal the importance of the entry margin, but the direction of that importance again differs across races. Although the increase in the student-teacher ratio experienced by white students is due to school entry, that margin enhances the within-school changes they experience in peer college attendance, per-student expenditures, and per-teacher compensation (Figure 3). Black students (Figure 4) instead see school entry intensify detrimental effects on the student-teacher ratio, instructional hours, and peer college attendance, while more than reversing within-school improvements in per-student expenditure and per-teacher compensation.

This decomposition demonstrates the importance of considering the effects of vouchers on the market as a whole, and suggests that the school entry margin is a major contributor to the overall market effects, but it does not completely address how that margin is affecting educational inputs. If new schools are typically worse (or better) than existing schools, vouchers could affect the inputs students experience through school entry by changing the likelihood that students attend new schools. If vouchers affect the likelihood of attending new schools differentially across races, that could give rise to differential effects of vouchers on inputs. Alternatively, vouchers could change the inputs deployed by new schools, with overall effects on inputs coming from differences between the average experience of new-school students before and after vouchers, rather than from differences in the number or composition of new-school students.

To address these possibilities, I estimate the effects of vouchers on inputs deployed at new schools, as well as their effect on the likelihood of attending a new school. The pattern of changes to inputs in

response to vouchers at new schools, reported in Table 6, is similar to the overall pattern, but the magnitudes are often much larger, particularly for black students.[45] Student-teacher ratios rise for both races, but point estimates indicate that changes in instructional hours and peer college-going are beneficial to white students and detrimental to black students.[46]

Vouchers also have differential impacts on the likelihood of attending a new school across races. As reported in Table 7, the introduction of a voucher-style program makes white students nearly 2.6 percentage points less likely to attend a new school, on a base of about 9 percent. There is little change in black students' likelihood of attending new schools. The change in the probability that a black student attends a new school is not statistically significant, though the point estimate indicates that that probability increases by nearly 7 percent (0.8 percentage points on a base of 12.4).

For some inputs, the average experience of white and black students in new schools, shown in Table 6, is worse than the average experience overall (Table 5), and voucher-style programs may affect the average educational experience in part by exposing fewer students to these input levels. The reduction in the probability of white students attending new schools could contribute to the increase in peer college-going, for example. But new schools are not worse than existing schools across the board, and the biggest changes in inputs are experienced by black students, who see a smaller, not significant change in their likelihood of attending new schools. On the other hand, changes in the inputs students experience at new schools are consistently in line with the changes they experience overall. This indicates that vouchers' effects on the entry margin are due primarily to changes they cause in the inputs deployed by new schools.

Vouchers also make students of both races less likely to attend schools that will close within two years. This effect is especially large for black students (a reduction of 6.9 percentage points on a base of 13.5 percent). However, this effect is unlikely to be a significant contributor to the worsening in inputs experienced by black students. Black students at schools that will close within two years experience a higher student-teacher ratio, less peer college-going, and lower per-student expenditure than black private school students overall. Moving away from schools that will soon close would improve these inputs on average, working against the negative effects I find. Black students do experience more instructional hours and per-teacher compensation at soon-to-close schools, so movement out of these schools may contribute to some extent to the negative effects on these estimates I find for black students. However, these contributions are likely to be small; the instructional hours and per-teacher compensation black students experience at soon-to-close schools are only about three percent higher than the average level black private school students experience overall.

[45] This analysis excludes financial outcomes, since only a handful of matched observations are available per state-year on average.
[46] These effects are also evident in school-level analysis that excludes new schools. The magnitudes of changes in inputs decline substantially when new schools are excluded. See Appendix Table 3.

Changes in the likelihood of attending new or closing schools could arise from changes in the rate of school entry and exit or from changes in the distribution of students across schools. Table 8 reports the effects of vouchers on rates of school entry and exit. Point estimates indicate that vouchers reduce the rate of school entry by 1.6 percentage points (on a base of 19.5 percent) and the rate of school exit by 1.0 percentage point (on a base of 19.8 percent), though neither effect is statistically significant at the five-percent level.[47] These estimates suggest that changes in the distribution of students across schools is likely more important than changes in rates of school entry and exit. For example, although there is substantial overlap between their confidence intervals, the reduction in the likelihood of a black student attending a school that would go on to close within two years is more than six times as large as the reduction in the rate at which schools close. The point estimate for black students also indicates that they become more likely to attend new schools, even without an increase in the rate at which schools open.

Aside from any contributions school exit may make to changes in inputs, it can also be thought of as an educational input in its own right. Research mentioned above indicates that disruptions in schooling, e.g. moving to a new school or neighborhood, are negatively correlated with educational performance. One element of a school closure is a forced change in schooling for its students. All else equal, fewer school closures lead to fewer educational disruptions. Although other estimates presented here indicate that all else is not, in fact, equal, the reduction in the likelihood of attending schools that go on to close within two years could be seen, in part, as an improvement in educational stability.

5.3. Robustness Checks

The results from Tables 4 and 5 for the five main outcomes are robust to a variety of alternative specifications. In Tables 9-13, I estimate models using student-teacher ratios, instructional hours, the share of graduating students who attend four-year colleges, per-student expenditures, and per-teacher compensation, but I change the basic specification by omitting time-varying state controls, including demographic controls in levels rather than shares, and using quadratic rather than linear trends, produce similar results, both in terms of magnitudes of effects and comparing estimates for white students to those for black students.[48]

---

[47] The school entry effect is significant at the ten-percent level under the Conley-Taber bootstrap method. The base rates of school entry and exit reported here may appear unusually high. This is due to the construction of the PSS's area frame, which samples different cities each year and includes schools that are identified as new by my method. Using only schools from the more stable list frame, the baseline rate of entry is 10.3 percent, and the baseline rate of exit is 10.7 percent. Estimates of the effects of voucher-style programs on school entry and exit produced using only schools in the list frame are not statistically significantly different from those presented here.

[48] When state-specific trends are omitted altogether, the magnitude of my estimates decline, and they generally become statistically insignificant, though the pattern of point estimates described here is largely preserved. Estimates similar to those presented in this paper can be produced without the use of state trends by restricting the analysis to years more closely surrounding (year of enactment, plus or minus five years) each treatment state's first policy change. That this alternative approach preserves the original effects is reassuring because it suggests that the effects do not arise from the excluded variation, which is more distant in time from the policy changes in question.

Although the five states included in my baseline analysis are fairly geographically and politically diverse, one might still be concerned that they are different in some important way from states that do not create voucher-style programs. To address this concern, I estimate a specification in which the control group is limited to states that subsequently adopted voucher-style programs.[49] If these estimates were substantially different from the baseline estimates, one might be concerned that the baseline estimates were driven by other aspects of states' private school markets that are related to the eventual creation of voucher-style programs. However, if the two sets of estimates are similar, they are more likely to driven by the actual creation of voucher-style programs, and those concerns would be lessened. Again, both the magnitudes and the pattern of estimates across races are similar to the baseline results in this specification.

Programs in three states (Georgia, Rhode Island, and Utah) were excluded from the baseline analysis on the basis of being either too small or adopted too late in the period covered by my data. Rhode Island and Utah impose very low limits on the amount of money their programs can spend.[50] Georgia's STO tax credit, created in 2008, saw limited utilization in its early years, and only three years of post-policy data are available for the Georgia voucher program, which was created in 2007. I repeat the analysis above with these marginal programs included in the treatment group. As one might expect after the inclusion of smaller, less-established programs, the magnitudes of the effects of vouchers on educational inputs shrink. However, most estimates that were statistically significant in the baseline analysis remain so in this specification, and the inputs experienced by black students continue to be differentially, adversely affected by voucher-style programs.

As a check for changes in input trends prior to the adoption of voucher-style programs or the potential reverse causality discussed above, I estimate effects of voucher-style programs on educational inputs with pre-law change placebo dummies included. If estimates for these placebo variables show effects similar to those estimated for the *index* variable in the baseline specification, one might be concerned that changes are not due to the policy change. For non-financial inputs, which are observed only every other year, I include a dummy that is equal to one for the observations that are either one or two years prior to the policy change, and another that is equal to one for observations three or four years prior to the policy change. For financial inputs, which are observed every year, I include such dummy variables for each of the five years prior to the policy change. As Appendix Tables 3a and 3b show, the placebo dummies do not show changes in inputs prior to policy adoption that would cast doubt upon the baseline results.

---

[49] These states include Alabama, Colorado, Georgia, Indiana, Kansas, Louisiana, Mississippi, Nevada, New Hampshire, North Carolina, Oklahoma, Rhode Island, South Carolina, Utah, and Virginia. Also included are states that have long pre-existing histories of offering smaller-scale support for private school attendance: Maine and Vermont provide assistance to students living in towns without their own schools, while Minnesota has long offered a tax deduction for private school expenditures and a means-tested tax credit for non-tuition educational expenses.

[50] Rhode Island initially spent no more than $1 million annually on its program, which it created in 2007; it has subsequently increased the limit to $1.5 million. Utah originally allocated $1.5 million to its program, which was established in 2005; it has since increased that amount to $3.75 million.

The index variable used in the baseline specification is a somewhat blunt instrument. More important perhaps than the differences between programs discussed earlier, states spend different amounts of money on them. I estimate a specification that uses the amount of money spent on voucher-style programs in place of the index variable from the baseline analysis to get a more continuous estimate of how inputs change in response to these programs. To remove endogeneity associated with the amount states actually spend on vouchers, I use a measure of the maximum amounts they are willing to spend as an instrument.[51] In Table 14, I report IV estimates of the effects of vouchers on inputs, scaled up by average spending per year in states with voucher-style programs (nearly $60 million) for ease of interpretation. With the exception of estimates for instructional hours, both the magnitudes and pattern of IV estimates across races are again consistent with the baseline analysis.

5.4. Comparing Estimates to the Public-Private Difference in Inputs

Putting the estimates presented above in the context of the differences between educational inputs experienced in public school vs. private school reveals that the changes vouchers cause within the private school market are large and important for black students. Table 15 compares point estimates from my baseline analysis to the differences between average levels of educational inputs in public and private school, by race. The increase in the student-teacher ratio that black students experience due to voucher-style programs is large enough to more than reverse the reduction they would experience by moving from public to private school, under average circumstances.

Recent estimates of the effects of class size on income can also be used to characterize the magnitude of this effect on black students. Rescaling Chetty et al. (2011)'s estimates suggests that the roughly 12.5 percent increase in student-teacher ratio that black students experience due to vouchers costs each student about $140 in income at age 27. This translates into a lifetime earnings reduction of nearly $3,600 in present value per student or nearly $72,000 for a class of 20 students.[52]

The market effects of voucher-style programs also reverse half the gain from switching in terms of instructional hours and nearly 90 percent of the increase in per-teacher compensation for black students. Overall, these estimates indicate that the market effects of voucher-style programs worsen the experience of switching from public to private school for black students while enhancing that experience for white students.

[51] Where legislated maximums are available, I use them as my instrument. Where program spending is not capped statutorily, I use measures based on the eligible population in the year prior to program adoption and, in some cases, nationwide growth in eligible populations, to impute the maximum amounts states are willing to spend. See the appendix for more details. In a regression of actual state voucher spending on my instrument using my main analysis sample, the instrument has an F-statistic of 17.99.

[52] In addition to the various strong assumptions imposed by Chetty et. al to produce their original estimates, these rescaled estimates also assume that the income effect of class sizes is linear in the change in class size, and that the effects of increases and decreases in class size are symmetric.

6. Discussion and Conclusion

Early voucher programs in the United States were introduced at the city-level, and analysis of those programs focused primarily on students who received vouchers. Since the late 1990s, however, legislators have increasingly focused on creating larger, state-level, voucher-style programs that have the potential to affect not just participating students but the market for private education as a whole. These market effects may alter the ability of voucher-style programs to improve students' educational experience overall and equity in that experience across groups of students. Earlier studies of smaller programs do not account for these potential market effects of vouchers and so might provide a misleading or incomplete picture of how larger programs may affect students.

I evaluate the effects of voucher-style programs on the educational experience of private school students, both overall and by race, in terms of the educational inputs they experience. These inputs include student-teacher ratios, instructional hours, college attendance among graduating peers, per-student expenditures, and per-teacher compensation. Measures of educational inputs are collected from a combination of the Private School Universe Survey and a unique dataset of financial information collected from non-profit tax returns.

I find that voucher-style programs do affect the educational inputs that private school students experience, and that these effects differ across races. Large voucher-style programs tend to slightly improve the experience of white students while worsening the experience of black students. Although they experience an increase in the student-teacher ratio they experience due to vouchers, white students see small improvements in other inputs. Black students, on the other hand, see an increase in the student-teacher ratio that is about 2.5 times as large as the increase for white students, along with a decline in instructional hours, fewer peers attending four-year colleges, and lower per-teacher compensation. Comparisons of estimates based on different school and state level analyses indicate that the worsening in educational inputs experienced by black students is due largely to changes in the inputs they experience at newly established schools.

The question of why voucher-style programs might lead the inputs deployed by new private schools to worsen remains to be addressed. A few possibilities come to mind. Students who are induced to (re)consider private schooling by the availability of voucher-style programs may be more sensitive to price than those who chose to attend private schools in the absence of such programs. The availability of a larger pool of price-sensitive potential students could lead new schools to cut costs, potentially by deploying worse combinations of educational inputs, in order to appeal to them. Alternatively, some voucher-style programs require participating schools to abide by testing, reporting, or hiring standards, which could lead new schools to direct resources away from educational inputs.

Why would these channels affect new schools but not existing schools? Students likely have limited ability to evaluate the quality of new schools, while it is relatively easy to determine the quality of existing schools. This could make it difficult for existing schools to cut costs or change strategies without endangering their positions in the market. Also, existing schools may have some market power (possibly arising from the costs to their students of switching schools) that enables them to address additional burdens without sacrificing quality.

Comparing the effects estimated here to differences in inputs between public and private schools illustrates the magnitude and significance of the changes black students experience due to vouchers. Once market effects are incorporated, voucher-style programs reverse more than 100 percent of the gain in student teacher ratio, half the gain in instructional hours, and nearly 90 percent of the gain in per-teacher compensation that black students would experience by moving from public to private school, under average circumstances. White students, on the other hand, see the experience of moving from public to private school slightly enhanced by vouchers. Both the presence of significant market effects and their differential impact across races suggest that extrapolation from prior studies of smaller programs may be inappropriate for thinking about how larger voucher-style programs ultimately affect students.

References

Alliance for School Choice. "School Choice Yearbook." Various years. http://afcgrowthfund.org/yearbook/

Angrist, Joshua, Eric Bettinger, and Michael Kremer. 2006. "Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia." *American Economic Review* 96(3), 847-862.

Arizona Department of Revenue. "Individual Income Tax Statistics." Various years. https://www.azdor.gov/ReportsResearch/IndividualTaxStats.aspx

Arizona Department of Revenue. "Tax Expenditure Reports." Various years. https://www.azdor.gov/ReportsResearch/TaxExpenditures.aspx

Averett, Nancy and James Wilkerson. 2002. "Tax Law Little Aid to Poor Students." *The Morning Call.* August 4.

Broughman, Stephen, Nancy Swaim, Randall Parmer, Allison Zotti, and Sarah Dial. 2014. "Private School Universe Survey (PSS): Public-Use Data File User's Manual for School Year 2011-12." United States Department of Education, National Center for Education Statistics.

Bush, Jeb. 2014. "We Need School Choice Now." *National Review*, January 27. http://www.nationalreview.com/article/369444/we-need-school-choice-now-jeb-bush

Card, David and Alan B. Krueger. 1992a. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." *Journal of Political Economy* 100(1), 1-40.

Card, David and Alan B. Krueger. 1992b. "School Quality and Black-White Relative Earnings: A Direct Assessment." *Quarterly Journal of Economics* 107(1), 151-200.

Card, David and Alan B. Krueger. 1996. "School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina." *Journal of Economic Perspectives* 10(4), 31-50.

Card, David and A. Abigail Payne. 2002. "School finance reform, the distribution of school spending, and the distribution of student test scores." *Journal of Public Economics* 83(1), 49-82.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2011. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR." *Quarterly Journal of Economics* 126(4), 1593-1660.

Conley, Timothy G. and Christopher R. Taber. 2011. "Inference with "Difference in Differences" with a Small Number of Policy Changes." *Review of Economics and Statistics* 93(1), 113-125.

De Giorgi, Giacomo, Michelle Pellizzari, and Silvia Redaelli. 2010. "Identification of Social Interactions through Partially Overlapping Peer Groups." *American Economic Journal: Applied Economics* 2(2), 241-275.

Dee, Thomas S. 2004. "Teachers, Race, and Student Achievement in a Randomized Experiment." *Review of Economics and Statistics* 86(1), 195-210.

Ding, Weili and Steven F. Lehrer. 2007. "Do Peers Affect Student Achievement in China's Secondary Schools?" *Review of Economics and Statistics* 89(2), 300-312.

Engec, Necati. 2006. "Relationship between Mobility and Student Performance and Behavior." *Journal of Educational Research* 99(3), 167-178.

Epple, Dennis and Richard E. Romano. 1998. "Competition between Private and Public Schools, Vouchers, and Peer-Group Effects." *American Economic Review* 88(1), 33-62.

Ferreyra, Maria Marta. 2007. "Estimating the Effects of Private School Vouchers in Multidistrict Economies." *American Economic Review* 97(3), 789-817.

Figlio, David N. 1997. "Teacher salaries and teacher quality." *Economic Letters* 55, 267-271.

Figlio, David N. 2002. "Can Public Schools Buy Better-Qualified Teachers?" *Industrial and Labor Relations Review* 55(4), 686-699.

Figlio, David N. 2007. "Boys named Sue: Disruptive children and their peers." *Education Finance and Policy* 2(4), 376-394.

Finn, Jeremy D., Susan B. Gerber, and Jayne Boyd-Zaharias. 2005. "Small Classes in the Early Grades, Academic Achievement, and Graduating from High School." *Journal of Educational Psychology* 97(2), 214-223.

Florida Department of Education. "Florida Tax Credit Scholarship Program Quarterly Reports." Various years. http://www.fldoe.org/schools/school-choice/k-12-scholarship-programs/ftc/quarterly-reports.stml

Florida Department of Education. "McKay Scholarship Program Quarterly Reports." Various years. http://www.fldoe.org/schools/school-choice/k-12-scholarship-programs/mckay/quarterly-reports.stml

Friedman Foundation for Educational Choice. "The ABCs of School Choice." Various years. http://www.edchoice.org/research/the-abcs-of-school-choice/

Graham, Bryan S. 2008. "Identifying Social Interactions through Conditional Variance Restrictions." *Econometrica* 76(3), 643-660.

Gibbons, Stephen, Sandra McNally, and Martina Viarengo. 2012. "Does Additional Spending Help Urban Schools? An Evaluation Using Boundary Discontinuities." IZA Discussion Paper No. 6281.

Greene, Jay P., Paul E. Peterson, Jiangtao Du, Leesa Boeger, and Curtis L. Frazier. 1996. "The Effectiveness of School Choice in Milwaukee: A Secondary Analysis of Data from the Program's Evaluation." Harvard University Program in Education Policy and Governance Occasional Paper 96-3.

Gullickson, Angela. 2012. "Iowa's School Tuition Organization Tax Credit: Tax Credits Program Evaluation Study." Iowa Department of Revenue, Tax Research and Program Analysis Section.

Haegeland, Torbjorn, Oddbjorn Raaum, and Kjell G. Salvanes. 2012. "Pennies from heaven? Using exogenous tax variation to identify effects of school resources on pupil achievement." *Economics of Education Review* 31(5), 601-614.

Hansen, Benjamin. 2008. "School Year Length and Student Performance: Quasi-Experimental Evidence." Unpublished manuscript, University of California Santa Barbara.

Hanushek, Eric A. 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature* 24(3), 1141-1177.

Hanushek, Eric A. 1989. "Expenditures, Efficiency, and Equity in Education: The Federal Government's Role." *American Economic Review* 79(2), 46-51.

Hanushek, Eric A. 1996. "A More Complete Picture of School Resource Policies." *Review of Educational Research* 66(3), 397-409.

Hanushek, Eric A. 1997. "Assessing the Effects of School Resources on Student Performance: An Update." *Educational Evaluation and Policy Analysis* 19(2), 141-164.

Hanushek, Eric A. 1998. "The Evidence on Class Size." University of Rochester, W. Allen Wallis Institute of Political Economy Occasional Paper No. 98-1.

Hanushek, Eric A. 2002. "Evidence, politics, and the class size debate," in *The Class Size Debate*, eds. Lawrence Mishel and Richard Rothstein (Washington, DC: Economic Policy Institute), 37-65.

Hanushek, Eric A., John F. Kain, Jacob M. Markman, and Steven G. Rivkin. 2003. "Does Peer Ability Affect Student Achievement?" *Journal of Applied Econometrics* 18, 527-544.

Hanushek, Eric A. John F. Kain, and Steven G. Rivkin. 2004. "Disruption versus Tiebout improvement: the costs and benefits of switching schools." *Journal of Public Economics* 88(9), 1721-1746.

Howell, William G., Patrick J. Wolf, David E. Campbell, and Paul E. Peterson. 2002. "School Vouchers and Academic Performance: Results from Three Randomized Field Trials." *Journal of Policy Analysis and Management* 21(2), 191-217.

Hoxby, Caroline. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." NBER Working Paper No. 7867.

Huerta, Luis A. and Chad d'Entremont. 2007. "Education Tax Credits in a Post-Zelman Era: Legal, Political, and Policy Alternatives to Vouchers?" *Educational Policy* 21(1), 73-109.

Hungerman, Daniel and Kevin Rinz. 2015. "Where Does Voucher Funding Go? How Large-Scale Subsidy Programs Affect Private-School Revenue, Enrollment, and Prices." NBER Working Paper No. 21687.

Hyman, Joshua. 2013. "Does Money Matter in the Long Run? Effects of School Spending on Educational Attainment." Working paper, University of Michigan.

Illinois Department of Revenue. 2013. "Annual Report of Collections and Distributions, Fiscal Year 2011 and Fiscal Year 2012." http://tax.illinois.gov/publications/AnnualReport/Annual-Report-2011-2012.pdf

Ingersoll, Gary M., James P. Scamman, and Wayne D. Eckerling. 1989. "Geographic Mobility and Student Achievement in an Urban Setting." *Educational Evaluation and Policy Analysis* 11(2), 143-149.

Jackson, C. Kirabo, Rucker Johnson, and Claudia Persico. 2014. "The Effect of School Finance Reforms on the Distribution of Spending, Academic Achievement, and Adult Outcomes." NBER Working Paper No. 20118.

Krueger, Alan B. 1999. "Experimental Estimates of Education Production Functions." *Quarterly Journal of Economics* 114(2), 497-532.

Krueger, Alan B. 2002. "Understanding the magnitude and effect of class size on student achievement," in *The Class Size Debate*, eds. Lawrence Mishel and Richard Rothstein (Washington, DC: Economic Policy Institute), 7-35.

Krueger, Alan B. and Dianne M. Whitmore. 2001. "The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR." *Economic Journal* 111(468)

Lavy, Victor. 2012. "Expanding School Resources and Increasing Time on Task: Effects of a Policy Experiment in Israel on Student Academic Achievement and Behavior." NBER Working Paper No. 18369.

LeFevre, Andrew. 2014. "Ten Thousand Lifeboats: Improving Students' Educational Futures via Pennsylvania's Scholarship Tax Credit Programs." Commonwealth Foundation for Public Policy Alternatives.

Matsudaira, Jordan D. 2008. "Mandatory summer school and student achievement." *Journal of Econometrics* 142(2), 829-850.

Minnesota House of Representatives Research Department. 2011. "Income Tax Deductions and Credits for Public and Nonpublic Education in Minnesota." http://www.house.leg.state.mn.us/hrd/pubs/educcred.pdf

Marcotte, Dave E. and Steven W. Hemelt. 2008. "Unscheduled School Closings and Student Performance." *Education Finance and Policy* 3(3), 316-338.

Mehana, Majida and Arthur J. Reynolds. 2004. "School mobility and achievement: a meta-analysis." *Children and Youth Services Review* 26, 93-119.

Mishel, Lawrence and Richard Rothstein (Eds.) 2002. *The Class Size Debate*. Washignton, DC: Economic Policy Institute.

Nechyba, Thomas J. 2000. "Mobility, Targeting, and Private-School Vouchers." *American Economic Review* 90(1), 130-146.

Peterson, Paul E., William G. Howell, and Jay P. Greene. 1999. "An Evaluation of the Cleveland Voucher Program after Two Years." Program on Education Policy and Governance, Harvard University.

Pischke, Jörn-Steffen. 2007. "The Impact of Length of the School Year on Student Performance and Earnings: Evidence from the German Short School Years." *Economic Journal* 117(523), 1216-1242.

Rogers, Bob and Angela Gullickson. 2012. "Iowa's Tuition and Textbook Tax Credit: Tax Credits Program Evaluation Study." Iowa Department of Revenue, Tax Research and Program Analysis Section.

Rouse, Cecilia Elena. 1998. "Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program." *Quarterly Journal of Economics* 113(2), 553-602.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2015. *Integrated Public Use Microdata Series: Version 6.0* [Machine-readable database]. Minneapolis: University of Minnesota.

Rumberger, Russell W. and Katherine A. Larson. 1998. "Student Mobility and the Increased Risk of High School Dropout." *American Journal of Education* 107(1), 1-35.

Sacerdote, Bruce. 2001. "Peer Effects with Random Assignment: Results for Dartmouth Roommates." *Quarterly Journal of Economics* 116(2), 681-704.

Schanzenbach, Diane Whitmore. 2006a. "Classroom Gender Composition and Student Achievement: Evidence from a Randomized Experiment." Unpublished manuscript, University of Chicago.
Schanzenbach, Diane Whitmore. 2006b. "What Have Researchers Learned from Project STAR?" *Brookings Papers on Education Policy* 9, 205-228.
Stinebrickner, Todd R. and Ralph Stinebrickner. 2006. "What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds." Unpublished manuscript, University of Western Ontario.
Temple, Judy A. and Arthur J. Reynolds. 1999. "School Mobility and Achievement: Longitudinal Findings From an Urban Cohort." *Journal of School Psychology* 37(4), 355-377.
United States Bureau of Labor Statistics. *Unemployment Rate in Alabama*, retrieved from FRED, Federal Reserve Bank of St. Louis. https://research.stlouisfed.org/fred2/series/ALUR/ (and similar for all other states)
United States Bureau of the Census. *Median Household Income in Alabama*, retrieved from FRED, Federal Reserve Bank of St. Louis. https://research.stlouisfed.org/fred2/series/MEHOINUSALA646N/ (and similar for all other states)
United States Bureau of the Census. "CPS October 2013 – Detailed Tables, Table 7." http://www.census.gov/hhes/school/data/cps/2013/tables.html
United States Bureau of the Census. "Public Education Finances." Various years. http://www2.census.gov/govs/school/13f33pub.pdf
United States Department of Education, National Center for Education Statistics. Common Core of Data, Local Education Agency Finance Survey microdata, 2011-12.
United States Department of Education, National Center for Education Statistics. Common Core of Data, Public Elementary/Secondary School Universe Survey microdata, 2013-14.
United States Department of Education, National Center for Education Statistics. *Individualized Education Program Students*, retrieved from Elementary/Secondary Information System (ElSi), United States Department of Education, National Center for Education Statistics. http://nces.ed.gov/ccd/elsi/
United States Department of Education, National Center for Education Statistics. Private School Universe Survey microdata, 1989-90 through 2011-12.
United States Department of Education, National Center for Education Statistics, Schools and Staffing Survey. "Public School Data File." 2007-08. https://nces.ed.gov/surveys/sass/tables/sass0708_035_s1s.asp
Urban Institute, National Center for Charitable Statistics. Core Data Files microdata, 1989-2012.
Urban Institute, National Center for Charitable Statistics. Digitized Data Master Header v2005b microdata.
Urban Institute, National Center for Charitable Statistics. Digitized Data Functional Expenses v2005b microdata.
Welner, Kevin. 2008. *NeoVouchers*. Lanham, MD: Rowman & Littlefield Publishers, Inc.
Witte, John F., Troy D. Sterr, and Christopher A. Thorn. 1995. "Fifth-Year Report: Milwaukee Parental Choice Program." LaFollette School Working Paper No. 1995-001.
Wolf, Patrick. Babette Gutmann, Michael Puma, Brian Kisida, Lou Rizzo, Nada Eissa, and Matthew Carr. 2010. "Evaluation of the DC Opportunity Scholarship Program." United States Department of Education.
Wolf, Patrick J., Paul E. Peterson, and Martin R. West. 2001. "Results of a School Voucher Experiment: The Case of Washington, D.C. after Two Years." Program on Education Policy and Governance, Harvard University.
Zimmerman, David J. 2003. "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment." *Review of Economics and Statistics* 85(1), 9-23.

Appendix

Matching Procedure and Additional Data Details

Financial information from tax forms is linked to other school information from the PSS using a crosswalk that matches the Employer Identification Number (EIN) used by the IRS to the internal PSS identifier (PIN) created by the NCES. The correspondence between these identifiers is created by matching on institution name and address information. The annual extracts of "Core" tax data provided by the NCCS contain the name and street address of each organization in each year.[53] That information also appears in the PSS, though only in more recent rounds. The PSS provides school names beginning in 2005-06 and street addresses beginning in 2007-08. The crosswalk used here was created using the 2005-60, 2007-08, and 2009-10 rounds of the PSS and the NCCS Core data files for 1989-2010.

Before conducting the matching, organization names and addresses were adjusted to eliminate differences in case and remove extraneous punctuation. Four types of matches across the two datasets were then collected: 1) exact address matches, based on street address, ZIP code, and state; 2) exact name matches, within the same ZIP code; 3) similar names within the same "ZIP plus four" area; and 4) exact name matches within the same state, allowing for ZIP code mismatches. In the cases 3 and 4, matches were identified manually from algorithmically determined candidates. In a small number of cases, this matching procedure generates matches between multiple EINs and a single PIN, or vice versa.[54] These matches are discarded.

The tax and PSS samples were created for different purposes, and the set of schools that appear in both of them may differ from the schools each contains on its own. As discussed in the text, the PSS was designed to collect information on all schools that are not primarily supported by public funding, provide classroom instruction in at least one K-12 grade, and have at least one teacher. This includes schools of a wide variety of sizes and religious affiliations.

The tax data, on the other hand, contain information on non-profits generally, and the criteria for inclusion (non-profits incorporated under section 501(c)(3) of the Internal Revenue Code with usual gross receipts of at least $50,000 annually) were not designed to ensure the presence of any particular types of school.[55] In fact, some types of school are practically excluded from the tax data. Schools that are affiliated with churches or religious orders are not required to file Form 990, so many religious (and especially Catholic) schools do not appear in the tax data. As Appendix Table 2 indicates, the rates at which Catholic schools from the PSS appear in the crosswalk and are matched to tax records are low. Other religious schools also match to the tax data at lower rates than do secular schools, though at higher rates than Catholic schools. On the other hand, some schools that appear in the tax data are excluded from the PSS, such as those that offer only pre-kindergarten grades.

Another difference between the PSS and tax data arises when the crosswalk is used to connect a school's records from the two datasets to the same year. The PSS records information as of October for each

---

[53] Each of these extracts includes data from the most recent available tax form for each organization. As such, if the timing of an organization's filing or the timing of the creation of the extracts varies, some schools may see single fiscal years appear in multiple extracts. The year in which each extract is created is included in the data and referred to as its "circa" year. When a single fiscal year appears in multiple extracts, the version from the most recent circa year is retained, in case any updates were filed subsequently to the point at which the data from the initial form was recorded.

[54] Although both the EIN and PIN are intended to be permanent identifiers, it is possible for multiple matches to occur in both directions. If a single organization operates multiple schools, one EIN could match with multiple PINs. On the other hand, if management of a school changes, on PIN could match with multiple EINs.

[55] As noted in a footnote in the main text, the gross receipts threshold that triggered the requirement to file was $25,000 before 2010.

academic year it surveys. The tax data, on the other hand, records information for schools' fiscal years, which schools get to choose and which may differ from their academic years. Fortunately, most schools select fiscal years that correspond well with their academic years. In general, tax records are assigned to academic years based on the calendar year in which the October of the fiscal year falls. To illustrate with an example, if one school submits a tax form that covers the fiscal year from July 1, 2011 to June 30, 2012, that form would be matched to the 2011-12 academic year, because the October contain in that fiscal year is October 2011, which lines up with the information collected in October 2011 for the 2011-12 PSS. If another school's tax form covered the fiscal year running from January 1, 2009 to December 31, 2009, that form would be matched to the 2009-10 PSS, since the October it contains is October 2009.

State covariates used in this analysis are obtained from several sources. State unemployment rates and median incomes come from the Bureau of Labor Statistics and the Census Bureau, respectively. Population-related measures, including the shares in various age by race/ethnicity groups, noncitizens, and those born abroad, are estimated using 1990 and 2000 census data, as well as data from the American Community Survey for 2001-2011, all of which was obtained from the Minnesota Population Center's Integrated Public Use Microdata Series (IPUMS). For the age by race/ethnicity groups, values for 1991-1999 are calculated by assigning respondents from the 2000 census to the categories they would have been in for each of those years, based on their state of residence in 2000. Other population-based measures are linearly interpolated between censuses. States' land areas, which are used to calculate population densities, come from the 2002 Statistical Abstract of the United States.

Crowdflower Details

In order to collect information on schools' compensation expenditures after 2003, I collect PDF copies of their tax returns from GuideStar.org. From that site, I downloaded all available tax forms for the schools that appear in the crosswalk described above. I use a website called Crowdflower.com, which functions similarly to Amazon's Mechanical Turk service, to record data from those PDFs. Customers provide instructions for completing tasks and set a piece rate that is paid to contributors that complete them. Crowdflower distributes these tasks to its users, collects their responses, performs quality control, distributes payment to users who successfully complete tasks, and aggregates responses for use by customers.

To enable Crowdflower users to process my data, I provided access to the PDF tax forms, instructions about which information to record and how, examples of correctly recorded values, and explanations as to why these answers are correct. Crowdflower distributed these to its users according to the quality control settings I selected.[56] The accuracy of contributors' responses can be tested by comparing them to "gold standard" responses provided by customers, as well as by comparing them to the responses of other contributors.

In order to participate in my tasks, contributors were first required to complete a "quiz" composed of requests to record data from tax forms that I had already recorded and provided to Crowdflower. Users that recorded the correct values in response at least 80 percent of these test questions were allowed to continue transcribing data. Similar test questions are also randomly included alongside requests to transcribe new data as contributors continue working, and contributors' responses to them are continuously monitored. Contributors who fall below 80 percent accuracy on test questions at any point in their participation in the task are excluded, and their prior responses are not included in the results reported to the customer.

Each piece of new data to be recorded was distributed to multiple Crowdflower contributors, and their responses were compared to each other to ensure accuracy. If the first two contributors who completed

[56] Here I describe the quality control settings I actually used for the vast majority of my data collection. Generally, these various measures can be used to greater or lesser extents, or in some cases not at all.

a particular task submitted the same value, that value was considered correct, and no further responses were solicited. If the first two contributors disagreed, a third response was requested. If the third response matched one of the first two, the matched value was considered correct. If all three submissions differed, a fourth was solicited. If the fourth matched one of the first three, the matched value was considered correct. If all four submissions differed, I manually determined the correct value. In the vast majority of cases, the correct value was determined by matching responses from contributors.

Because the organization of the form has changed over the years, I created separate sets of tasks for each form-year. The particular lines from which I collected compensation expenditure data in each year are given here:

Form 990 2005 and earlier	Form 990 2006-2007	Form 990 2008 and later	Form 990-EZ All years
Part II, line 25, column (A) – compensation of officers, directors, etc. *Part II, line 26, column (A)* – other salaries and wages *Part II, line 27, column (A)* – pension plan contributions *Part II, line 28, column (A)* – other employee benefits *Part II, line 29, column (A)* – payroll taxes	*Part II, line 25a, column (A)* – compensation of current officers, etc. *Part II, line 25b, column (A)* – compensation of former officers, etc. *Part II, line 25c, column (A)* – compensation and other distributions to disqualified persons *Part II, line 26, column (A)* – other salaries and wages *Part II, line 27, column (A)* – pension plan contributions *Part II, line 28, column (A)* – other employee benefits *Part II, line 29, column (A)* – payroll taxes	*Part IX, line 5, column (A)* – compensation of current officers, directors, trustees, and key employees *Part IX, line 6, column (A)* – compensation to disqualified persons *Part IX, line 7, column (A)* – other salaries and wages *Part IX, line 8, column (A)* – pension plan contributions *Part IX, line 9, column (A)* – other employee benefits *Part IX, line 10, column (A)* – payroll taxes	Part I, line 12 – salaries, other compensation, and employee benefits

Maximum State Spending Instrument Construction

As an instrument for the amount of money states actually spend on voucher-style programs in each state-year, I construct a measure of the maximum amount they are willing to spend. Where states have explicitly imposed caps on the amount of money particular programs can spend, I use those in constructing my instrument. Such caps are available for the corporate tax credit programs in Arizona, Florida, Iowa, and Pennsylvania.

Where programs do not have legislated maximum spending amounts, I impute an amount based on the eligible population in the year before the program was created. For the own-child tax credit programs in Illinois and Iowa, on where actual state spending increased dramatically at their introduction/expansion and then saw smaller changes thereafter, I calculate the amount of money that would have been spent on these tax credits if they had been claimed on behalf of every private school student in each state in the year prior to

the creation of these programs.[57] I then use these values, held constant over time, as the maximum amounts Illinois and Iowa are willing to spend on their own-child tax credits in each year of their existence.

For Arizona's original individual STO tax credit program and Florida's McKay voucher program, which saw initially low levels of actual spending grow substantially over time, I calculate each state's initial willingness to spend using the prior-year population of eligible students. I then allow each state's willingness to spend on these programs to grow proportionally with national growth in the size of the relevant population and the cost of education.

For Arizona, I calculate initial willingness to spend using tables available from the state's Department of Revenue that give the distribution of individual taxpayers across income and the type of filing made (i.e. single, married filing jointly, etc.). I use the tables from the earliest available year, 2000, to calculate willingness to spend for 1998, implicitly assuming that the distribution of Arizona taxpayers does not change between 1998 and 2000. I calculate the amount of money Arizona would have spent on its individual STO tax credit if all taxpayers had claimed the maximum credit available to them, given their tax liability and the program's initial parameters. I then use one percent of this value as Arizona's willingness to spend on this program in its first year.

For Florida's McKay voucher program, I calculate the state's willingness to spend in the program's first year using the number of students in Florida with Individualized Education Programs and the state's cost of public education per pupil in the prior year (i.e. 1998). I take the product of these two numbers and again use one percent of that value as Florida's willingness to spend on this program in its first year.

For both Arizona and Florida, I allow state willingness to spend on these programs to grow in subsequent years as the relevant student populations and the cost of education grow nationally. For Arizona's individual STO tax credit program, the relevant population is all private school students. For Florida's McKay voucher program, the relevant population is students with Individualized Education Programs.

In the construction of my instrument for these programs, states are separately sensitive to growth in the population of relevant students and growth in the cost of education. They increase their willingness to spend in proportion to growth in each by increasing the share of the maximum possible spending for the year prior to program adoption that they are willing to cover (i.e. increase from the one percent assumed for each program's first year). States never decrease their willingness to spend on these programs. That share increases according to

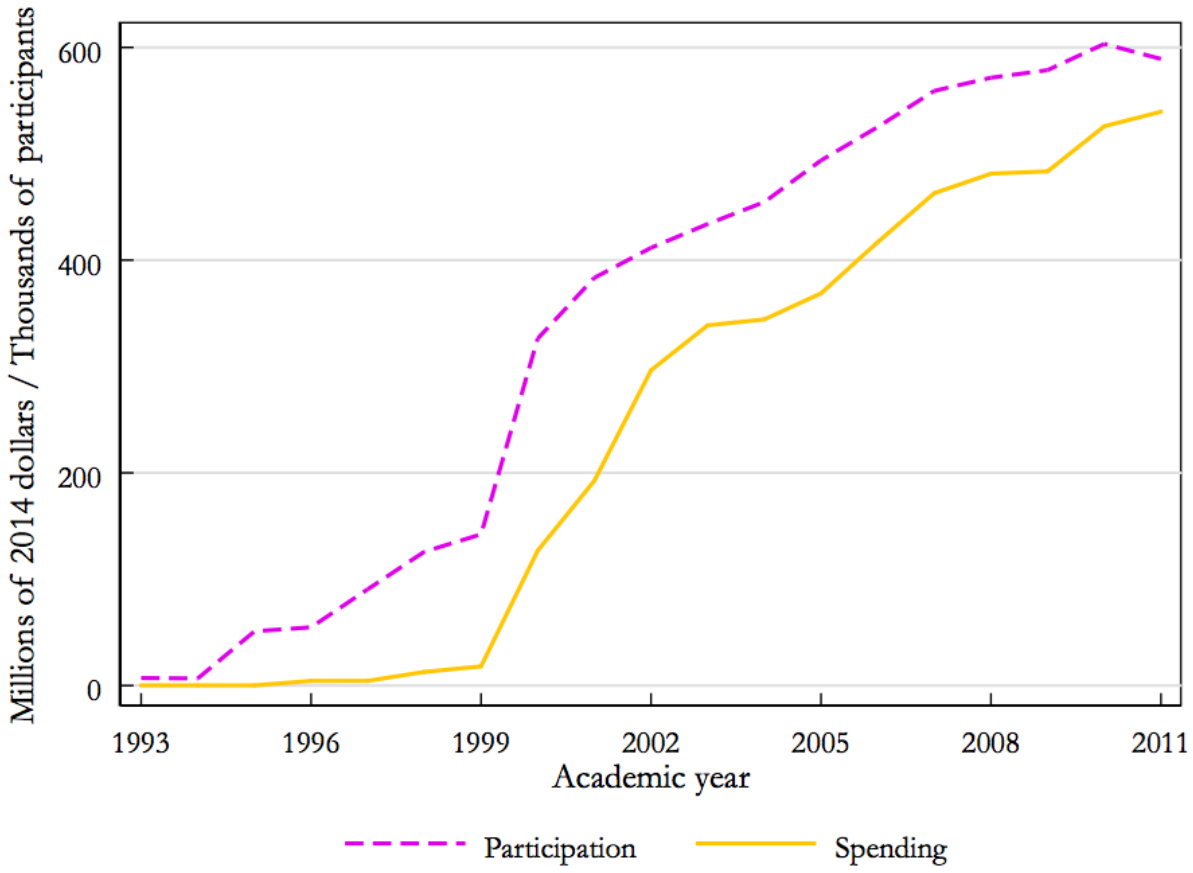
$$Share_{t+1} = Share_t + max\left(\frac{max(\%PopGrowth_t, 0) + max(\%CostGrowth_t, 0)}{10}, .002\right) \qquad \text{(A1)}$$

where percent growth in the relevant population and the cost of education for time *t* are calculated using periods *t* and *t-1*. The .002 term ensures that nominal willingness to spend will increase at least 0.2 percent annually, or about one tenth the targeted rate of overall inflation. Each state's willingness to spend in each year is then this share for each state-year times the maximum possible spending calculated for the year prior to program adoption.

[57] In Illinois, I use 1998-99 private school enrollment for this calculation. In Iowa, I use 1997-98 private school enrollment. Since Iowa's own-child tax credit program existed at a much smaller scale prior to that expansion, I also use 1997-98 enrollment and previous program parameters to calculate maximum spending for years prior to 1998.

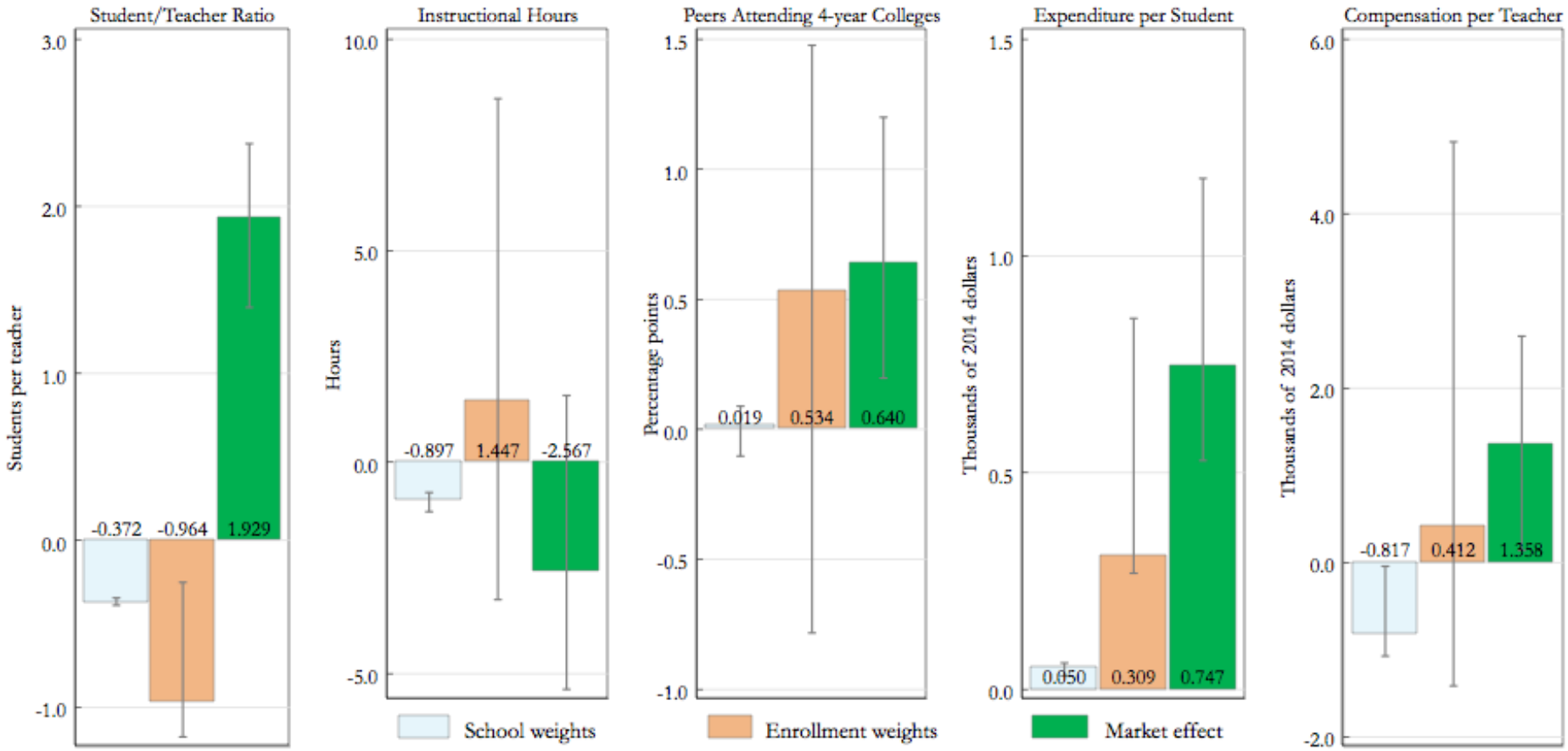
Figures and Tables

Figure 1: Participation in and Spending on Main Voucher-Style Programs



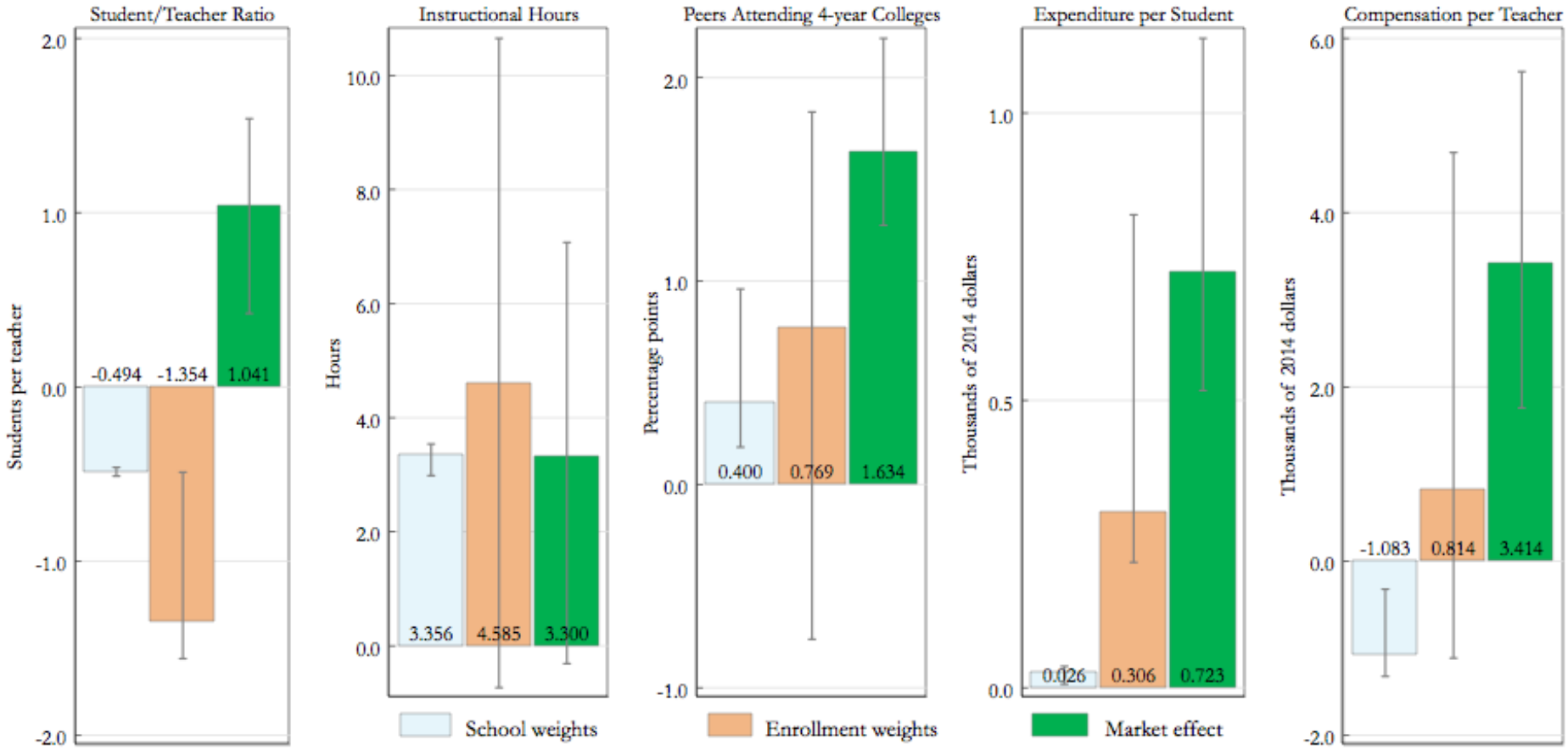
Note: The figure reports participation in and spending on eight voucher-style programs in Arizona, Florida, Illinois, Iowa, and Pennsylvania. The count of participants is equal to the sum of scholarships awarded by tax credit-supported STOs, traditional vouchers awarded by Florida's McKay voucher program, and tax returns that claim a credit for spending on a dependent child's private education. Total spending is the sum of money awarded by STOs as scholarships, money spent on McKay vouchers, and tax revenue foregone due to own-child tax credits. Own-child tax credit spending is measured by states on a calendar year basis; here it is assigned to the academic year beginning in the calendar year for which the corresponding credits were claimed. Academic years are labeled using the calendar year in which they begin.

Figure 2: Margins of Adjustment to Voucher-Style Programs, All Students



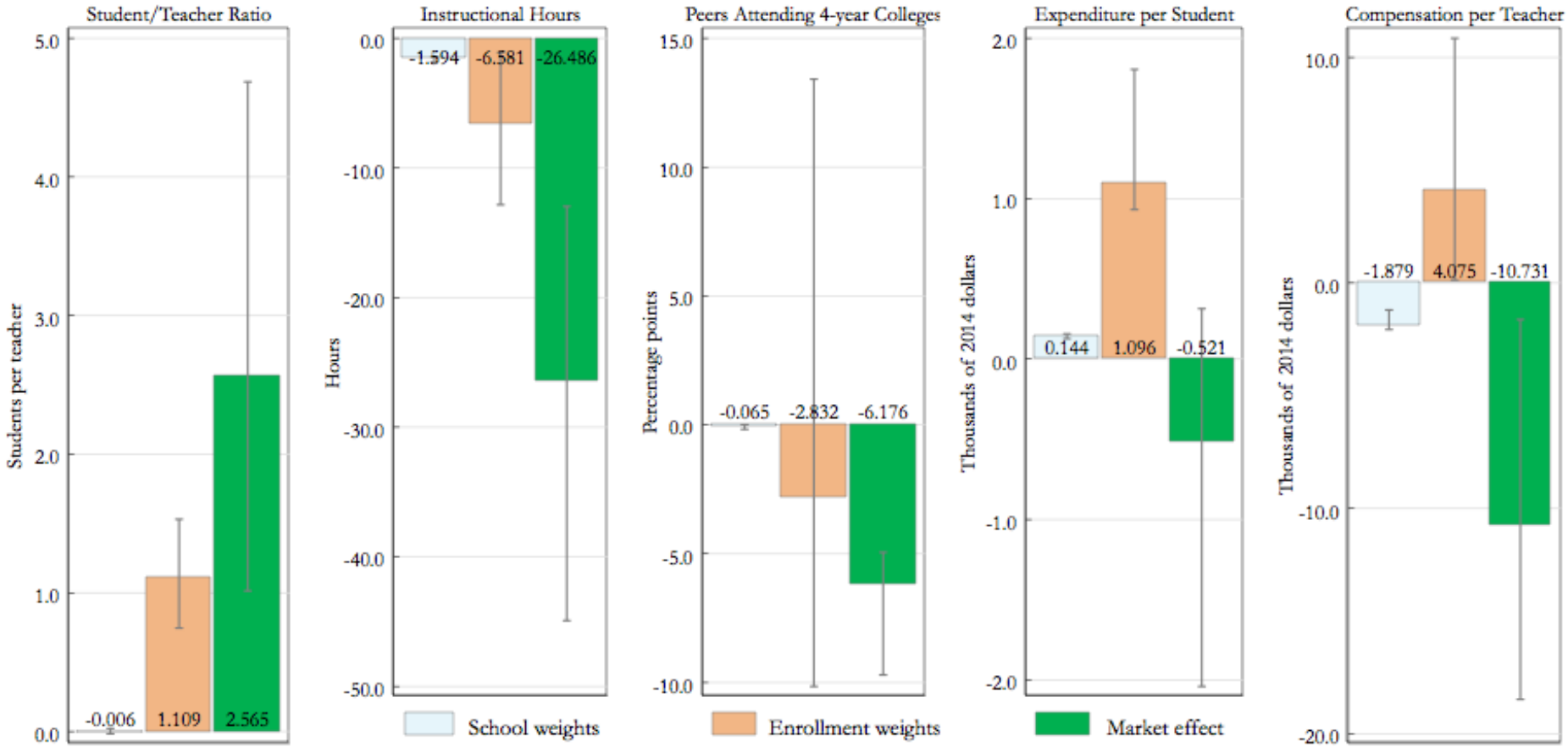
Note: Figure presents estimates from three sets of regressions: school-level regressions with school fixed effects, weighted with PSS sample weights (left bar, shown in blue); school-level regressions with school fixed effects, weighted with enrollment (middle bar, shown in orange); and state-level regressions weighted with total enrollment (right bar, shown in green). More details are provided in Section 5.2. All regressions include the set of state-level controls listed in Section 4. Error bars show 95% confidence intervals, constructed using an approach based on Conley and Taber (2011), as described in Section 4.

Figure 3: Margins of Adjustment to Voucher-Style Programs, White Students



Note: Figure presents estimates from three sets of regressions: school-level regressions with school fixed effects, weighted with PSS sample weights (left bar, shown in blue); school-level regressions with school fixed effects, weighted with enrollment (middle bar, shown in orange); and state-level regressions weighted with total enrollment (right bar, shown in green). More details are provided in Section 5.2. All regressions include the set of state-level controls listed in Section 4. Error bars show 95% confidence intervals, constructed using an approach based on Conley and Taber (2011), as described in Section 4.

Figure 4: Margins of Adjustment to Voucher-Style Programs, Black Students



Note: Figure presents estimates from three sets of regressions: school-level regressions with school fixed effects, weighted with PSS sample weights (left bar, shown in blue); school-level regressions with school fixed effects, weighted with enrollment (middle bar, shown in orange); and state-level regressions weighted with total enrollment (right bar, shown in green). More details are provided in Section 5.2. All regressions include the set of state-level controls listed in Section 4. Error bars show 95% confidence intervals, constructed using an approach based on Conley and Taber (2011), as described in Section 4.

Table 1: Details of Main Voucher-Style Programs Used

State	Year Enacted/ Expanded	Financing Mechanism	Major Student Eligibility Restrictions	Average Annual Spending
Arizona	1997	STO tax credit (individual)	None	$36.0 million
Arizona	2006	STO tax credit (corporate)	Means-tested; requires prior public school attendance	$6.8 million
Florida	1999	Traditional voucher	Students must have an Individualized Education Plan or a Section 504 plan to be eligible; requires prior public school attendance	$106.4 million
Florida	2001	STO tax credit (corporate)	Means-tested; requires prior public school attendance	$91.9 million
Illinois	1999	Own-child tax credit	None	$84.4 million
Iowa	1987, expanded 1996/1998	Own-child tax credit	1987-1996: Means-tested 1996 on: None	1993-2011: $11.5 million 1993-1997: $1.4 million 1998-2011: $16.6 million
Iowa	2006	STO tax credit (both individual and corporate)	Means-tested	$7.6 million
Pennsylvania	2001	STO tax credit (corporate)	Means-tested	$42.8 million

Note: STO tax credits are awarded in return for donations to non-profit, scholarship granting organizations (sometimes known as School Tuition Organizations, or STOs). For STO tax credits, the types of taxpayer that are eligible for the credit are listed in parentheses. Own-child tax credits are awarded in to taxpayers who spend money on the private education of their own dependent children. This type of tax credit is only available to individual taxpayers. The eligibility restrictions listed here are on top of requirements shared by all programs, such as requirements that students live and attend school within the state offering the program and that students be of roughly traditional school age. Average annual spending for each program is calculated from the program's first year through 2011-12 and is reported in 2014 dollars. Information in this table was drawn from reports by the Research Department of the Minnesota House of Representatives (2011), the Friedman Foundation (various years), and the Alliance for School Choice (various years), as well as various state sources.

Table 2: Educational Input Means, by Race

	Student-Teacher Ratio	Instructional Hours	Graduating Peers Attending 4-year Colleges (percent)	Per-Student Expenditure	Per-Teacher Compensation
All Students	21.01	1210.7	77.362	12.604	97.317
	(0.621)	(6.616)	(1.094)	(0.441)	(3.858)
White Students	20.83	1206.4	78.033	12.273	93.436
	(0.608)	(8.335)	(1.039)	(0.420)	(3.445)
Black Students	20.55	1230.9	70.973	13.966	113.566
	(0.476)	(6.183)	(1.821)	(1.058)	(5.252)
Number of Schools	288,165	288,165	41,197	31,710	26,361

Note: Means are calculated using all available data from 1993-2011. Financial outcomes are in thousands of 2014 dollars. Compensation includes salary, benefits, and payroll taxes. College attendance among graduating peers is reported in percentage points, i.e. 1.0 = 1%. Standard errors are clustered at the state level.

Table 3: Effects of Voucher-Style Programs on Private School Enrollment, by Race

	Total	White	Black
Enrollment Levels			
Index	7,528	3,787	2,330
	(5,385 ~ 9,150)	(2,219 ~ 5,115)	(1,897 ~ 2,612)
	[11,048]	[8,803]	[1,072]
N	444	444	444
R^2	0.997	0.996	0.992
Outcome Mean	118,837	89,824	11,560
Enrollment Shares			
Index		-0.636	0.187
		(-0.982 ~ -0.304)	(-0.009 ~ 0.398)
		[0.455]	[0.231]
N		444	444
R^2		0.99	0.967
Outcome Mean		75.586	9.728

Note: Two measures of precision are reported below each estimate: 95% confidence intervals constructed using an approach based on Conley and Taber (2011) are presented in parentheses; and standard errors from OLS regressions, clustered at the state level, are reported in brackets. College attendance among graduating peers is reported in percentage points. Financial outcomes are in thousands of 2014 dollars. The state controls used in these regressions include median income, the unemployment rate, population density, share of residents who are non-citizens, share of residents born abroad, and shares of the population in several race/ethnicity by age bins, as discussed in Section 4. Share regressions are weighted using the total private school enrollment in each state-year. All regressions include time-varying state controls, state fixed effects, year fixed effects, and state-specific linear time trends.

Table 4: Effects of Voucher-Style Programs on Educational Inputs, All Students

	Student-Teacher Ratio	Instructional Hours	Graduating Peers Attending 4-year Colleges (percent)	Per-Student Expenditure	Per-Teacher Compensation
Index	1.929	-2.567	0.64	0.747	1.358
	(1.644 ~ 2.356)	(-4.816 ~ 0.146)	(-1.074 ~ 2.614)	(0.573 ~ 1.112)	(-0.616 ~ 4.159)
	[0.695]	[4.455]	[0.998]	[0.214]	[1.616]
N	444	444	261	847	663
R^2	0.719	0.934	0.928	0.947	0.815
Outcome Mean	21.01	1,210.74	77.36	12.60	97.32

Note: Two measures of precision are reported below each estimate: 95% confidence intervals constructed using an approach based on Conley and Taber (2011) are presented in parentheses; and standard errors from OLS regressions, clustered at the state level, are reported in brackets. College attendance among graduating peers is reported in percentage points. Financial outcomes are in thousands of 2014 dollars. The state controls used in these regressions include median income, the unemployment rate, population density, share of residents who are non-citizens, share of residents born abroad, and shares of the population in several race/ethnicity by age bins, as discussed in Section 4. All regressions include time-varying state controls, state fixed effects, year fixed effects, and state-specific linear time trends.

Table 5: Effects of Voucher-Style Programs on Educational Inputs, by Race

	Student-Teacher Ratio	Instructional Hours	Graduating Peers Attending 4-year Colleges (percent)	Per-Student Expenditure	Per-Teacher Compensation
White Students					
Index	1.041	3.300	1.634	0.723	3.414
	(0.689 ~ 1.502)	(0.986 ~ 6.744)	(0.042 ~ 3.664)	(0.551 ~ 1.032)	(0.596 ~ 9.021)
	[0.783]	[5.096]	[1.078]	[0.224]	[2.417]
N	444	444	261	847	663
R^2	0.731	0.929	0.923	0.941	0.788
Outcome Mean	20.8	1,206.4	78.0	12.273	93.436
Black Students					
Index	2.565	-26.486	-6.176	-0.521	-10.731
	(1.075 ~ 4.008)	(-35.939 ~ -12.657)	(-10.086 ~ -1.120)	(-1.224 ~ 0.198)	(-27.580 ~ -3.980)
	[1.428]	[12.425]	[1.855]	[1.179]	[6.085]
N	444	444	261	841	662
R^2	0.583	0.79	0.906	0.91	0.704
Outcome Mean	20.5	1,230.9	71.0	13.966	113.566

Note: Two measures of precision are reported below each estimate: 95% confidence intervals constructed using an approach based on Conley and Taber (2011) are presented in parentheses; and standard errors from OLS regressions, clustered at the state level, are reported in brackets. College attendance among graduating peers is reported in percentage points. Financial outcomes are in thousands of 2014 dollars. The state controls used in these regressions include median income, the unemployment rate, population density, share of residents who are non-citizens, share of residents born abroad, and shares of the population in several race/ethnicity by age bins, as discussed in Section 4. All regressions include time-varying state controls, state fixed effects, year fixed effects, and state-specific linear time trends.

Table 6: Effects of Voucher-Style Programs on Educational Inputs at New Schools, by Race

	Student-Teacher Ratio	Instructional Hours	Graduating Peers Attending 4-year Colleges (percent)
All Students			
Index	5.301	-26.405	-3.113
	(-2.025 ~ 20.588)	(-66.898 ~ 11.802)	(-16.118 ~ 9.407)
	[6.748]	[20.31]	[6.129]
N	34,222	34,222	2,985
R^2	0.039	0.076	0.167
Outcome Mean	34.4	1,196.9	67.1
White Students			
Index	5.199	5.507	5.767
	(-3.066 ~ 25.603)	(-34.141 ~ 45.403)	(-8.569 ~ 18.366)
	[8.490]	[29.52]	[5.926]
N	30,929	30,929	2,801
R^2	0.05	0.101	0.186
Outcome Mean	35.3	1,184.5	67.9
Black Students			
Index	6.088	-126.435	-29.516
	(0.028 ~ 14.191)	(-187.079 ~ -88.113)	(-46.664 ~ -16.170)
	[3.740]	[28.50]	[13.60]
N	19,443	19,443	2,010
R^2	0.036	0.06	0.266
Outcome Mean	28.9	1,255.2	57.3

Note: Two measures of precision are reported below each estimate: 95% confidence intervals constructed using an approach based on Conley and Taber (2011) are presented in parentheses; and standard errors from OLS regressions, clustered at the state level, are reported in brackets. College attendance among graduating peers is reported in percentage points. The state controls used in these regressions include median income, the unemployment rate, population density, share of residents who are non-citizens, share of residents born abroad, and shares of the population in several race/ethnicity by age bins, as discussed in Section 4. All regressions include time-varying state controls, state fixed effects, year fixed effects, and state-specific linear time trends.

Table 7: Effects of Voucher-Style Programs on the Likelihood of Attending New and Closing Schools

	All Students		White Students		Black Students	
	Attend New School	Attend Closing School	Attend New School	Attend Closing School	Attend New School	Attend Closing School
Index	-2.554	-2.882	-2.574	-2.193	0.825	-6.859
	(-3.606 ~ -1.181)	(-3.978 ~ -0.545)	(-3.672 ~ -1.214)	(-3.296 ~ -0.444)	(-3.299 ~ 6.211)	(-10.972 ~ -2.429)
	[2.362]	[2.234]	[2.807]	[2.274]	[2.344]	[3.694]
N	444	364	444	364	444	364
R^2	0.587	0.565	0.604	0.538	0.603	0.529
Outcome Mean	9.674	9.503	9.392	8.968	12.433	13.464

Note: Two measures of precision are reported below each estimate: 95% confidence intervals constructed using an approach based on Conley and Taber (2011) are presented in parentheses; and standard errors from OLS regressions, clustered at the state level, are reported in brackets. Outcomes are reported in percentage points. The state controls used in these regressions include median income, the unemployment rate, population density, share of residents who are non-citizens, share of residents born abroad, and shares of the population in several race/ethnicity by age bins, as discussed in Section 4. All regressions include time-varying state controls, state fixed effects, year fixed effects, and state-specific linear time trends.

Table 8: Effects of Voucher-Style Programs on Private School Entry and Exit

	School Entry	School Exit
Index	-1.621	-1.01
	(-4.174 ~ 0.124)	(-5.295 ~ 1.512)
	[4.617]	[5.814]
N	444	364
R^2	0.926	0.918
Outcome Mean	19.531	19.826

Note: Two measures of precision are reported below each estimate: 95% confidence intervals constructed using an approach based on Conley and Taber (2011) are presented in parentheses; and standard errors from OLS regressions, clustered at the state level, are reported in brackets. Outcomes are reported in percentage points. The state controls used in these regressions include median income, the unemployment rate, population density, share of residents who are non-citizens, share of residents born abroad, and shares of the population in several race/ethnicity by age bins, as discussed in Section 4. Regressions are weighted using the number of schools in each state-year. Outcome means are based on data from 1995 on. All regressions include time-varying state controls, state fixed effects, year fixed effects, and state-specific linear time trends.

Table 9: Robustness Checks, Student-Teacher Ratio

	Baseline	No Covariates	Quadratic Trends	Covariates in Levels	Future Adopters	More Programs
All Students						
Index	1.929	1.523	0.671	1.43	2.724	1.314
	(1.644 ~ 2.356)	(1.239 ~ 2.097)	(0.448 ~ 0.987)	(1.139 ~ 1.831)	(2.562 ~ 3.092)	(1.044 ~ 1.708)
	[0.695]	[1.554]	[0.731]	[1.122]	[0.815]	[0.873]
N	444	444	444	444	204	454
R^2	0.719	0.554	0.768	0.774	0.794	0.72
White Students						
Index	1.041	1.021	0.053	0.756	1.718	0.582
	(0.689 ~ 1.502)	(0.625 ~ 1.578)	(-0.221 ~ 0.459)	(0.419 ~ 1.226)	(1.409 ~ 2.106)	(0.213 ~ 1.007)
	[0.783]	[1.616]	[0.912]	[1.274]	[1.333]	[0.808]
N	444	444	444	444	204	454
R^2	0.731	0.632	0.771	0.775	0.769	0.732
Black Students						
Index	2.565	1.654	3.091	3.382	2.564	1.714
	(1.075 ~ 4.008)	(0.136 ~ 3.386)	(1.438 ~ 3.971)	(1.876 ~ 5.008)	(1.106 ~ 3.649)	(0.432 ~ 2.916)
	[1.428]	[1.261]	[1.455]	[1.980]	[1.607]	[1.037]
N	444	444	444	444	204	454
R^2	0.583	0.54	0.628	0.593	0.673	0.585
State Controls	Yes	No	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
State Trends	Linear	Linear	Quadratic	Linear	Linear	Linear

Note: Two measures of precision are reported below each estimate: 95% confidence intervals constructed using an approach based on Conley and Taber (2011) are presented in parentheses; and standard errors from OLS regressions, clustered at the state level, are reported in brackets. Regressions reported in the "No Covariates" column do not include state-level economic and demographic controls. In the "Covariates in Levels" regressions, controls measuring the population in various race/ethnicity by age groups enter in levels (i.e. population counts) rather than as shares of the total population. The "Future Adopters" regressions restrict the control group to states that eventually adopt voucher-style programs (AL, CO, GA, IN, LA, MS, NV, NH, NC, OK, RI, SC, UT, VA) or have long, pre-existing histories of offering smaller scale support for private school attendance (ME, MN, VT). Programs in Georgia, Rhode Island, and Utah are added to the treatment group in the "More Programs" regressions.

Table 10: Robustness Checks, Instructional Hours

	Baseline	No Covariates	Quadratic Trends	Covariates in Levels	Future Adopters	More Programs
All Students						
Index	-2.567	1.824	-2.851	-3.709	-5.479	-1.447
	(-4.816 ~ 0.146)	(-0.931 ~ 4.468)	(-4.653 ~ -0.370)	(-5.902 ~ -0.988)	(-7.312 ~ -3.339)	(-4.050 ~ 1.113)
	[4.455]	[4.532]	[6.335]	[5.927]	[5.653]	[3.335]
N	444	444	444	444	204	454
R^2	0.934	0.93	0.953	0.937	0.945	0.935
White Students						
Index	3.3	5.663	1.17	2.294	1.1	3.529
	(0.986 ~ 6.744)	(3.102 ~ 9.329)	(-1.023 ~ 4.079)	(0.010 ~ 5.971)	(-0.913 ~ 4.352)	(1.206 ~ 6.596)
	[5.096]	[5.535]	[8.120]	[6.636]	[7.570]	[3.569]
N	444	444	444	444	204	454
R^2	0.929	0.923	0.949	0.931	0.918	0.929
Black Students						
Index	-26.486	-12.862	-21.459	-32.19	-16.004	-12.008
	(-35.939 ~ -12.657)	(-21.465 ~ 1.071)	(-28.053 ~ -5.365)	(-40.664 ~ -18.236)	(-23.045 ~ -2.595)	(-18.889 ~ 1.086)
	[12.425]	[11.237]	[11.289]	[11.828]	[15.321]	[10.272]
N	444	444	444	444	204	454
R^2	0.79	0.764	0.824	0.798	0.794	0.787
State Controls	Yes	No	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
State Trends	Linear	Linear	Quadratic	Linear	Linear	Linear

Note: Two measures of precision are reported below each estimate: 95% confidence intervals constructed using an approach based on Conley and Taber (2011) are presented in parentheses; and standard errors from OLS regressions, clustered at the state level, are reported in brackets. Regressions reported in the "No Covariates" column do not include state-level economic and demographic controls. In the "Covariates in Levels" regressions, controls measuring the population in various race/ethnicity by age groups enter in levels (i.e. population counts) rather than as shares of the total population. The "Future Adopters" regressions restrict the control group to states that eventually adopt voucher-style programs (AL, CO, GA, IN, LA, MS, NV, NH, NC, OK, RI, SC, UT, VA) or have long, pre-existing histories of offering smaller scale support for private school attendance (ME, MN, VT). Programs in Georgia, Rhode Island, and Utah are added to the treatment group in the "More Programs" regressions.

Table 11: Robustness Checks, Share of Peers Attending Four-Year Colleges

	Baseline	No Covariates	Quadratic Trends	Covariates in Levels	Future Adopters	More Programs
All Students						
Index	0.64	0.524	2.027	0.524	-0.101	0.552
	(-1.074 ~ 2.614)	(-1.156 ~ 2.959)	(1.230 ~ 3.144)	(-1.109 ~ 2.774)	(-1.653 ~ 0.962)	(-0.444 ~ 1.761)
	[0.998]	[0.727]	[2.642]	[1.066]	[1.220]	[0.848]
N	261	261	261	261	117	270
R^2	0.928	0.918	0.952	0.94	0.949	0.928
White Students						
Index	1.634	1.202	2.377	1.926	1.468	1.597
	(0.042 ~ 3.664)	(-0.452 ~ 3.744)	(1.636 ~ 3.345)	(0.315 ~ 4.117)	(0.177 ~ 2.443)	(0.588 ~ 2.787)
	[1.078]	[0.858]	[2.556]	[1.163]	[1.586]	[0.906]
N	261	261	261	261	117	270
R^2	0.923	0.91	0.948	0.934	0.946	0.923
Black Students						
Index	-6.176	-3.503	-3.31	-8.698	-9.017	-5.901
	(-10.086 ~ -1.120)	(-7.196 ~ -1.883)	(-6.177 ~ -0.239)	(-12.175 ~ -3.475)	(-12.471 ~ -3.691)	(-7.708 ~ -2.330)
	[1.855]	[1.648]	[6.783]	[2.399]	[3.514]	[1.116]
N	261	261	261	261	117	270
R^2	0.906	0.88	0.941	0.912	0.915	0.912
State Controls	Yes	No	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
State Trends	Linear	Linear	Quadratic	Linear	Linear	Linear

Note: Two measures of precision are reported below each estimate: 95% confidence intervals constructed using an approach based on Conley and Taber (2011) are presented in parentheses; and standard errors from OLS regressions, clustered at the state level, are reported in brackets. Regressions reported in the "No Covariates" column do not include state-level economic and demographic controls. In the "Covariates in Levels" regressions, controls measuring the population in various race/ethnicity by age groups enter in levels (i.e. population counts) rather than as shares of the total population. The "Future Adopters" regressions restrict the control group to states that eventually adopt voucher-style programs (AL, CO, GA, IN, LA, MS, NV, NH, NC, OK, RI, SC, UT, VA) or have long, pre-existing histories of offering smaller scale support for private school attendance (ME, MN, VT). Programs in Georgia, Rhode Island, and Utah are added to the treatment group in the "More Programs" regressions.

Table 12: Robustness Checks, Per-Student Expenditures

	Baseline	No Covariates	Quadratic Trends	Covariates in Levels	Future Adopters	More Programs
All Students						
Index	0.747	0.766	0.58	0.801	0.662	0.265
	(0.573 ~ 1.112)	(0.571 ~ 1.053)	(0.433 ~ 0.797)	(0.638 ~ 1.156)	(0.492 ~ 0.853)	(0.088 ~ 0.538)
	[0.214]	[0.204]	[0.199]	[0.268]	[0.226]	[0.237]
N	847	847	847	847	391	864
R^2	0.947	0.944	0.963	0.952	0.968	0.947
White Students						
Index	0.723	0.724	0.563	0.751	0.656	0.208
	(0.551 ~ 1.032)	(0.537 ~ 0.966)	(0.372 ~ 0.738)	(0.590 ~ 1.067)	(0.512 ~ 0.870)	(0.032 ~ 0.455)
	[0.224]	[0.181]	[0.141]	[0.253]	[0.225]	[0.239]
N	847	847	847	847	391	864
R^2	0.941	0.938	0.962	0.946	0.956	0.94
Black Students						
Index	-0.521	-0.506	-1.43	-0.586	-0.563	-0.815
	(-1.224 ~ 0.198)	(-1.426 ~ 0.119)	(-1.959 ~ -0.936)	(-1.235 ~ 0.116)	(-0.903 ~ 0.090)	(-1.508 ~ -0.122)
	[1.179]	[1.035]	[1.789]	[1.030]	[1.107]	[0.717]
N	841	841	841	841	391	858
R^2	0.91	0.891	0.933	0.912	0.932	0.908
State Controls	Yes	No	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
State Trends	Linear	Linear	Quadratic	Linear	Linear	Linear

Note: Two measures of precision are reported below each estimate: 95% confidence intervals constructed using an approach based on Conley and Taber (2011) are presented in parentheses; and standard errors from OLS regressions, clustered at the state level, are reported in brackets. Regressions reported in the "No Covariates" column do not include state-level economic and demographic controls. In the "Covariates in Levels" regressions, controls measuring the population in various race/ethnicity by age groups enter in levels (i.e. population counts) rather than as shares of the total population. The "Future Adopters" regressions restrict the control group to states that eventually adopt voucher-style programs (AL, CO, GA, IN, LA, MS, NV, NH, NC, OK, RI, SC, UT, VA) or have long, pre-existing histories of offering smaller scale support for private school attendance (ME, MN, VT). Programs in Georgia, Rhode Island, and Utah are added to the treatment group in the "More Programs" regressions.

Table 13: Robustness Checks, Per-Teacher Compensation

	Baseline	No Covariates	Quadratic Trends	Covariates in Levels	Future Adopters	More Programs
All Students						
Index	1.358	2.048	0.457	0.692	2.031	-3.21
	(-0.616 ~ 4.159)	(0.015 ~ 4.890)	(-0.543 ~ 2.691)	(-1.445 ~ 3.243)	(0.467 ~ 4.587)	(-5.415 ~ -1.058)
	[1.616]	[2.579]	[3.225]	[2.163]	[2.443]	[2.439]
N	663	663	663	663	303	680
R^2	0.815	0.798	0.837	0.824	0.861	0.809
White Students						
Index	3.414	3.716	4.118	0.42	2.77	-3.23
	(0.596 ~ 9.021)	(0.917 ~ 9.263)	(2.680 ~ 9.001)	(-2.700 ~ 5.688)	(-0.465 ~ 8.419)	(-6.121 ~ 3.149)
	[2.417]	[2.238]	[3.942]	[2.988]	[3.805]	[3.358]
N	663	663	663	663	303	680
R^2	0.788	0.773	0.813	0.8	0.818	0.763
Black Students						
Index	-10.731	-14.645	-6.585	-9.136	-4.236	-15.2
	(-27.580 ~ -3.980)	(-32.265 ~ -8.258)	(-14.164 ~ -0.719)	(-26.151 ~ -2.196)	(-22.474 ~ 1.364)	(-28.044 ~ -10.233)
	[6.085]	[7.783]	[17.684]	[6.761]	[8.388]	[4.004]
N	662	662	662	662	303	679
R^2	0.704	0.681	0.75	0.704	0.794	0.702
State Controls	Yes	No	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
State Trends	Linear	Linear	Quadratic	Linear	Linear	Linear

Note: Two measures of precision are reported below each estimate: 95% confidence intervals constructed using an approach based on Conley and Taber (2011) are presented in parentheses; and standard errors from OLS regressions, clustered at the state level, are reported in brackets. Regressions reported in the "No Covariates" column do not include state-level economic and demographic controls. In the "Covariates in Levels" regressions, controls measuring the population in various race/ethnicity by age groups enter in levels (i.e. population counts) rather than as shares of the total population. The "Future Adopters" regressions restrict the control group to states that eventually adopt voucher-style programs (AL, CO, GA, IN, LA, MS, NV, NH, NC, OK, RI, SC, UT, VA) or have long, pre-existing histories of offering smaller scale support for private school attendance (ME, MN, VT). Programs in Georgia, Rhode Island, and Utah are added to the treatment group in the "More Programs" regressions.

Table 14: Effects of Voucher-Style Programs on Educational Inputs, Instrument Variable Estimates, by Race

	Student-Teacher Ratio	Instructional Hours	Graduating Peers Attending 4-year Colleges (percent)	Per-Student Expenditure	Per-Teacher Compensation
All Students					
Actual Spending	1.566	9.138	0.324	0.29	0.113
	[1.204]	[4.716]	[1.248]	[0.259]	[2.174]
N	444	444	261	847	663
R^2	0.716	0.934	0.928	0.947	0.815
White Students					
Actual Spending	0.737	11.673	2.909	0.499	-0.539
	[1.276]	[4.898]	[1.550]	[0.257]	[1.925]
N	444	444	261	847	663
R^2	0.73	0.929	0.922	0.94	0.788
Black Students					
Actual Spending	3.287	19.371	-4.697	-1.137	-18.677
	[1.292]	[9.133]	[1.547]	[0.628]	[7.017]
N	444	444	261	841	662
R^2	0.586	0.774	0.903	0.911	0.707

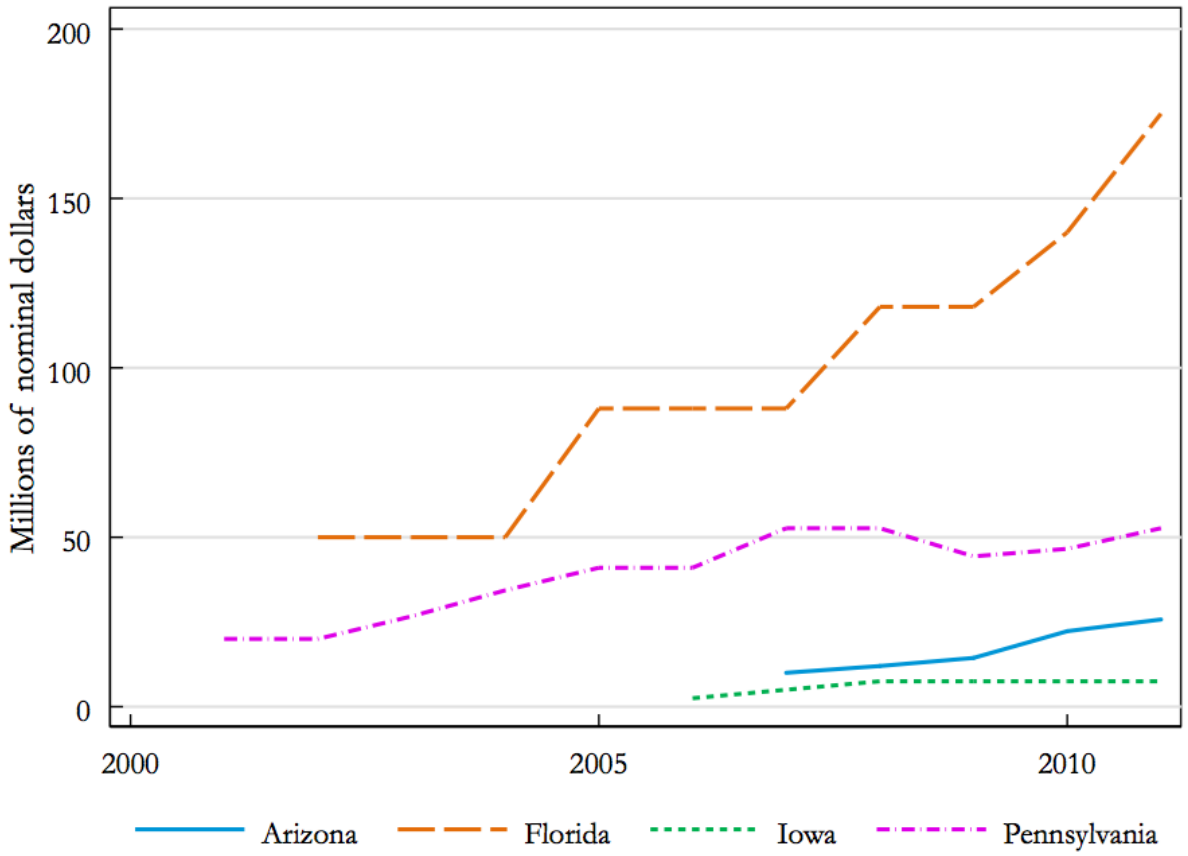
Note: A measure of maximum state spending on voucher-style programs is used as an instrument for actual spending. Legislated maximums are used where available. Where no maximum is imposed by law, maximum spending is constructed as described in the appendix. Estimates presented here have been scaled up by average spending on voucher-style programs in state-years with positive spending (about $60 million). All regressions include time-varying state controls, state fixed effects, year fixed effects, and state-specific linear time trends.

Table 15: Effects of Voucher-Style Programs vs. Differences between Public and Private School Educational Inputs

	Public School	Private School	Difference (Private - Public)	Voucher Effect	Voucher Effect (% of Difference)
White Students					
Student/Teacher Ratio	28.8	18.6	-10.2	1.041	-10.2%
Instructional Hours	1195	1234	39	3.3	8.5%
4-year College Matriculation	39.1%	77.0%	37.9%	1.63%	4.3%
Per-Student Expenditure	$13,230	$12,913	-$317	$723	228.1%
Per-Teacher Compensation	$107,175	$100,899	-$6,276	$3,414	54.4%
Black Students					
Student/Teacher Ratio	19.1	17.7	-1.4	2.565	-183.2%
Instructional Hours	1195	1247	52	-26.486	-50.9%
4-year College Matriculation	34.4%	70.0%	35.6%	-6.18%	-17.4%
Per-Student Expenditure	$15,595	$15,303	-$292	-$521	-178.4%
Per-Teacher Compensation	$104,795	$117,163	$12,368	-$10,731	-86.8%

Note: Private school input estimates were produced using the 2011-12 PSS and matched tax data. Public school student-teacher ratios by race are estimated using the 2013-14 Common Core of Data (CCD) School Survey. Public school expenditure per student and compensation per teacher by race are estimated using the 2011-12 CCD School District surveys. Public school instructional hours were obtained from published estimates based on the 2007-08 Schools and Staffing Survey; estimates were not available by race. Public school four-year college matriculation was backed out for each race using overall four-year college matriculation rates (obtained from the 2013 October supplement to the Current Population Survey), the share of students in private school, and the private-school estimate of four-year college matriculation. Financial outcomes are in 2014 dollars. The sign of the voucher effect as a percentage of the difference between public and private school reflects whether the voucher effect represents an improvement or a worsening in the input in question.

Appendix Figure 1: Legislated Spending Limits for Corporate Tax Credit STO Programs



Note: Information in this table was drawn from reports by the Research Department of the Minnesota House of Representatives (2011), the Friedman Foundation (various years), and the Alliance for School Choice (various years), as well as various state sources.

Appendix Table 1: Additional Details of Main Voucher-Style Programs Used

State	Program	Taxpayer Value Limit	Aggregate Value Limit	Scholarship Value Limit	Student Eligibility	Other Notes
Arizona	Individual STO credit	100% of gift, up to a limit that varies over time and by filer type 1998-99: $500 2000-04: $500 (individuals)/ $625 (married) 2005: $500 (individual)/$825 (married) 2006-11: $500 (individual)/ $1,000 (married)	None	None	All Arizona students in grades K-12 are eligible	STOs must spend at least 90% of revenue on scholarships, must consider income as one of the criteria use to evaluate potential scholarship recipients, may not limit recipients to only one private school; Donors may not make gifts to STOs that are specifically for the use of any particular child, though for much of the period covered by this study this behavior was not explicitly prohibited and does appear to have taken place; credit in excess of tax liability can be carried forward up to five years
Arizona	Corporate STO credit	100% of gift, with no taxpayer-specific limit on the value of the credit	Legislated, see Appendix Figure 1	Scholarships initially capped at $4,200 for grades K-8 and $5,500 for grades 9-12 in 2006, with those limits increasing by $100 each year	Arizona students must have family income ≤ 185% of the eligibility cutoff for free/reduced price school lunches (342% FPL), AND have been enrolled in public school 90 days in the prior year or one semester in the current year OR be enrolled in private kindergarten or private pre-K for students with disabilities OR be the dependent of an active-duty member of the military stationed in Arizona OR be a prior scholarship recipient	Donors cannot designate their gifts for use by particular students; credit in excess of tax liability can be carried forward up to five years; aggregate limit increases by 20% each year
Florida	McKay voucher program	N/A	None	The amount the state would have spent on the student in public school, or private school tuition and fees, whichever is less	Florida students with Individualized Education Programs or Section 504 accommodation plans who have been enrolled in public school for at least one year are eligible	Schools receiving voucher students must be approved by the state; school personnel working with voucher students must pass federal background checks

State	Program	Taxpayer Value Limit	Aggregate Value Limit	Scholarship Value Limit	Student Eligibility	Other Notes
Florida	Corporate STO credit	100% of gift, with the value of the credit capped at 75% of total tax liability	Legislated, see Appendix Figure 1	Through 2005: $3,500 2006-08: $3,750 2009: $3,950 2010 on: calculated based on state public school spending	Florida students are eligible for scholarships if they received a scholarship in the previous year OR re in foster care OR are entering grades K-5 AND are eligible for free/reduced price school lunch (185% FPL) OR are a sibling of a scholarship recipient; scholarships are renewable as long as a student's family income remains below 230% FPL; prior to 2008, new students were eligible for scholarships only when entering grades K-1, and scholarships were renewable only up to 200% FPL	STOs must spend at least 97% of revenue on scholarships, and must spend at least 75% of revenue in the year it was received; STOs must award scholarships on a first-come, first-served basis, with priority given to renewals, and may not award scholarships to children of their owners; scholarship values are reduced by 25%-50% when family income exceeds 185% FPL; for most of the period analyzed here, tax credits in excess of the annual taxpayer-specific limit could be carried forward three to five years, with state approval; the aggregate limit increases by 25% following each year in which at least 90% of available tax credits were awarded
Illinois	Individual own-child credit	25% of spending in excess of $250 on fees and tuition, with maximum credit value of $500 per family	None	N/A	Spending on any Illinois student under 21 years of age attending grades K-12 in public, private, or home school who is the taxpayer's dependent is eligible for the tax credit	Eligible expenses include those on tuition, books, and lab or activity fees; the inclusion of non-tuition expenses makes guardians of public school students eligible for the tax credit
Iowa	Individual own-child credit	Pre 1996: 5% of spending, maximum $50 per dependent 1996-97: 10% of spending, maximum $100 per dependent 1998 on: 25% of spending, maximum $250 per dependent	None	N/A	Pre 1996: families with federal adjusted gross income below $45,000 could receive the credit for spending on dependent students in grades K-12 at any accredited Iowa school not operated for profit (public or private) 1996 on: families can receive the credit for spending on dependent students in grades K-12 at any accredited Iowa school not operated for profit (public or private), regardless of income	Eligible expenses include those on tuition, books, and lab or activity fees; the inclusion of non-tuition expenses makes guardians of public school students eligible for the tax credit

State	Program	Taxpayer Value Limit	Aggregate Value Limit	Scholarship Value Limit	Student Eligibility	Other Notes
Iowa	Individual and corporate STO credit	65% of gift, with no taxpayer-specific limit on the value of the credit	Legislated, see Appendix Figure 1	None, scholarship amounts are determined by STOs	Iowa students with family incomes up to 300% FPL are eligible for scholarships	STOs must spend at least 90% of revenue on scholarships; taxpayers must obtain a certificate from STOs in order to claim the credit; credit can also be claimed for non-cash contributions to a dependent student's education; credits in excess of tax liability can be carried forward up to five years; donors may not designate their gifts for use by particular students; before July 2009, only individual taxpayers could claim the tax credit for donations to STOs; schools affiliate with specific STOs, and no school may be affiliated with more than one STO
Pennsylvania	Corporate STO credit	75% of gift, or 90% with contributions in consecutive years, with the value of the credit capped at $300,000 per corporation annually	Legislated, see Appendix Figure 1	None, scholarship amounts are determined by STOs	Eligibility is restricted based on family income: Pre 2011: Students eligible with family income up to $50,000 plus $10,000 per child 2011: Students eligible with family income up to $60,000 plus $12,000 per child	STOs must spend at least 80% of revenue on scholarships; credits in excess of tax liability may not be carried forward

Note: In this table, FPL refers to the Federal Poverty Level. Information in this table was drawn from reports by the Research Department of the Minnesota House of Representatives (2011), the Friedman Foundation (various years), and the Alliance for School Choice (various years), as well as various state sources.

Appendix Table 2: School and Enrollment Counts by Year and Match Status

		Number of Schools				Enrollment			
		PSS	Crosswalk	Matched	Used	PSS	Crosswalk	Matched	Used
All Schools									
	1993	36,610	4,445	2,256	1,639	5,081,001	1,066,952	534,252	412,840
	1995	28,597	4,950	2,635	1,929	5,364,291	1,211,904	639,159	490,778
	1997	30,226	5,453	3,462	2,540	5,563,975	1,307,259	832,639	629,827
	1999	29,047	5,591	3,598	2,648	5,481,507	1,322,903	857,246	648,227
	2001	30,750	6,330	4,194	3,070	5,831,389	1,499,062	1,009,290	756,776
	2003	30,006	6,750	4,663	3,389	5,512,103	1,517,496	1,062,988	785,539
	2005	29,723	7,390	5,220	3,734	5,468,818	1,622,372	1,161,618	839,566
	2007	28,390	7,409	5,284	3,727	5,197,073	1,610,093	1,165,512	828,742
	2009	28,157	7,565	5,529	3,878	5,004,387	1,633,924	1,206,122	827,185
	2011	26,930	6,806	4,926	3,442	4,702,540	1,491,083	1,088,336	741,802
Catholic									
	1993	8,392	553	142	52	2,532,895	221,395	53,810	17,675
	1995	8,102	586	174	59	2,572,665	238,618	65,346	18,000
	1997	8,113	597	210	84	2,595,685	244,279	83,649	26,458
	1999	7,921	595	221	83	2,547,860	238,676	87,838	26,778
	2001	8,037	631	243	100	2,587,164	251,503	88,623	28,906
	2003	7,669	636	271	114	2,392,303	246,804	96,515	32,318
	2005	7,285	671	296	115	2,273,875	251,977	103,687	32,677
	2007	6,976	674	316	134	2,157,251	249,042	111,986	37,053
	2009	6,826	665	320	131	2,079,364	247,877	109,315	33,973
	2011	6,383	619	280	120	1,928,851	231,047	93,897	29,133
Other Religious									
	1993	15,458	1,718	710	587	1,690,508	417,532	186,112	162,372
	1995	12,221	1,914	857	717	1,819,572	486,988	236,635	203,707
	1997	12,737	2,129	1,149	965	1,895,397	528,314	303,564	264,550
	1999	12,537	2,229	1,231	1,022	1,914,129	550,202	324,359	276,857
	2001	13,604	2,644	1,553	1,274	2,099,241	650,746	413,451	350,435

	2003	13,175	2,813	1,753	1,454	1,965,167	650,043	436,486	371,972
	2005	13,409	3,182	2,069	1,683	2,041,747	727,829	504,903	420,459
	2007	12,795	3,212	2,152	1,746	1,955,197	732,294	511,438	425,046
	2009	12,584	3,260	2,289	1,865	1,836,888	719,536	518,148	431,762
	2011	11,833	2,887	2,000	1,638	1,728,072	654,456	468,926	389,946
Secular									
	1993	6,998	2,121	1,381	1,000	857,598	428,025	294,330	232,793
	1995	8,274	2,450	1,604	1,153	972,054	486,298	337,178	269,071
	1997	9,376	2,727	2,103	1,491	1,072,893	534,666	445,426	338,819
	1999	8,589	2,767	2,146	1,543	1,019,518	534,025	445,049	344,592
	2001	9,109	3,055	2,398	1,696	1,144,984	596,813	507,216	377,435
	2003	9,162	3,301	2,639	1,821	1,154,633	620,649	529,987	381,249
	2005	9,029	3,537	2,855	1,936	1,153,196	642,566	553,028	386,430
	2007	8,619	3,523	2,816	1,847	1,084,625	628,757	542,088	366,643
	2009	8,747	3,640	2,920	1,882	1,088,135	666,511	578,659	361,450
	2011	8,714	3,300	2,646	1,684	1,045,617	605,580	525,513	322,723

Note: "PSS" columns give the number of schools/total enrollment that in the PSS each year. "Crosswalk" columns give the number schools in the PSS each year for which I have been able to match the internal PSS identifier to the Employer Identification Number that appears in the tax data, and the total enrollment at those schools. "Matched" columns give the number of schools in each year that appear in both the PSS and the tax data and the total enrollment at those schools. "Used" columns give the number of schools in each year that appear in both the PSS and the tax data and satisfy the sample restrictions employed in the baseline analysis. Both school and enrollment counts in this table are unweighted. The sum of school counts and enrollment across religious affiliation categories may not equal the total for the full PSS due to missing information on religious affiliation.

Appendix Table 3: Effects of Voucher-Style Programs on Education Inputs, by School Type and Race

	Student-Teacher Ratio	Instructional Hours	Graduating Peers Attending 4-year Colleges (percent)
All Students			
All Schools	1.929	-2.567	0.64
	(1.644 ~ 2.356)	(-4.816 ~ 0.146)	(-1.074 ~ 2.614)
	[0.695]	[4.455]	[0.998]
New Schools	5.301	-26.4	-3.113
	(-2.025 ~ 20.588)	(-66.898 ~ 11.802)	(-16.118 ~ 9.407)
	[6.748]	[20.31]	[6.129]
Existing Schools	1.064	1.384	0.404
	(0.369 ~ 1.907)	(-3.565 ~ 14.646)	(-1.676 ~ 3.105)
	[0.816]	[3.277]	[0.698]
White Students			
All Schools	1.041	3.300	1.634
	(0.689 ~ 1.502)	(0.986 ~ 6.744)	(0.042 ~ 3.664)
	[0.783]	[5.096]	[1.078]
New Schools	5.199	5.507	5.767
	(-3.066 ~ 25.603)	(-34.141 ~ 45.403)	(-8.569 ~ 18.366)
	[8.490]	[29.52]	[5.926]
Existing Schools	0.863	3.707	0.794
	(0.106 ~ 1.826)	(-2.284 ~ 20.392)	(-1.646 ~ 3.400)
	[0.935]	[3.429]	[0.945]
Black Students			
All Schools	2.565	-26.486	-6.176
	(1.075 ~ 4.008)	(-35.939 ~ -12.657)	(-10.086 ~ -1.120)
	[1.428]	[12.425]	[1.855]
New Schools	6.088	-126.4	-29.52
	(0.028 ~ 14.191)	(-187.079 ~ -88.113)	(-46.664 ~ -16.170)
	[3.740]	[28.50]	[13.60]
Existing Schools	1.395	-10.04	-3.515
	(0.586 ~ 3.839)	(-20.305 ~ -0.927)	(-5.726 ~ 1.286)
	[1.152]	[11.68]	[1.550]

Note: New schools are those making their first appearance in the PSS. Existing schools are all schools except new schools. Two measures of precision are reported below each estimate: 95% confidence intervals constructed using an approach based on Conley and Taber (2011) are presented in parentheses; and standard errors from OLS regressions, clustered at the state level, are reported in brackets. Estimates for all schools are reproductions of the baseline estimates in Tables 4 and 5. Estimates for new schools are reproductions of those found in Table 6. Estimates for existing schools were produced at the school level, in the same manner as the estimates for new schools.

Appendix Table 4a: Effects of Voucher-Style Programs on Non-Financial Educational Inputs, with Pre-Law-Change Placebo Dummies

	Student-Teacher Ratio	Instructional Hours	Graduating Peers Attending 4-year Colleges (percent)
All Students			
Index	1.981	-2.131	0.63
	[0.902]	[5.594]	[1.216]
One/Two Year Lead	0.256	0.528	-0.068
	[0.591]	[4.406]	[2.454]
Three/Four Year Lead	-0.374	0.549	0.45
	[0.484]	[2.563]	[2.345]
N	444	444	261
R^2	0.719	0.934	0.928
White Students			
Index	1.011	5.887	0.936
	[1.115]	[6.761]	[1.369]
One/Two Year Lead	0.185	3.364	-1.269
	[0.765]	[6.158]	[2.599]
Three/Four Year Lead	-0.541	2.217	-0.894
	[0.658]	[3.222]	[1.908]
N	444	444	261
R^2	0.731	0.929	0.923
Black Students			
Index	2.456	-35.486	-4.368
	[1.713]	[9.799]	[2.146]
One/Two Year Lead	0.243	-8.207	3.432
	[1.522]	[12.290]	[3.644]
Three/Four Year Lead	-0.941	-18.121	11.263
	[1.224]	[8.618]	[3.904]
N	444	444	261
R^2	0.584	0.793	0.912

Note: The specification used here matches the baseline specification, except for the inclusion of the lead dummies, which are equal to one the specified number of years before a policy change and equal to zero otherwise. Standard errors from OLS regressions, clustered at the state level, are reported in brackets. College attendance among graduating peers is reported in percentage points. All regressions include time-varying state controls, state fixed effects, year fixed effects, and state-specific linear time trends.

Appendix Table 4b: Effects of Voucher-Style Programs on Financial Educational Inputs, with Pre-Law-Change Placebo Dummies

	Per-Student Expenditure			Per-Teacher Compensation		
	All Students	White Students	Black Students	All Students	White Students	Black Students
Index	0.722	0.752	1.492	0.996	3.743	-6.616
	[0.273]	[0.446]	[0.830]	[2.434]	[3.322]	[15.497]
One Year Lead	0.094	0.199	2.222	-1.522	0.225	-4.119
	[0.372]	[0.530]	[1.228]	[5.040]	[5.227]	[20.370]
Two Year Lead	0.053	0.087	2.935	-0.733	0.796	7.149
	[0.434]	[0.579]	[1.980]	[3.545]	[5.799]	[18.067]
Three Year Lead	-0.24	-0.186	2.526	2.31	1.479	22.215
	[0.386]	[0.545]	[1.407]	[3.826]	[4.672]	[16.310]
Four Year Lead	-0.249	-0.23	2.597	0.682	0.757	11.211
	[0.356]	[0.429]	[1.450]	[4.376]	[5.000]	[12.928]
Five Year Lead	-0.117	0.007	2.152	-1.28	-1.273	19.697
	[0.249]	[0.301]	[1.311]	[4.828]	[4.371]	[10.086]
N	847	847	841	663	663	662
R^2	0.948	0.941	0.916	0.815	0.788	0.709

Note: The specification used here matches the baseline specification, except for the inclusion of the lead dummies, which are equal to one the specified number of years before a policy change and equal to zero otherwise. Standard errors from OLS regressions, clustered at the state level, are reported in brackets. Outcomes are reported in 2014 dollars. All regressions include time-varying state controls, state fixed effects, year fixed effects, and state-specific linear time trends.