Illusory Gains from Chile's Targeted School Voucher Experiment

Benjamin Feigenberg, Steven Rivkin, and Rui Yan*

February 10, 2017

Abstract

In 2008, Chile implemented a targeted voucher program that increased voucher values for disadvantaged students at participating schools by approximately 50%. Although disadvantaged students made substantial fourth grade test score gains that other studies have attributed to the program, our analysis raises serious doubts that the program had a substantial effect on cognitive skills. First, there was only a minor reduction in class size and little evidence of increases in any inputs. An audit showed that many schools were not using additional revenues for permitted expenditures, and estimates that exploit a discontinuity in the revenues allocated to schools show no evidence of positive effects of allocated funds on achievement growth. In addition, there is limited evidence of competitive or incentive effects on school quality or that disadvantaged students transitioned to higher quality schools. The much smaller gains made by disadvantaged students in low-stakes eighth grade test scores along with an increased rate of missing scores on fourth grade tests is consistent with extensive strategic behavior by schools. In contrast, increases in parental education and income among disadvantaged children indicate a primary role for improvements in family circumstances of tested students in explaining the meaningful decline in the achievement gap.

*Feigenberg: University of Illinois at Chicago, bfeigenb@uic.edu. Rivkin: University of Illinois at Chicago, sgrivkin@uic.edu. Yan: University of Illinois at Chicago, ryan57@uic.edu.

1 Introduction

Students from economically disadvantaged families potentially face a number of impediments in voucher-funded schooling markets that can limit the benefits of competition. There may be a positive association between family income and the academic and social skills that schools value. In addition, geographical segregation and more limited family resources may elevate the costs of operating schools. If private schools are able to charge tuition above the value of a voucher and selectively admit applicants, disadvantaged students may have few high-quality options.

In Chile, these concerns and evidence that disadvantaged children had realized limited benefits from the Chilean voucher program prompted the passage of a targeted voucher reform known as the Subvención Escolar Preferencial (SEP) in 2008.[1] The SEP program raised the value of school vouchers by 50% for students from the lowest-socioeconomic status (SES) households. In order to receive these additional revenues, both public and private voucher schools were required to sign contracts with the Chilean Ministry of Education that defined anticipated test score gains over the subsequent years, required detailed accounting of SEP program expenditures, eliminated screening of SEP-eligible students based on past academic performance or family background, and prohibited schools from charging SEP-eligible students additional tuition or fees (Correa et al., 2014).

The targeted voucher would be expected to increase the quality of instruction and achievement for low-income students through a number of channels. First, public schools with low-income students and private schools participating in the SEP program would receive additional revenue for each low-income student that could be used to reduce class

[1] Early research, including McEwan (2001) and Hsieh and Urquiola (2006), has concluded that gains associated with the increased market competition were small and did not differentially benefit poorer students. More recent work, including Hanushek et al. (2012), Bravo et al. (2010), and Gallego (2013), identifies more substantial gains associated with voucher program-induced competition but does not provide evidence of any significant convergence in academic achievement based on student socioeconomic status.

size, improve technology, or purchase other resources. Importantly, the law does not permit schools to use the program revenue to raise teacher pay. Second, the higher revenue provides schools an incentive to become more attractive to low-income families, and raising the quality of instruction could be one component of such an effort.[2] Third, the expansion of the set of schools in which low-SES students could enroll at no cost and without the possibility of rejection based on background would be expected to induce some SEP-eligible students to switch to a higher quality school that was previously not available.

In this paper we investigate the impact of SEP on the achievement deficit of low-SES students and the contributions of specific channels to observed changes. Our analysis of the effect of SEP on the achievement gap begins with a simple multi-year differences-in-differences research design that compares low- and high-SES students' test scores before and after the SEP reform and identifies a greater than 0.2 standard deviation improvement in the relative performance of low-SES students in the period after SEP is introduced. This closing of the gap reproduces findings from a growing body of work that attributes a significant reduction in inequality to SEP. Neilson (2013) and Correa et al. (2014) argue that achievement gains among disadvantaged students can be explained primarily by school quality improvements. Navarro-Palau (2015) also finds a significant though more modest effect of SEP on the achievement gap, and concludes based on a regression discontinuity design analysis that the positive program impact was driven by within-public school improvements. The findings in Murnane et al. (2016) also point to the combination of increased school funding and greater accountability as the primary mechanism through which SEP raised achievement.

Yet although we replicate the large decline in the achievement gap, we believe that the body of evidence as a whole provides little support for the belief that the SEP reform

[2] Work including Hoxby (2000), Card et al. (2010), and Lavy (2010) highlights the potential benefits associated with increased market competitiveness in alternative settings.

substantially reduced the achievement gap in Chile. Rather, the improvements for low-SES students appear to be largely illusory, and the findings suggest that relative gains in parental education and log household income account for much if not all of the small decline in the actual achievement differential. We base these conclusions on detailed investigations of the primary channels through which SEP would have been expected to raise the quality of instruction, changes over time in the patterns of missing test scores, changes over time in achievement measures that are not high-stakes for schools that receive SEP funding, and information on family background and its association with achievement.

First, there is little evidence that SEP had a substantial effect on school inputs or that it altered the education market in a manner that raised achievement for low-SES children. Despite the 50% increase in the voucher value for disadvantaged students at participating schools, there was only a minor reduction in class size and results suggest that teacher characteristics became relatively less positive following program implementation. An audit showed that many schools were not using the additional revenues for permitted expenditures, and estimates that exploit a discontinuity in the revenues allocated to schools show little or no evidence of a positive effect of allocated funds on achievement growth. Second, similar to Navarro-Palau (2015), we find limited evidence that disadvantaged students transitioned to higher quality schools during the period of large achievement gains for low-SES students. Third, although difficult to measure, there is little direct evidence of a competitive or incentive effect on school quality.

In contrast, we find evidence consistent with the notion that the incentives to raise fourth grade achievement in order to qualify for an unconditional renewal of SEP funding after four years led schools to engage in strategic behavior that accounts for much of the observed achievement gains. First, the low-stakes eighth grade test score gap declined by only half as much as the fourth grade gap, on average, for the cohorts of students exposed to the SEP program for whom eighth grade scores are available. Second, the rate of missing

fourth grade test scores for likely low-SES students increased substantially in the period immediately following the introduction of the SEP program. Consequent relative increases in parental education and log family income among disadvantaged children who were tested are consistent with the notion that changes in student composition played a primary role in raising the relative achievement of low-SES children.[3] Finally, disadvantaged students who are SEP-eligible but enroll in schools that have not qualified to receive SEP funding experience test score gains that are remarkably similar to the gains experienced by the subset of socioeconomically comparable students who do participate in the SEP program.

2 Educational Data

To conduct our analysis, we draw from a number of sources to assemble a rich database that tracks primary school students across schools and years for the 2005-2014 study period. These include administrative records on matriculation, academic performance, family background, school and teacher characteristics, and SEP eligibility and program participation. Unique school and student identifiers make it possible to track students over time and across schools and merge information from the various data sets.

The restricted-access administrative records provided by the Chilean Ministry of Education for all grade levels for the years 2005-2014 include matriculation and academic performance data. The matriculation records contain information on school attended, grade, attendance rate, and part-time enrollment, and the supplementary academic performance data includes students' grade point average in each year (on a one-to-seven scale), as well as indicators for grade progression, transfer status, academic probation and dropping out.

[3] As we describe in more detail in Section 2, parental education and household income data is based on a household survey that is only conducted among test-takers.

Test score data contain results from the Sistema de Medición de la Calidad de la Educación (SIMCE). The SIMCE battery of exams provides national standardized test scores in math and Spanish for grade four students in every year between 2005 and 2014 and for grade eight students in the years 2007, 2009, 2011, 2013, and 2014. Based on the existing literature, we exploit the comparability of the SIMCE test across years (Neilson, 2013). We are able to match SIMCE test score data to matriculation records using the unique student identifiers provided. Throughout the analysis, we follow the existing literature in normalizing grade four SIMCE math and Spanish test scores by the corresponding means and standard deviations from 2005 (Neilson, 2013; Navarro-Palau, 2015). In the analysis, we focus on the average of these two normalized scores, and we ignore tests conducted in other subjects that are irregularly administered.

Information on parental educational attainment and household income is available only for students who take the SIMCE examinations. We use mother's education and household income to determine socio-economic status. Specifically, we rank students in each grade four cohort based on mother's education and categorize the 40% of students with the lowest level of mother's education as low-SES. For students with the same level of mother's education, we use household income to break ties, where household income is reported as a categorical variable with thirteen-fifteen values in each year. The remaining 60% of students in each cohort are characterized as high-SES students. This dichotomy is based on the structure of the SEP program, which provides targeted vouchers to the 40% of students in each cohort at the bottom of the socioeconomic distribution.

The Ministry of Education provides data on all schools and teachers. We use information on teacher's educational attainment, contract status, and years of experience in the analysis.

Finally, the Ministry also provides data on SEP program eligibility and participation starting with 2008 when the program was first introduced. In addition, the data report

whether the student qualified for SEP program participation based on (1) household enrollment in the Chile Solidario social program, (2) the child's household being identified as among the one-third of most vulnerable households in the Ficha de Protección social safety net program and/or being in Group A of the FONASA public health insurance program, or (3) the child's household reporting sufficiently low household income, low parental education, rural residency status, and/or the child living in a municipality with a high local poverty rate. In addition, SEP files identify all participating schools in each school year as well as the year in which each school officially joined the SEP program.

3 Institutional Details

This section provides an overview of the Chilean education system and describes the SEP program. It highlights issues relevant to the deficit in school quality for economically-disadvantaged children. Students from economically-disadvantaged families potentially face a number of impediments in voucher-funded schooling markets that can limit the benefits of competition. There may be a positive association between family income and the academic and social skills that schools value. In addition, geographical segregation and more limited family resources may elevate the costs of operating schools. If private schools are able to charge tuition above the value of a voucher and selectively admit applicants, disadvantaged students may have quite limited options.

3.1 The Chilean Education System

Influenced by Friedman (1962), Chile's military government adopted a national school voucher program in 1981. Supporters of the voucher program argued that increased competition in the market for primary and secondary education would lead to improved academic achievement across the distribution (Bettinger, 2011). Given the unique scale

of Chile's voucher program, it has attracted substantial academic attention (Correa et al., 2014). Evidence on the efficacy of Chile's voucher program, however, has been mixed, and distributional analyses suggest that the voucher program may have exacerbated stratification based on socioeconomic status (Hsieh and Urquiola, 2006).

Over the past 35 years school funding has primarily been a function of enrollment levels and the annual value of the grade-specific nationwide voucher that goes to public and private voucher schools. Since 1994, private voucher schools in Chile have been permitted to charge tuitions up to three times the value of the nationwide school voucher and to impose their own eligibility criterion in the admissions process. In contrast, public schools have not been allowed to turn away students unless oversubscribed or to charge tuition over and above the school voucher (Urquiola, 2016). The result of this policy regime has been substantial inter-school stratification based on socioeconomic status: as of the mid 2000s, 69% of low-SES students but only 35% of higher-SES students attended public schools.

3.2 The SEP Program

To address the large, persistent gap in achievement based on student background, the Chilean Ministry of Education launched the SEP program in 2008. It was designed to improve educational outcomes for SEP-eligible priority students by encouraging primary schools to enroll these students and focus additional resources on improving their academic performance. In order to incentivize primary schools to accept low-SES applicants, schools were allocated the product of approximately an additional 50% of the baseline voucher payment for each enrolled priority student multiplied by the student's attendance rate (defined to take on a value between zero and one). In addition, schools received supplementary revenue as a function of the share of priority students enrolled in the school and the average attendance rate of these students. For grade levels one through four, schools

received up to an additional 9.8% of the standardized national voucher payment unit if between 15% and 30% of enrolled students were classified as priority. This multiplier increased to 16.8% for enrollment rates between 30% and 45%, to 22.4% for enrollment rates between 45% and 60%, and to 25.2% for priority enrollment rates above 60%.[4] In all cases, the supplementary funding was based on the product of the formula-based multiplier and the average attendance rate of priority students in the relevant grades.

In exchange for receiving these additional funds, participating schools had to sign contracts that ensured that SEP funds would be spent appropriately and that all expenditures would be documented. Appropriate expenditures included spending on additional personnel or school resources, while increased salaries, bonuses and other expenditure categories (debt repayments, school celebrations, etc.) were excluded. In addition to submitting a plan for educational improvement that outlined planned expenditures and anticipated test score gains over subsequent years, schools were required to significantly alter admissions and student retention systems. Specifically, schools could no longer charge tuition or fees to priority students in excess of the voucher revenues received by schools, schools could not selectively admit priority students based on past educational achievement or family background, and schools could not expel priority students for failing a grade before allowing them at least one opportunity to repeat each grade level (SEP, 2008).

In preparation for the introduction of the SEP program in 2008, the Ministry of Education engaged in an information campaign to make school administrators aware of the key features of the program, including additional associated revenues and requirements for program participants. The SEP enrollment period for schools was shortened for the 2008 school year, but 77% of public and voucher private schools nonetheless enrolled in the program in year one (including 51% of all voucher private schools). In 2008, all

[4] For comparison, in 2008, primary schools received 275% of the standardized national voucher payment unit for each full-time student enrolled. This corresponded to approximately $92 USD per month.

priority students in grades one through four were eligible to receive SEP funding if their school enrolled in the program. During subsequent school years, each cohort that had previously been eligible maintained eligibility while incoming first graders were added to the program. As a result, students in grade levels one through eight were eligible by 2012.

In 2008, the SEP program enrolled approximately two-thirds of the number of students that were enrolled in subsequent school years. This lower initial enrollment rate was due primarily to a lower share of students being classified as SEP-eligible. The share of all students in grades one through four who received SEP funding increased from 26.9% in 2008 to 41.9% in 2009 and remained stable thereafter.[5]

4 Academic Achievement Gap

This section describes differences in family and school characteristics by SEP eligibility and then illustrates changes over time in the achievement gap. SEP is designed to cover the bottom two quintiles of the SES distribution, but prior to program implementation in 2008 there is no comprehensive information on eligibility. Therefore, following Neilson (2013), we use information on parental education and income to estimate SES and eligibility status.

In Table 1 we report summary statistics for achievement, family background variables and school sector by time period (pre- or post-reform) and disadvantaged status for a series of measures of disadvantage. Because SEP priority status and recipiency are not available prior to program implementation, only the disadvantage measure based upon family background is shown in the pre-reform period. In the post-reform period, disadvantaged status is measured by this family background variable as well as SEP eligibility and whether an eligible student attended a school that participated in the program.

[5] The share of participating voucher private schools also increased to 61% by 2009 and continued to increase in the following years.

Across all three definitions of disadvantage, Table 1 shows that disadvantaged students have lower levels of parental education and household income, lower grade point averages, and lower SIMCE test scores in both time periods. In addition, these students are more likely to reside in rural areas, are more likely to be enrolled in public schools and are less likely to be enrolled in voucher schools.

Table 1: Variable means by SES and SEP Program Status

	(1) Mother's Education (Years)	(2) Father's Education (Years)	(3) Household Income (Pesos)	(4) GPA (1-7 Scale)	(5) Normalized SIMCE Score	(6) Rural School	(7) Public School	(8) Voucher Private School	(9) Observations
Panel A: Pre-SEP program (2005-2007)									
Low-SES	7.51	8.62	149,288	5.69	-0.36	0.22	0.69	0.31	244,008
High-SES	12.98	12.42	454,958	6.00	0.26	0.05	0.35	0.54	365,226
Panel B: SEP program in place (2008-2014)									
Low-SES	8.15	9.02	213,466	5.71	-0.12	0.21	0.59	0.40	548,533
High-SES	13.18	12.68	624,458	5.96	0.37	0.06	0.28	0.60	822,493
Priority	9.81	9.83	244,310	5.69	-0.04	0.18	0.55	0.45	762,656
Non-Priority	12.20	12.32	625,965	5.91	0.32	0.07	0.30	0.56	934,136
SEP Recipient	9.57	9.57	221,750	5.67	-0.08	0.20	0.63	0.37	670,159
Non-SEP Recipient	12.14	12.26	605,576	5.90	0.31	0.07	0.28	0.59	1,026,633

Notes: Table displays mean values over relevant years for fourth grade students. Household Income measures monthly household income in Chilean Pesos, higher GPA values reflect better academic performance.

However, the top two rows of Panels A and B also reveal the sizable decline in the SIMCE test-score gap following the introduction of the SEP program, as the average differential between high- and low-SES students declines from 0.62 standard deviations in the pre-reform period to 0.49 standard deviations post-reform. To formally characterize the yearly change in relative test performance, we estimate multi-year differences-in-differences models. The first set of specifications examine achievement gains for low-SES

students as a function of SEP recipient status:

$$Testscore_{it} = \alpha + \gamma_t + \sum_{t=2005}^{2014} (LowSES_{it} \cdot \gamma_t)\delta_{1t} + \sum_{t=2008}^{2014} (LowSES_{it} \cdot SEP_{it} \cdot \gamma_t)\delta_{2t} + \sum_{t=2008}^{2014} (HighSES_{it} \cdot SEP_{it} \cdot \gamma_t)\delta_{3t} + \epsilon_{it} \quad (1)$$

The second set of specifications examine corresponding achievement gains for priority students as a function of SEP recipient status:

$$Testscore_{it} = \alpha + \gamma_t + \sum_{t=2008}^{2014} (Priority_{it} \cdot \gamma_t)\delta_{1t} + \sum_{t=2008}^{2014} (Priority_{it} \cdot SEP_{it} \cdot \gamma_t)\delta_{2t} + \epsilon_{it} \quad (2)$$

In the equations above, $Testscore_{it}$ represents the normalized test score of student i in year t, $LowSES_{it}$ is an indicator variable defined based on SIMCE survey responses, $HighSES_{it}$ is an indicator variable representing those students not classified as low SES, SEP_{it} is an indicator variable defined by whether a student is a SEP voucher recipient, γ_t represent year fixed effects, and $Priority_{it}$ is an indicator for whether student i is classified by the Ministry of Education as a SEP-eligible priority student in year t. The sample used to estimate Equation (1) includes all years between 2005 and 2014 and the sample used to estimate Equation (2) includes all years between 2008 and 2014.

Results from the specifications defined by Equations (1)-(2) are presented in Table 2. Column (1) estimates Equation (1) and average gains for low-SES students appear similar to previous estimates, including those presented in Neilsen (2013), which indicate that low-SES students increased their relative test scores by roughly 0.2 standard deviations between 2007 (the year before the SEP program was introduced) and 2014 (the last year for which test score data is currently available). The majority of these test score gains occurred between 2007 and 2011. In Column (2), we add the following student-level

covariates: mother's educational attainment (in years), father's educational attainment and log household income. This specification also includes municipality-by-year fixed effects. There are a total of 346 municipalities in Chile with an average municipality-level population of approximately 50,000. Previous research has used municipalities to define local education markets given that over 90% of primary schools students attend a school in the same municipality in which they reside (Feigenberg, 2016). Column (2) estimates identify smaller gains for low-SES students during the sample period. Although we present direct evidence on changing household socio-demographic characteristics in the subsequent analysis, the decline in measured low-SES gains relative to Column (1) previews our finding that changes in the characteristics of tested students explain a substantial share of the test score convergence during the post-SEP period. Turning to the triple interaction coefficients that characterize differential improvement for low-SES students who are also SEP recipients, we find point estimates that are inconsistent in sign, indicating that test score gains for low-SES students are largely independent of SEP recipient status.

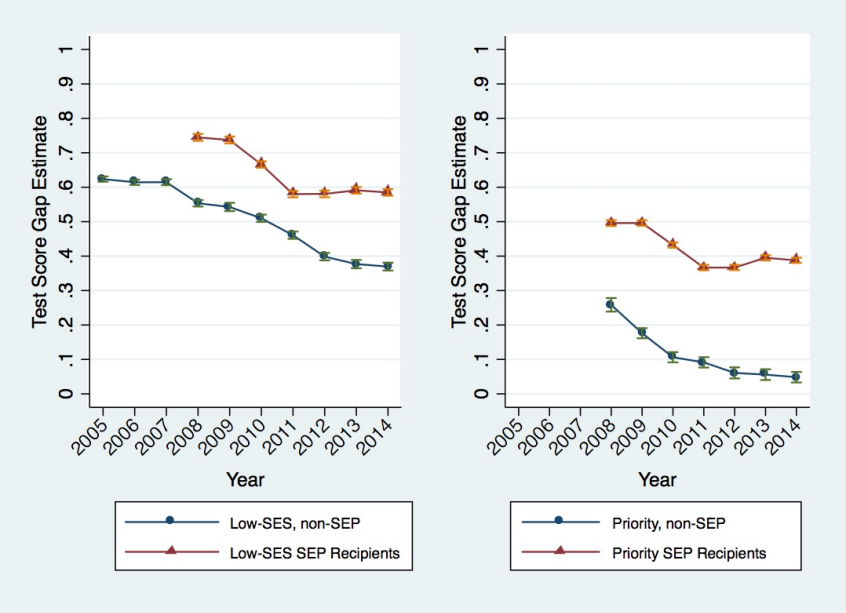
Columns (3)-(4) present parallel estimates based on Equation (2). Here, we identify a similar pattern of relative test score gains for priority students during the post-SEP period, although the magnitudes of relative gains are somewhat smaller than the corresponding estimates from Columns (1)-(2). Again, the triple interaction terms suggest that test score gains are no greater for the subset of disadvantaged students who are SEP recipients once family background is accounted for. In sum, Table 2 estimates show large test score gains for disadvantaged students following the introduction of the SEP program, consistent with a large program effect. At the same time, evidence that these gains are independent of SEP status strengthens the doubts about the role of the SEP program in driving test score convergence. A graphical comparison of Table 2 estimates is presented in Figure 1 and highlights the lack of an association between SEP recipient status and test score gain.

Table 2: Estimated Achievement Deficits for Disadvantaged Students by SEP Program Participation Status and Year

	(1) SIMCE Score	(2) SIMCE Score	(3) SIMCE Score	(4) SIMCE Score
Disadvantaged	-0.624***	-0.171***	-0.259***	-0.044***
	(0.004)	(0.004)	(0.010)	(0.011)
Disadvantaged·2006	0.009*	-0.011*		
	(0.006)	(0.006)		
Disadvantaged·2007	0.009	-0.012**		
	(0.006)	(0.006)		
Disadvantaged·2008	0.070***	0.030***		
	(0.006)	(0.006)		
Disadvantaged·2009	0.081***	0.036***	0.082***	0.035***
	(0.007)	(0.008)	(0.013)	(0.013)
Disadvantaged·2010	0.114***	0.077***	0.152***	0.056***
	(0.007)	(0.007)	(0.013)	(0.013)
Disadvantaged·2011	0.163***	0.134***	0.167***	0.063***
	(0.007)	(0.007)	(0.013)	(0.013)
Disadvantaged·2012	0.225***	0.143***	0.198***	0.076***
	(0.007)	(0.007)	(0.013)	(0.014)
Disadvantaged·2013	0.247***	0.153***	0.203***	0.101***
	(0.007)	(0.007)	(0.013)	(0.013)
Disadvantaged·2014	0.254***	0.164***	0.210***	0.088***
	(0.007)	(0.007)	(0.013)	(0.013)
Disadvantaged·SEP	-0.191***	-0.111***	-0.237***	-0.149***
	(0.006)	(0.006)	(0.010)	(0.011)
Disadvantaged·SEP·2009	-0.003	-0.003	-0.083***	-0.027*
	(0.009)	(0.010)	(0.013)	(0.014)
Disadvantaged·SEP·2010	0.035***	0.038***	-0.088***	0.011
	(0.008)	(0.009)	(0.013)	(0.014)
Disadvantaged·SEP·2011	0.072***	0.062***	-0.038***	0.053***
	(0.008)	(0.009)	(0.013)	(0.014)
Disadvantaged·SEP·2012	0.009	0.029***	-0.069***	0.024*
	(0.009)	(0.009)	(0.013)	(0.014)
Disadvantaged·SEP·2013	-0.023***	0.018**	-0.102***	-0.009
	(0.009)	(0.009)	(0.013)	(0.014)
Disadvantaged·SEP·2014	-0.024***	0.006	-0.102***	-0.010
	(0.009)	(0.009)	(0.013)	(0.014)
Additional Controls		X		X
Municipality-Year Fixed Effects		X		X
Disadvantaged Measure Used	Low SES	Low SES	Priority	Priority
Observations	1,939,551	1,824,304	1,484,663	1,254,936

Notes: Robust standard errors are in parentheses and all specifications include year fixed effects. The dependent variable is the normalized fourth grade test score (normalized by 2005 mean and standard deviation). The low socioeconomic status indicator is determined based on mother's years of education and family income as measured by SIMCE parental surveys from the years 2005-2014. The priority status of a student is designated by the Ministry of Education. Additional controls are mother's years of education, father's years of education, and log household income. Columns (1)-(2) omit coefficients that characterize differential achievement growth for SEP recipients who are not low-SES students.
* significant at 10 percent level ** significant at 5 percent level *** significant at 1 percent level.

Figure 1: Estimated Achievement Deficits for Disadvantaged Students by SEP Program Participation Status and Year



Notes: The figure plots point estimates and 95% confidence intervals from Columns (1) and (3) of Table 2.

5 Potential Channels of SEP Program Effects

The achievement trends both illustrate the substantial closing of the gap and sew doubts about the role of the SEP reform, and in this section we seek to provide additional evidence by examining the primary channels through which the reform would have been expected to raise the quality of instruction for disadvantaged children. First, we assess the effects of SEP on the quantity of school inputs.[6] Second, we investigate the effects of SEP on the distribution of students among schools to assess the possibility that the program led to extensive quality upgrading for low-SES students. Third, we assess whether there is evidence consistent with the SEP program having increased competition in a manner that reduced the achievement gap.

[6] This potential mechanisms is the one least consistent with the findings presented in Table 2.

5.1 SEP Effects on School Inputs

To investigate the extent of SEP-induced increases in school inputs we examine alternative sources of variation in school revenues and expenditures. First, we present the findings from an audit study conducted by the Chilean Comptroller's Office (Comptroller's Office, 2012). This audit compared SEP funding inflows for the 2008-2011 period to documented SEP expenditures for 77 of Chile's 346 municipalities.[7] On average, only 65% of received funds could be linked to validated expenditures during the audit period. Moreover, municipality-level regressions of the change in the test-score gap on funds spent with and without demographic controls reveal little or no evidence of a substantial effect of SEP spending on the within-municipality gap (Table 3, Columns 1 and 2), and subsequent regressions also provide no evidence that disadvantaged students benefitted from increased SEP spending (Table 3, Columns 3 and 4).[8]

Of course these estimates do not capture causal effects, but fortunately the structure of the SEP funding formula discussed in detail in Section 2 enables the use of regression discontinuity methods to identify the causal effects of additional revenue. Specifically, funding increases discontinuously with the share of disadvantaged students. The four sharp cutoffs can be used to measure the effect of additional funding on the achievement differential.

Equation (3) presents the relationship between test scores and share disadvantaged that provides an estimate of the reduced form, intent-to-treat SEP revenue effects[9]:

[7] Although 77 municipalities were included in the audit, estimated expenditures were not provided in the audit report for four of these municipalities.

[8] Although schools within a municipality determined the share of SEP funds to spend within the set of permitted categories, failing to spend SEP funds appropriately was in direct violation of SEP regulations and could potentially limit future inflows. Thus, it seems likely that schools/municipalities which failed to spend SEP funds were, if anything, relatively less efficient than schools that did spend funds as required. This would in turn suggest that our estimates likely provide upper bounds of the effect of funding on test score gains for disadvantaged students.

[9] Implicitly, the first-stage dependent variable is $Bonus_{st}$, the value of concentration bonus funds received as a fraction of non-bonus SEP funds allocated to school s in year t. Since we impute bonus funds based on the concentration bonus formula, there is mechanically a sharp discontinuity in this

$$Testscore_{st} = \alpha + \beta_j Threshold_{j,st} + \tau f(Priority_{st}) + \epsilon_{ist} \quad (3)$$

In this specification, $Threshold_{j,st}$ is defined as an indicator variable for whether school s passes concentration formula threshold j in year t, $Priority_{st}$ is the school-level share of priority students (the basis for the concentration formula), and $f(Priority_{st})$ represents a local linear polynomial that is estimated separately on each side of the relevant threshold. To provide a better sense of the underlying variation in the data, Figure 2 presents a histogram of school-level priority shares and graphs the bonus funding measure as well as student test scores and mother's years of education as a function of priority student share for the first concentration bonus threshold.

We estimate separate RD specifications around each threshold with separate samples restricted to schools with values of $Priority_{st}$ share within 0.075 of the relevant cutoff. Column (1) of Table 4 reports the reduced form effects on test scores. We find no evidence that additional SEP funding leads to test score improvements based on these estimates. In Column (2), mother's years of education replaces student test score as the dependent variable and we find little evidence of a discontinuity in this student-level characteristic around the relevant thresholds, indicating that sorting cannot likely explain the lack of test score impacts. Columns (3)-(4) re-produce Columns (1)-(2) specifications for the subsample of low-SES students in SEP schools and findings are comparable. Finally, Column (5) presents results from a school-level specification that tests for manipulation (i.e. bunching) around each cutoff. The test, based on McCrary (2008), reveals evidence of manipulation only at the fourth threshold. While estimates of test score effects at threshold four should consequently be interpreted with caution, the estimates based on the three other thresholds provide consistent evidence that there is little return to additional

measure at each of the bonus thresholds. Crossing the bonus threshold is associated with increases in bonus funding of 7, 5, 4 and 2 percentage points at thresholds one through four, respectively.

SEP revenues during the study period. Moreover, the fourth threshold corresponds to the smallest jump in bonus funding. To the extent that the return to additional funds is not increasing as funds rise, we would expect to see larger effects at the lower thresholds, where differences in test scores are not statistically significant.

Table 3: OLS Estimated Effects of SEP Program Expenditure on SIMCE Scores by SES

	(1)	(2)	(3)	(4)
	2008-2011 Gains for:			
	Low SES Relative to High SES	Low SES Relative to High SES	SEP School Students	LowSES SEP School Students
% SEP Funds Spent	-0.073	-0.042	-0.029	-0.096
	(0.050)	(0.056)	(0.098)	(0.094)
Additional Controls		X	X	X
Observations	73	73	73	73

Notes: Robust standard errors are in parentheses. In Columns (1)-(2), the dependent variable is the change in the normalized test score gap between low-SES and high-SES students over the 2008-2011 period (positive values indicate convergence). Columns (3)-(4) measure average 2008-2011 gains for students in SEP schools and for low-SES students in SEP schools, respectively. Column (1)-(4) specifications are at the municipality-level and include those municipalities that were audited by the Ministry of Education in 2012 (observations are weighted by the 2008 number of fourth grade students in the municipality). In Columns (2)-(4), regressions control for municipality-level log number of students and fraction low-SES students.
* significant at 10 percent level ** significant at 5 percent level *** significant at 1 percent level.

To bring additional evidence to bear on SEP-induced changes in school inputs and teacher quality we now describe differential changes in class size and teacher characteristics for low-SES students during the study period by estimating regressions of the school and teacher characteristics on full interactions between year and low-SES dummy variables. The estimates reported in Table 5 reveal only a small decrease in relative class size for low-SES students that is at most half a student per class (Column 5). Moreover, the results suggest that teacher characteristics for low-SES students became relatively less positive following program implementation. There is a small decline in the fraction of teachers with a college degree (1 percentage point), and there is a modest (1-2 percentage point) increase in the share of inexperienced teachers (defined as teachers with less than two years of teaching experience). There are also significant increases in the share of teachers who work part time and the share that work in multiple schools. Together the results suggest that the hiring of part-time teachers who are marginally less qualified than

Table 4: RDD Specification Checks and Estimated Effects of SEP Funding on SIMCE Scores by SES

	(1)	(2)	(3)	(4)	(5)
	SEP School Students		*LowSES SEP Students*		
	SIMCE Score	Mother's Education (Years)	SIMCE Score	Mother's Education (Years)	Density Test
Threshold 1	0.007	-0.117	0.121	-0.041	0.048
	(0.074)	(0.278)	(0.121)	(0.170)	*(0.67)*
Threshold 2	-0.063	-0.072	-0.054	0.118	-0.033
	(0.044)	(0.124)	(0.051)	(0.111)	*(0.52)*
Threshold 3	-0.017	-0.039	-0.014	-0.019	0.026
	(0.043)	(0.151)	(0.036)	(0.077)	*(0.42)*
Threshold 4	0.0005	-0.260	0.013	-0.076	0.184
	(0.036)	(0.201)	(0.031)	(0.095)	*(0.001)*
Observations	805,067	748,406	344,722	353,417	43,361

Notes: The dependent variable in Columns (1), (3) is the student's normalized fourth grade test score (normalized by 2005 mean and standard deviation) and the dependent variable in Columns (2), (4) is mother's years of education. Columns (1)-(4) specifications are estimated at the student level and include data from the years 2008-2014. Each threshold refers to a given Concentration Bonus discontinuity. Specifications in Columns (1)-(4) are estimated separately for each threshold and include a local linear polynomial in Priority share that is estimated separately on each side of the relevant concentration formula threshold. For each regression estimated in Columns (1)-(4), the sample is limited to include schools with a Priority share within 0.075 of the cutoff and standard errors are clustered at the school level. Column (5) presents discontinuity estimates and corresponding p-values from school-level tests for manipulation (i.e. bunching) around each cutoff.
* significant at 10 percent level ** significant at 5 percent level *** significant at 1 percent level.

those already on staff contributed to the small reduction in class size. Yet even ignoring any decline in teacher quality, existing evidence suggests that a less than one-half student reduction in class size would have only a small effect on achievement (Rivkin et al., 2005; Krueger, 1999). Although we cannot rule out the possibility that increased spending contributed to the decline in the achievement gap, this channel clearly plays a minor role in explaining observed gains.[10] Importantly, the prohibition on using SEP funds to raise teacher salaries may have dampened gains from the program.

5.2 School Upgrading

We next test the hypothesis that the SEP program improved test scores for disadvantaged students by raising the quality of the schools that they attended. Previous research has found that post-SEP period changes in enrollment patterns were limited (Navarro-Palau, 2015). We examine this question by estimating changes in the school quality gap in

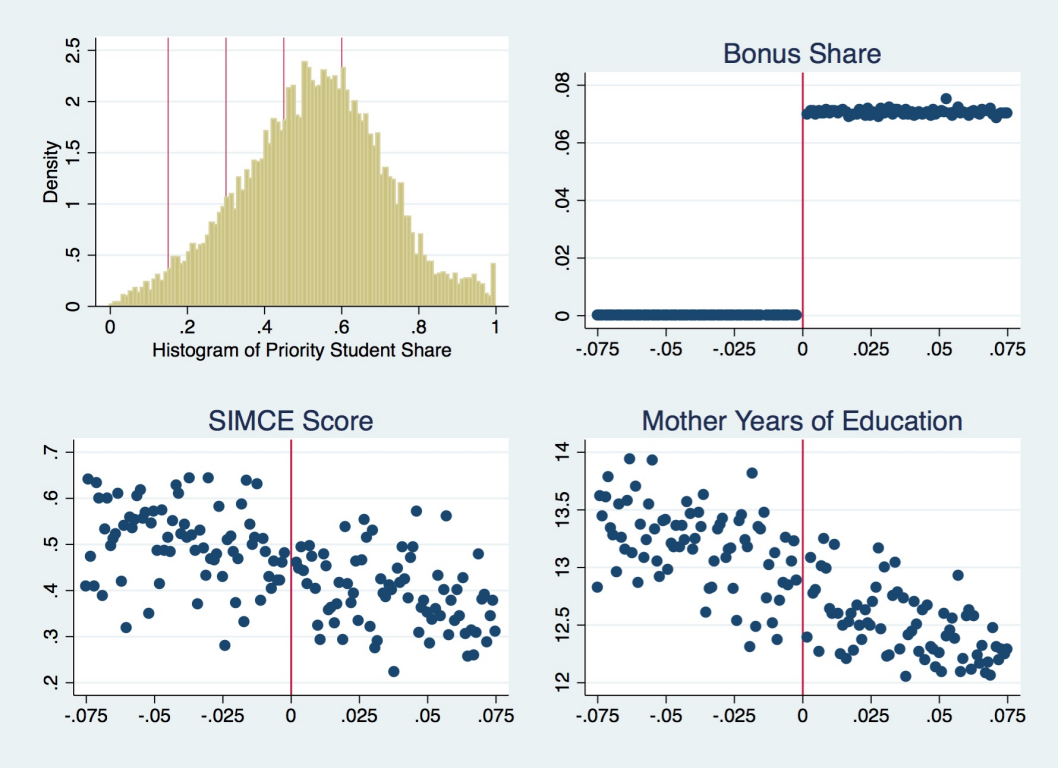
[10] Corresponding estimates for priority students are presented in Appendix Table 1. Results there indicate a relative increase in class size for priority students during the post-SEP period.

Table 5: Estimated Differences in Class Size and Teacher Characteristics for Low-SES Students, by Year

	(1)	(2)	(3)	(4)	(5)
	Percentage of Teachers with:				
	College Degree	$\leq$ 1 Year Experience	$\leq$ 20 Contract Hours	Employment in Multiple Schools	Class Size (# Students)
LowSES	-0.017***	-0.013***	-0.020***	0.001	-3.397***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.141)
LowSES ·2006	-0.0005	0.001	0.005***	-0.002	-0.122
	(0.001)	(0.002)	(0.001)	(0.002)	(0.104)
LowSES ·2007	-0.001	-0.004*	0.008***	0.004*	-0.246**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.124)
LowSES ·2008	0.001	0.008***	0.010***	0.007***	-0.296***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.114)
LowSES ·2009	0.001	0.018***	0.013***	0.009***	-0.327***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.119)
LowSES ·2010	0.0002	0.012***	0.017***	0.009***	-0.351***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.123)
LowSES ·2011	-0.001	0.005	0.021***	0.009***	-0.550***
	(0.002)	(0.003)	(0.003)	(0.002)	(0.131)
LowSES ·2012	-0.001	0.010***	0.026***	0.010***	0.062
	(0.002)	(0.003)	(0.002)	(0.002)	(0.125)
LowSES ·2013	-0.005**	0.023***	0.029***	0.009***	-0.067
	(0.002)	(0.002)	(0.002)	(0.002)	(0.126)
LowSES ·2014	-0.008***	0.019***	0.032***	0.008***	-0.192
	(0.002)	(0.003)	(0.002)	(0.002)	(0.130)
Observations	1,961,683	1,963,218	1,963,218	1,963,218	1,963,218

Notes: Standard errors are clustered at the school level and shown in parentheses. All specifications are estimated at the student-level and include data from the years 2005-2014 as well as year fixed effects. Low socioeconomic status is determined based on mother's years of education and family income as measured by SIMCE parental surveys.
* significant at 10 percent level ** significant at 5 percent level *** significant at 1 percent level.

Figure 2: Regression Discontinuity Design Histogram, First Stage, Reduced Form and Specification Check



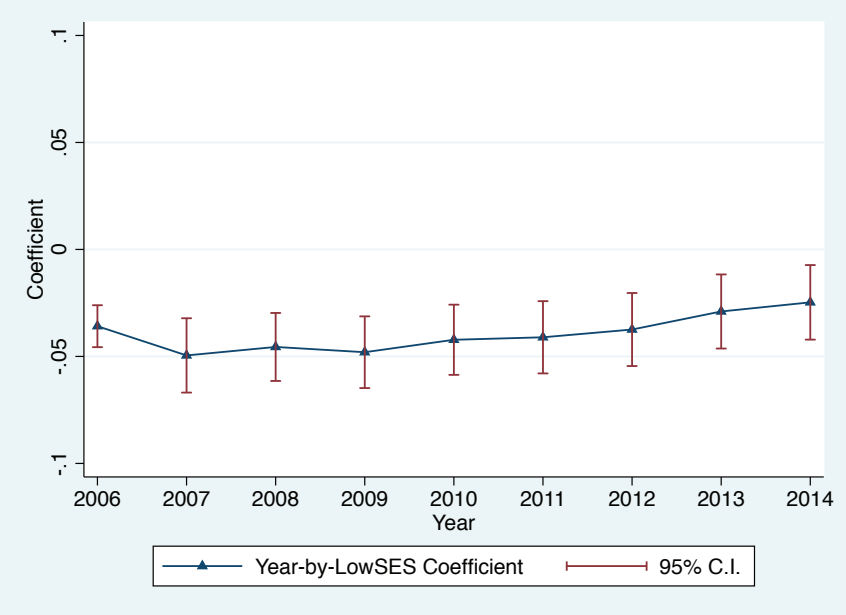
Notes: The figure displays a histogram of school-level priority shares and presents plots of concentration bonus funding, student test scores, and mother's years of education as a function of the share of priority students in the school. Except in the histogram, the first concentration bonus threshold is normalized to 0 in each panel and only data points within 0.075 priority share of the cutoff are included.

terms of 2005 test scores, i.e., school quality prior to the adoption of SEP. This approach ensures that program-induced school improvement is excluded, and only changes in the distribution of students among schools affects the school quality gap as measured by 2005 test scores.

In Figure 3, we plot coefficients and confidence intervals from a regression that re-estimates the specification employed in Table 5 while replacing the dependent variable with the average 2005 normalized fourth grade test score of the school in which a student is enrolled. The regression controls for the following student-level covariates: father's educational attainment, mother's educational attainment, and log household income. Figure

3 shows that there is no discernible improvement in the quality of schools attended by low-SES students who are in fourth grade during the post-SEP period. In the corresponding specification that excludes student-level socio-demographic controls, we observe a modest (0.05 SD) relative increase in the average baseline test scores of schools attended by low-SES students. However, none of this improvement takes place prior to 2012, by which point the majority of the 2005-2014 decline in the achievement gap has already taken place.

Figure 3: Estimated Deficits in Average School Quality for Low-SES Students, by Year



Notes: The figure presents point estimates and confidence intervals constructed from a regression in which the dependent variable is the average 2005 normalized fourth grade test score of the school in which a student is enrolled in a given year. The included regressors are the full set of interactions between year and low-SES dummy variables, and controls for father's educational attainment, mother's educational attainment, and log household income.

5.3 SEP Effects on Competition

The third channel we investigate is whether the evidence is consistent with the hypothesis that SEP reduced the achievement gap by fostering competition for low-SES students. An

increase in the voucher value associated with socioeconomically disadvantaged students increases the return to enrolling and retaining these students and should lead to greater competition. If socioeconomically disadvantaged students are informed participants in the primary school market and have multiple primary schools in their choice set, this should in turn incentivize schools that desire to enroll low-SES students to improve. Prior research on the Chilean education market suggests that the magnitude of such competitive pressure may be limited, however, given schools' market power, which is driven in part by parents' strong preferences to send their children to primary schools in close proximity to their homes (Chumacero et al., 2011; Feigenberg, 2016).

In Table 6, we test the hypothesis that socioeconomically disadvantaged students' test score improvement is driven by increased market competition using two alternative proxies for competitiveness. First, we construct a municipality-level Herfindahl Index in order to investigate whether test score gains for socioeconomically disadvantaged students are higher in markets that are less concentrated (i.e., have lower Herfindahl Index values) and so are more competitive. Second, given that municipalities with lower population density have schools that are more geographically dispersed, we test for heterogeneity in test score impacts based on whether a student resides in a municipality that is above the 50th percentile in population density. Across Table 6 specifications, we find little or no evidence of differences by level of competition in the gains made by disadvantaged students.

6 Alternative Explanations for the Closing of the Gap

We now consider alternative explanations for the test score gains of disadvantaged students, focusing on non-school inputs and strategic behavior by schools. There is widespread

Table 6: Estimated Achievement Deficits for Low-SES Students, by Intensity of School Competition, Measure of Competition, and Year

	(1)	(2)	(3)	(4)
	Population Density		Herfindahl Index	
	Above Median	Below Median	Above Median	Below Median
LowSES	-0.66***	-0.57***	-0.63***	-0.62***
	(0.01)	(0.01)	(0.01)	(0.01)
LowSES·2006	0.01	-0.001	0.01	-0.001
	(0.01)	(0.01)	(0.01)	(0.01)
LowSES·2007	-0.01*	0.01	0.02***	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
LowSES·2008	0.02**	0.04***	0.04***	0.02**
	(0.01)	(0.01)	(0.01)	(0.01)
LowSES·2009	0.03***	0.05***	0.05***	0.04***
	(0.01)	(0.01)	(0.01)	(0.01)
LowSES·2010	0.08***	0.08***	0.08***	0.09***
	(0.01)	(0.01)	(0.01)	(0.01)
LowSES·2011	0.15***	0.14***	0.15***	0.15***
	(0.01)	(0.01)	(0.01)	(0.01)
LowSES·2012	0.21***	0.17***	0.19***	0.21***
	(0.01)	(0.01)	(0.01)	(0.01)
LowSES·2013	0.22***	0.19***	0.20***	0.22***
	(0.01)	(0.01)	(0.01)	(0.01)
LowSES·2014	0.22***	0.20***	0.22***	0.23***
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	1,002,322	937,280	1,004,215	935,387

Notes: Robust standard errors are in parentheses and all specifications are estimated at the student-level. The dependent variable is the student's normalized fourth grade test score (normalized by 2005 mean and standard deviation). The Herfindahl Index is calculated at the municipality-level among low-SES fourth graders in 2005 and population density is measured based on 2002 Chilean Census data. Low socioeconomic status is determined based on mother's years of education and family income as measured by SIMCE parental surveys in 2005-2014.
* significant at 10 percent level ** significant at 5 percent level *** significant at 1 percent level.

agreement on the importance of families in the acquisition of human capital, and public programs including income supports (for which most priority students in this setting are eligible) have also been shown to influence schooling outcomes (Dahl and Lochner, 2012). Because the bottom 40 percent of the SES distribution was designated as low-SES regardless of the absolute level of parental education or income, parental education and household income differentials may have changed substantially over time. In addition, there is extensive evidence of opportunistic behavior on the part of schools in response to high-stakes testing requirements and accountability pressures (Cullen and Reback, 2006). The requirement to meet achievement targets in order to qualify for unconditional renewal of SEP funding may lead schools to attempt to raise achievement through channels including selective test-taking, teaching to the test or even outright cheating.

Table 7 describes the timing of changes in parental education and log household income for SIMCE test-takers using the same difference-in-differences approach used above but replacing SIMCE score with father's educational attainment (in years) in Column (1), with mother's educational attainment in Column (3), and with log household income in Column (5). Point estimates reveal that parental education levels and log household income increased significantly for socioeconomically disadvantaged test-takers (relative to their higher socioeconomic status counterparts) during the post-2007 period. The even-numbered columns of Table 7 re-estimate changes in parental education and household income while including school-by-year fixed effects. Estimates are similar to those in odd-numbered columns, suggesting that differences across schools do not account for the observed changes in the socio-demographic characteristics of tested students. In line with estimates presented in Correa et al. (2014), a regression of the normalized test score on the family background variables and school-by-year fixed effects shows that an additional year of mother's education is associated with a test score increase of 0.038 standard deviations, while an additional year of father's education is associated with a test score increase of

0.025 standard deviations and a one-unit change in log income is associated with a 0.078 standard deviation test score increase; all estimates are significant at the 1% significance level.[11] Based on these estimates, changes in parental education and household income can explain 0.057 standard deviations in test score gains between 2005 and 2014 for low socioeconomic status students.[12]

Table 7: Average Deficits in Parental Education and Household Income for Low-SES Students, by Year

	(1) Father's Education (Years)	(2) Father's Education (Years)	(3) Mother's Education (Years)	(4) Mother's Education (Years)	(5) Log Household Income (Pesos)	(6) Log Household Income (Pesos)
LowSES	-3.85***	-2.14***	-5.55***	-4.72***	-0.98***	-0.41***
	(0.013)	(0.015)	(0.009)	(0.009)	(0.003)	(0.003)
LowSES ·2006	0.07***	0.06***	0.11***	0.13***	0.01	-0.001
	(0.019)	(0.021)	(0.012)	(0.013)	(0.005)	(0.005)
LowSES ·2007	0.10***	0.07***	0.16***	0.17***	0.02***	0.004
	(0.020)	(0.022)	(0.013)	(0.013)	(0.005)	(0.005)
LowSES ·2008	0.11***	0.10***	0.13***	0.18***	0.03***	0.02***
	(0.019)	(0.021)	(0.012)	(0.013)	(0.005)	(0.005)
LowSES ·2009	0.14***	0.18***	0.39***	0.45***	-0.06***	-0.07***
	(0.020)	(0.022)	(0.013)	(0.014)	(0.005)	(0.005)
LowSES ·2010	0.20***	0.17***	0.49***	0.57***	-0.11***	-0.13***
	(0.019)	(0.021)	(0.013)	(0.013)	(0.005)	(0.005)
LowSES ·2011	0.22***	0.21***	0.57***	0.67***	-0.12***	-0.16***
	(0.019)	(0.021)	(0.013)	(0.014)	(0.005)	(0.005)
LowSES ·2012	0.52***	0.48***	0.67***	0.79***	0.11***	0.08***
	(0.020)	(0.021)	(0.014)	(0.013)	(0.005)	(0.005)
LowSES ·2003	0.60***	0.51***	0.73***	0.85***	0.12***	0.06***
	(0.020)	(0.021)	(0.014)	(0.013)	(0.005)	(0.005)
LowSES ·2014	0.63***	0.52***	0.93***	1.06***	0.07***	-0.02***
	(0.020)	(0.021)	(0.014)	(0.013)	(0.005)	(0.005)
School-by-Year Fixed Effects		X		X		X
Observations	1,880,008	1,880,008	1,980,260	1,980,260	1,980,260	1,980,260

Notes: Robust standard errors are in parentheses and all specifications are estimated at the student-level and include year fixed effects. Low socioeconomic status is determined based on mother's years of education and family income as measured by SIMCE parental surveys. All six columns include data from the years 2005-2014.
* significant at 10 percent level ** significant at 5 percent level *** significant at 1 percent level.

The analysis presented thus far is not able to explain fully the magnitude of fourth grade test score gains made by socioeconomically disadvantaged students during the period after the introduction of SEP. This suggests the possibility that a portion of these gains is illusory, driven by strategic behavior rather than real improvements in knowledge. Although we lack direct measures of any such behavior, comparisons with effects on lower-

[11] The sample is restricted to the pre-period years (2005-2007).

[12] Corresponding socio-demographic changes for priority students are presented in Appendix Table 2. Changes in parental and household characteristics can explain a 0.039 standard deviation test score gain for priority students between 2008 and 2014. However, observed changes in rural residency status of priority students and in whether they rank among the lowest 40% of the population based on reported mother's education and household income complicate year-on-year comparisons based on priority status.

stakes schooling outcomes and patterns of missing SIMCE data will provide evidence on the likely importance of such actions.

Table 8 reports changes over time in the SES-based gap in GPA and eighth grade SIMCE scores, two outcomes not directly connected to the SEP program. Re-estimation of the difference-in-differences model with these outcomes provides comparisons to the baseline results presented in Table 2. Point estimates in Column (1) of Table 8 show that disadvantaged students realized average GPA gains of 0.03 SD over the course of the 2008-2014 study period, where GPA is normalized by school and year. When we regress normalized student test scores on normalized GPA in a specification with school-by-year fixed effects, we estimate a coefficient of 0.6 with an R-squared of 0.6 (not shown). This estimate suggests that the decline in the GPA gap should have been associated with much smaller test score gains than we identified in Table 2. The divergence between the small decrease in the SES-based gap in GPA and the large decrease in the fourth grade SIMCE gap is striking.

In contrast, changes in the gap in eighth grade SIMCE scores presented in Columns (3) and (4) of Table 8 line up much more closely with the changes in GPA. Data are available for the eighth grade cohorts which took the exam in 2007, 2009, 2011, 2013, and 2014. Column (3) estimates test score changes for students who appear in both fourth and eighth grade in the SIMCE sample and defines socioeconomic status based on fourth grade survey responses, while Column (4) estimates test score changes for all eighth grade students and defines socioeconomic status based on eighth grade survey responses. Regardless of the sample, the results show that low-SES students exposed to the SEP program as fourth graders (i.e., those in eighth grade in 2012 or later) experience relative test score gains of between 0.02 and 0.07 SD. This finding further reinforces the notion that the substantial improvement of disadvantaged students on the fourth grade SIMCE tests cited in the previous literature as indicating a sizable program effect is largely illusory and

not evidence of substantial gains in the relative academic skills of disadvantaged children.

Table 8: Estimated GPA and Eighth Grade SIMCE Test Score Deficits for Low-SES Students, by Year

	(1) GPA	(2) GPA	(3) SIMCE (Grade 8)	(4) SIMCE (Grade 8)
LowSES	-0.25***	-0.12***	-0.54***	-0.53***
	(0.004)	(0.005)	(0.004)	(0.004)
LowSES·2006	0.01	0.004		
	(0.01)	(0.01)		
LowSES·2007	0.005	0.0002		
	(0.01)	(0.01)		
LowSES·2008	0.03***	0.03***		
	(0.01)	(0.01)		
LowSES·2009	0.05***	0.04***		-0.03***
	(0.01)	(0.01)		(0.01)
LowSES·2010	0.05***	0.04***		
	(0.01)	(0.01)		
LowSES·2011	0.08***	0.06***	0.03***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)
LowSES·2012	0.04***	0.02***		
	(0.01)	(0.01)		
LowSES·2013	0.03***	0.01	0.06***	0.07***
	(0.01)	(0.01)	(0.01)	(0.01)
LowSES·2014	0.03***	0.01	0.02***	0.06***
	(0.01)	(0.01)	(0.01)	(0.01)
Additional Controls		X		
Observations	1,962,854	1,863,598	589,009	928,965

Notes: Standard errors clustered at the school-level are presented in parentheses. All specifications are estimated at the student-level and include year fixed effects. Low socioeconomic status is determined based on mother's years of education and family income as reported in SIMCE parental surveys. All columns include data from the years 2005-2014. Low socioeconomic status is determined based on fourth grade survey data in Column (3) and based on eighth grade survey data in Column (4). Additional controls includes controls for mother's years of education, father's years of education, and log household income.
* significant at 10 percent level ** significant at 5 percent level *** significant at 1 percent level.

The data do not allow us to distinguish between direct gaming of the SIMCE exam and short-term improvements that manifested themselves as large fourth grade test score gains that disappeared by eighth grade. However, we can use GPA information that is available for all students to impute missing SIMCE scores and estimate the contribution of missing scores to the closing of the gap. Importantly, information on SES is not available for those with missing scores. Therefore we use attendance at a public school as a proxy for low-SES status. The noisiness of the proxy will attenuate the differences, but the comparison between trends based on all students and those based on students with non-missing scores will illuminate the contribution of missing data to the observed decline in the achievement gap.

The first column in Table 9 reports coefficients on interactions between public sector and year from a linear probability model that regresses an indicator for a missing score

on year dummies, a public school dummy and their interactions. The coefficients show an increase in the missing rate in public schools relative to private schools following the SEP reform that ranges between three and nine percentage points. The effects of an increase in missing data depend upon both the incidence and composition of students with missing tests, and we make use of the GPA information to estimate the effects of missing examinations on the observed closing of the achievement gap. Specifically, we impute test scores for all students using school-specific estimates of the linear relationship between SIMCE score and GPA for those with non-missing scores (school-specific estimates are constructed using data from the pre-2008 period). Changes over time in the achievement gap for the sample of students with non-missing scores can then be compared with changes for the full sample of students. Column 2 reports coefficients on the interactions between public and year dummies from a regression of SIMCE score on a public dummy, year dummies and their interactions for the sample of students with non-missing data, while Column 3 reports the same coefficients from a regression that uses imputed SIMCE score as the dependent variable. A comparison of the coefficients suggests that more than half of the relative gains observed for public school students during the SEP period can be explained by changes in the composition of test-takers, as estimates in Column 3 tend to be roughly half as large as those in Column 2 in the period following the full implementation of SEP.

7 Conclusion

Although the SES-based fourth grade SIMCE test score gap decreases by roughly 0.2 standard deviations following the implementation of the SEP program, our analysis does not support the belief that the SEP program had a substantial impact on the corresponding school-quality gap. Neither increases in school expenditure nor school quality upgrading

Table 9: Public-Private School Differences in the Rate of Missing Test Scores and Estimated Achievement Deficits for Low-SES Students by Treatment of Missing Test Scores and Year

	(1) Missing SIMCE(0/1)	(2) Actual SIMCE	(3) Imputed SIMCE
Public	0.02***	-0.47***	-0.49***
	(0.001)	(0.004)	(0.004)
Public·2006	-0.001	0.01	0.01
	(0.002)	(0.01)	(0.01)
Public·2007	0.001	-0.04***	-0.03***
	(0.002)	(0.01)	(0.01)
Public·2008	0.01***	-0.03***	-0.03***
	(0.002)	(0.01)	(0.01)
Public·2009	0.09***	-0.01*	-0.04***
	(0.002)	(0.01)	(0.01)
Public·2010	0.03***	0.03***	-0.002
	(0.002)	(0.01)	(0.01)
Public·2011	0.04***	0.10***	0.05***
	(0.002)	(0.01)	(0.01)
Public·2012	0.03***	0.09***	0.05***
	(0.002)	(0.01)	(0.01)
Public·2013	0.04***	0.08***	0.04***
	(0.002)	(0.01)	(0.01)
Public·2014	0.04***	0.07***	0.03***
	(0.002)	(0.01)	(0.01)
Observations	2,378,699	2,115,350	2,343,523

Notes: Standard errors clustered at the school-level are presented in parentheses. All specifications are estimated at the student-level. All columns include data from the years 2005-2014. Imputed test scores in Column (3) are predicted for missing observations based on student GPA and a school-specific estimate of the linear relationship between GPA and test scores.
* significant at 10 percent level ** significant at 5 percent level *** significant at 1 percent level.

appears to explain much if any of the apparent gains for low-SES students. Rather, the evidence suggests that convergence in family background characteristics of tested students can explain a meaningful share of the high-stakes fourth grade gains and virtually all of the declines in SES-based differences in GPA and eighth grade test scores, neither of which are high-stakes outcomes from the perspective of SEP participant schools.

The crucial questions for policy concern the lack of impact of the SEP reform on the academic outcomes of disadvantaged children. Specifically, it is critical to understand the relative importance of: (1) the lack of integrity of the policy implementation which caused the increase in validated school expenditures to be far smaller than the increase in revenues; and (2) the failure of the performance incentives to alter behavior in ways that improved the quality of instruction and learning for disadvantaged students.

Alternative explanations with different policy implications come to the forefront, and their divergent implications for policy highlights the importance of gaining a clear understanding of their contributions. First, program rules may have compromised program effectiveness. These include a prohibition on using the SEP funds to raise teacher salaries to attract more effective educators and a focus on fourth grade SIMCE scores that were too easily susceptible to strategic behavior. The adverse effects of these and other deficiencies in program structure may have been amplified by weak monitoring and enforcement.

Alternatively, it is possible that such a major reform requires time to take effect, as found in a study of Texas charter-school reforms (Baude et al., 2014). However, the absence of marked improvements in school quality for low-SES children and the limited market entry of new voucher schools serving low-income areas even five years after program implementation raises doubts that the program will have a large effect over the longer-term. An alternative explanation emphasized in Feigenberg (2016) suggests that the market power enjoyed by schools in a system in which many parents seem unwilling or unable to respond to differences in school quality is likely to dampen the benefits of

programs designed to raise school competition for disadvantaged children.

In sum, our findings indicate that the Chilean SEP experiment was not nearly as promising as it appeared and that additional evidence is needed on the question of whether targeted voucher policies can effectively serve those students most in need. Understanding the extent to which the price mechanism can be employed within educational markets like Chile's in order to mitigate adverse features of these markets remains an open question in the academic literature and one that is of first-order importance to educational policy-makers who seek to better understand the tradeoffs associated with voucher systems.

References

Baude, P., M. Casey, E. Hanushek, and S. Rivkin (2014). The Evolution of Charter School Quality. NBER Working Paper 20645.

Bettinger, E. (2011). Educational Vouchers in International Contexts. In E. A. Hanushek, S. Machin, and L. Woessmann (Eds.), *Handbook of the Economics of Education* (1 ed.), Volume 4, pp. 551–572. Elsevier.

Bravo, D., S. Mukhopadhyay, and P. E. Todd (2010). Effects of School Reform on Education and Labor Market Performance: Evidence from Chile's Universal Voucher System. *Quantitative Economics 1*(1), 47–95.

Card, D., M. Dooley, and A. Payne (2010). School Competition and Efficiency with Publicly Funded Catholic Schools. *American Economic Journal: Applied Economics 2*(4), 150–176.

Chilean Comptroller's Office (2012). Final SEP Report.

Chumacero, R., D. Gomez, F. Parro, and R. Paredes (2011). I Would Walk 500 Miles (if it Paid): Vouchers and School Choice in Chile. *Economics of Education Review 30*(5), 1103–1114.

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

Cullen, J. and R. Reback (2006). Tinkering Toward Accolades: School Gaming under a Performance Accountability System. In T. Gronberg and D. Jansen (Eds.), *Improving School Accountability*, Volume 14. Emerald Group Publishing Limited.

Dahl, G. and L. Lochner (2012). The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit. *American Economic Review 102*(5), 1927–1956.

Feigenberg, B. (2016). Priced Out: Aggregate Income Shocks and School Pricing in the Chilean Voucher Market. Working Paper.

Friedman, M. (1962). *Capitalism and Freedom*. Chicago: University of Chicago Press.

Gallego, F. (2013). When Does Inter-School Competition Matter? Evidence from the Chilean "Voucher" System. *B.E. Journal of Economic Analysis & Policy 13*(2), 525–562.

Hanushek, E. A., P. Peterson, and L. Woessmann (2012). Achievement Growth: International and U.S. State Trends in Student Performance. *Harvard's Program on Education Policy and Governance Education Next*.

Hoxby, C. (2000). Does Competition Among Public Schools Benefit Students and Taxpayers? *American Economic Review 90*(5), 1209–1238.

Hsieh, C.-T. and M. Urquiola (2006). The Effects of Generalized School Choice on Achievement and Stratification: Evidence from Chile's Voucher Program. *Journal of Public Economics 90*, 1477–1503.

Krueger, A. (1999). Experimental Estimates of Education Production Functions. *Quarterly Journal of Economics 114*(2), 497–532.

Lavy, V. (2010). Effects of Free Choice Among Public Schools. *The Review of Economic Studies 77*(3), 1164–1191.

McCrary, J. (2008). Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics 142*(2), 698–714.

McEwan, P. (2001). The Effectiveness of Public, Catholic, and Non-Religious Private Schools in Chile's Voucher System. *Review of Educational Research 9*(2), 103–128.

Murnane, R., M. Waldman, M. S. Bos, E. Vegas, and J. Willett (2016). The Consequences of Education Voucher Reform in Chile. Poster Paper.

Navarro-Palau, P. (2015). Effects of Differentiated School Vouchers: Evidence From a Policy Change and Date of Birth Cutoffs. Working Paper.

Neilson, C. (2013). Targeted Vouchers, Competition among Schools, and the Academic Achievement of Poor Students. Working Paper.

Preferential School Voucher (SEP) Law of 2008. Law 20248.

Rivkin, S., E. Hanushek, and J. Kain (2005). Teachers, Schools, and Academic Achievement. *Econometrica 73*(2), 417–458.

Urquiola, M. (2016). Competition Among Schools: Traditional Public and Private Schools. In E. A. Hanushek, S. Machin, and L. Woessmann (Eds.), *Handbook of the Economics of Education*, Volume 5, pp. 209–237. Elsevier.

8 Appendix

Table A1: Estimated Differences in Class Size and Teacher Characteristics for Priority Students, by Year

	(1)	(2)	(3)	(4)	(5)
	Percentage of Teachers with:				
	College Degree	$\leq$ 1 Year Experience	$\leq$ 20 Contract Hours	Employment in Multiple Schools	Class Size (# Students)
Priority	-0.005**	-0.01***	-0.01***	-0.002	-4.48***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.16)
Priority·2009	-0.01***	0.01***	0.01***	0.01***	1.67***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.13)
Priority·2010	-0.01***	0.01***	0.01***	0.01***	1.75***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.14)
Priority·2011	-0.01***	-0.001	0.01***	0.01***	1.98***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.15)
Priority·2012	-0.01***	0.01*	0.02***	0.01***	1.97***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.23)
Priority·2013	-0.02***	0.02***	0.02***	0.01***	1.56***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.35)
Priority·2014	-0.02***	0.02***	0.02***	0.01***	1.85***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.22)
Observations	1,679,675	1,680,329	1,680,329	1,680,329	1,680,329

Notes: Standard errors are clustered at the school level and shown in parentheses. All specifications are estimated at the student-level and include data from the years 2008-2014 as well as year fixed effects. The priority status of a student is designated by the Ministry of Education.
* significant at 10 percent level ** significant at 5 percent level *** significant at 1 percent level.

Table A2: Average Differences in Parental and School Characteristics for Priority Students, by Year

	(1) Father's Education (Years)	(2) Mother's Education (Years)	(3) Log Household Income (Pesos)	(4) LowSES	(5) Attend Rural School
Priority	-2.91***	-2.91***	-0.89***	0.39***	0.16***
	(0.04)	(0.04)	(0.01)	(0.004)	(0.005)
Priority ·2009	0.24***	0.36***	-0.04***	-0.04***	-0.04***
	(0.03)	(0.03)	(0.01)	(0.004)	(0.003)
Priority ·2010	0.42***	0.45***	0.02**	-0.04***	-0.04***
	(0.03)	(0.03)	(0.01)	(0.004)	(0.003)
Priority ·2011	0.49***	0.65***	0.04***	-0.06***	-0.05***
	(0.03)	(0.03)	(0.01)	(0.004)	(0.004)
Priority ·2012	0.55***	0.73***	0.06***	-0.11***	-0.05***
	(0.03)	(0.03)	(0.01)	(0.004)	(0.004)
Priority ·2013	0.27***	0.46***	-0.01	-0.10***	-0.05***
	(0.03)	(0.04)	(0.01)	(0.004)	(0.004)
Priority ·2014	0.45***	0.65***	0.07***	-0.10***	-0.06***
	(0.03)	(0.04)	(0.01)	(0.004)	(0.004)
Observations	1,325,659	1,383,655	1,390,764	1,371,026	1,696,783

Notes: Robust standard errors are in parentheses and all specifications are estimated at the student-level and include year fixed effects. Priority student status is determined by the Ministry of Education for the years 2008-2014.
* significant at 10 percent level ** significant at 5 percent level *** significant at 1 percent level.