Online Appendix
Screening and Recruiting Talent At Teacher Colleges*

Christopher A. Neilson † Sebastian Gallegos ‡ Franco Calle §

Mohit Karnani ¶

This version: February 7, 2022

*Usual disclaimers apply.

†Princeton University, and NBER cneilson@princeton.edu
‡UAI Business School, sebagallegos@uchicago.edu
§University of Chicago, francocalle@chicagobooth.edu
¶MIT, karnani@mit.edu

Contents

List of Figures

List of Tables

A Teacher Exit Exam

A.1 Description and Implementation

The teacher exit exam, called 'INICIA', consists of a set of tests taken by newly graduated teachers, implemented for the first time in year 2009 and up to 2015.[1] The exam intends to measure four dimensions: (1) disciplinary knowledge (e.g. math knowledge for math teachers); (2) pedagogical knowledge (intended to measure if test takers can explain concepts in a coherent way); (3) writing skills, and (4) capacity to use ICT (information and communication technologies) for teaching purposes. In 2016 the ministry of education administered 'Diagnostica' which also evaluates disciplinary and pedagogical skills and is taken the year before graduation in different universities.

INICIA and Diagnostica's main objective is to assess the qualification of recent teacher graduates. The information produced by the exit exams is thought to be useful for the institutions training teachers, policy makers and the test-takers themselves, although there are no associated consequences to its results.[2] Results are published at the institution level, with individual-level information remaining confidential. The exam's application was gradually expanded by year and by the level at which teachers specialize (i.e. pre-school, primary and secondary), as summarized in Table 1.[3]

[1] Typically, students in their last semester of class, or just graduated students taking the exam before getting a job.

[2] This may change in the near future. During 2015, the Ministry of Education sent a bill to the Congress in order to make the INICIA test mandatory and to establish minimum performance levels to be allowed to teach at least in the public sector.

[3] In 2013 the exam was not applied.

Table 1: Teacher Exit Exam: Tests Implemented by Year and Teacher Specialization Level

		Year						
Level	Test	2009	2010	2011	2012	2014	2015	2016
Pre-school	Disciplinary	✓	✓	✓	✓	✓	✓	✓
	Pedagogical		✓	✓	✓	✓	✓	✓
	Writing	✓	✓	✓	✓			
	ICT							
Primary	Disciplinary	✓	✓	✓	✓	✓	✓	✓
	Pedagogical	✓	✓	✓	✓	✓	✓	✓
	Writing	✓	✓	✓	✓			
	ICT	✓	✓	✓				
Secondary	Disciplinary				✓	✓	✓	✓
	Pedagogical				✓	✓	✓	✓
	Writing				✓			
	ICT							

Notes: 'Disciplinary' stands for the test measuring disciplinary knowledge; 'Writing' stands for the writing skills test; 'Pedagogical' stands for the Pedagogical knowledge test; 'ICT' stands for the test measuring the information and communications technology skills. Source: MINEDUC (2012).

The Inicia exam is voluntary. Formally, the Ministry of Education invites graduate institutions that train teachers (i.e. institutions offering teacher or education degrees) to participate in the INICIA exam every year. In the case of Diagnostica, the exam is mandatory and is administered to all students of pedagogy in certified institutions of education. Table 2 summarizes the number of institutions invited, and those that participated. From years 2009 to 2012, around 80% of the invited institutions participated, which means that at least some of its graduates took the test. Institutions can encourage their graduates to participate, but can not force them to do so.

Table 3 summarizes the number of potential test-takers, the ones that sign-up and those that take at least one test, by year. Every INICIA test before 2012 was held in December of each year, which coincides with the end of the academic year in Chile. Due to administrative issues, the 2012 INICIA test was held in April of 2013. This delay seems to be the reason behind the low take-up of that year's test (see Table 3). By that time of the year, most new teachers would be working, because the academic year starts in March. Also, it is likely that graduates lose the connection with their universities after a while. After this episode, the Ministry of Education decided to postpone the application of the 2013 INICIA, supposed to be held in December 2013, to December of 2014, combining evaluations 2013 and 2014 into a single sitting.

Table 2: Teacher Exit Exam: Invited and Participating Institutions by Year

	Application	Number of Institutions		Participation
Year	Date	Invited	Participating	Percentage
2009	Dec. 2009	54	43	80%
2010	Dec. 2010	56	43	77%
2011	Dec. 2011	59	49	83%
2012	Apr. 2013	58	50	86%
2014	Apr. 2014		50	-
2015	Dec. 2015		50	-
2016*	Apr. 2016		50	-

Notes: Invited institutions correspond to those that train primary school teachers (every year), pre-school teachers (years 2009-2012 and 2016) and secondary school teachers (year 2012 and 2016). Participating institutions are the ones for which at least one of their graduates takes one or more of the tests described in Table 1. Participation percentage displays the number of participating institutions as a percentage of the number of invited institutions. * In 2016 the corresponding exit exam was Diagnostica and was mandatory. Source: MINEDUC (2012).

A.2 Teacher Exit Exam Results and PSU Scores

Institutional Reports. The Ministry of Education publishes each year a presentation with the INICIA exam results.[4] According to these institutional reports, the results achieved by the education graduates are below what is needed to perform adequately as a teacher.

[4]For years 2008 to 2010, results were mainly published as the percentage of correct answers achieved in each test, without a statement on what was considered a good outcome. For the 2011 and later exams, the MINEDUC implemented three categories to classify test-takers according to their performance, based on the knowledge and skills necessary to begin their career as a classroom teachers: Outstanding, Acceptable and Unsatisfactory.

Table 3: Teacher Exit Exam: Test-Takers by Year

	Number of Test-Takers			
Year	Potential	Signed-up	Participated	Take-up
2008	5,250	3,006	1,994	38%
2009	7,979	4,527	3,223	40%
2010	8,594	4,681	3,616	42%
2011	8,069	4,874	3,271	41%
2012	10,351	2,443	1,443	14%
2014	15,013	714	682	4%
2015	14,472	1,993	1,916	13%
2016*	20,215	17,971	12,741	63%

Notes: Potential Test-Takers correspond to the number of graduates from previous year. Those that sign-up to take the test are displayed in column 2. The number individuals that took at least one test described in Table 1 is shown in column 3. Column 4 presents the number of actual test-takers as a percentage of the potential test-takers. * In 2016 the corresponding exit exam was Diagnostica and was mandatory. Source: MINEDUC (2012).

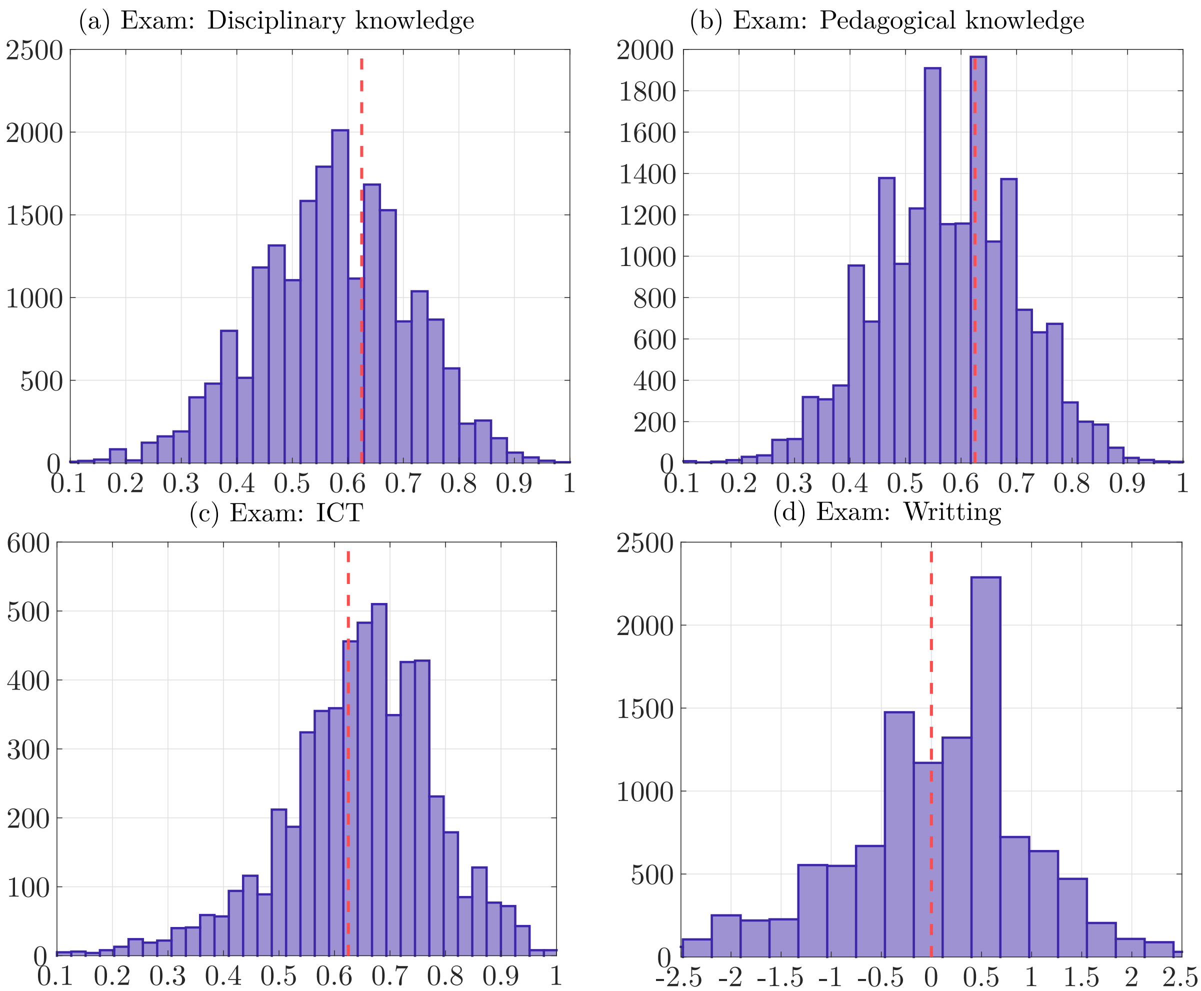
More than 60% of the test-takers that graduated as primary teachers fall in the 'unsatisfactory' category for the disciplinary tests in 2011 and 2012. The percentage is approximately 40% for the pedagogical test. For secondary teachers, the disciplinary tests by subject show the worse results in Mathematics, Biology, Physics and Chemistry, where about 70% of the test-takers fall in the 'unsatisfactory' category.

Microdata. The Ministry of Education provided us information from 2009 to 2015 on the INICIA exam, at the individual level and 2016 data on Diagnostica exam. We have microdata for more than 16K teachers with INICIA scores in at least one test and arround 13k teachers evaluated in Diagnostica. Table 4 provides summary statistics for the four available tests. The first three rows report the percent of correct answers for the disciplinary, pedagogical and the ICT tests.[5] The last row shows the scores in the standardized writing test.

Figure 1 shows histograms for the four tests, where the vertical dashed line indicates the

[5]The difference in the samples is explained by the fact that the Pedagogical test was not held in 2009, and the ICT test was not applied in 2012.

Figure 1: Exit Exams Histograms



Note: All four figures use information for teachers that took the respective tests from the INICIA exit exam between years 2009 and 2014. The dashed red line indicates the cutoff above which the performance in each test is considered 'acceptable'. These cutoffs are 0.61, 0.61, 0.65 and -0.09 for the Disciplinary, Pedagogical, ICT and Writing tests, respectively. Cutoff values vary slightly over the years, so they should be interpreted as proxies. The figures consider all test-takers with valid scores in the Disciplinary test (Figure 1(a), $N = 12,477$), the Pedagogical test (Figure 1(b), $N = 8,943$), the ICT test (Figure 1(c), $N = 6,249$) and the Writing test (Figure 1(d), $N = 10,665$).

Table 4: Exit Exam Summary statistics

Variable	Mean	Std. Dev.	Min	Max	N	Corr(PSU)
% of Correct Answers in:						
Disciplinary Test	0.57	0.14	0	1	20224	0.53
Pedagogical Test	0.58	0.13	0.05	1	18025	0.51
ICT Test	0.65	0.14	0.1	1	5517	0.51
Writing Test*	0.02	0.99	-6.65	3.26	11300	0.28

Notes: the last column displays the displays Spearman's rank correlations for each variable and PSU scores. The percentage of correct answers for the Disciplinary, Pedagogical and ICT tests has an associated a cutoff above which the performance is considered 'acceptable'. These cutoffs are .61, .61 and .65 for the Disciplinary, Pedagogical and ICT tests, respectively. These thresholds vary slightly over years, so they should be interpreted as proxies. For the writing test score, the cutoff is about -.09 SD from the mean. * Writing test is a standardized variable of the scores achieved by students by years.

cutoff above which the performance is considered acceptable.[6] Test-takers perform poorly: in the Disciplinary test, 62% of the test takers are *below* the threshold. For the Pedagogical test, the percentage is 58%. For the ICT and Writing tests, 39% and 42% of the test-takers have scores below acceptable.

Exit exam test results are strongly correlated with PSU scores, as suggested by the Spearman's rank correlations in the last column in Table 4 and nonparametric plots of the bivariate relation in Figure 2. The positive correlation ocurs in the whole spectrum of the PSU score according to the figures.

Bear in mind that the PSU test is administered right before beginning higher education studies, and the exit exams are administered right after completing their degree.

B Teacher Evaluation in the Public Sector

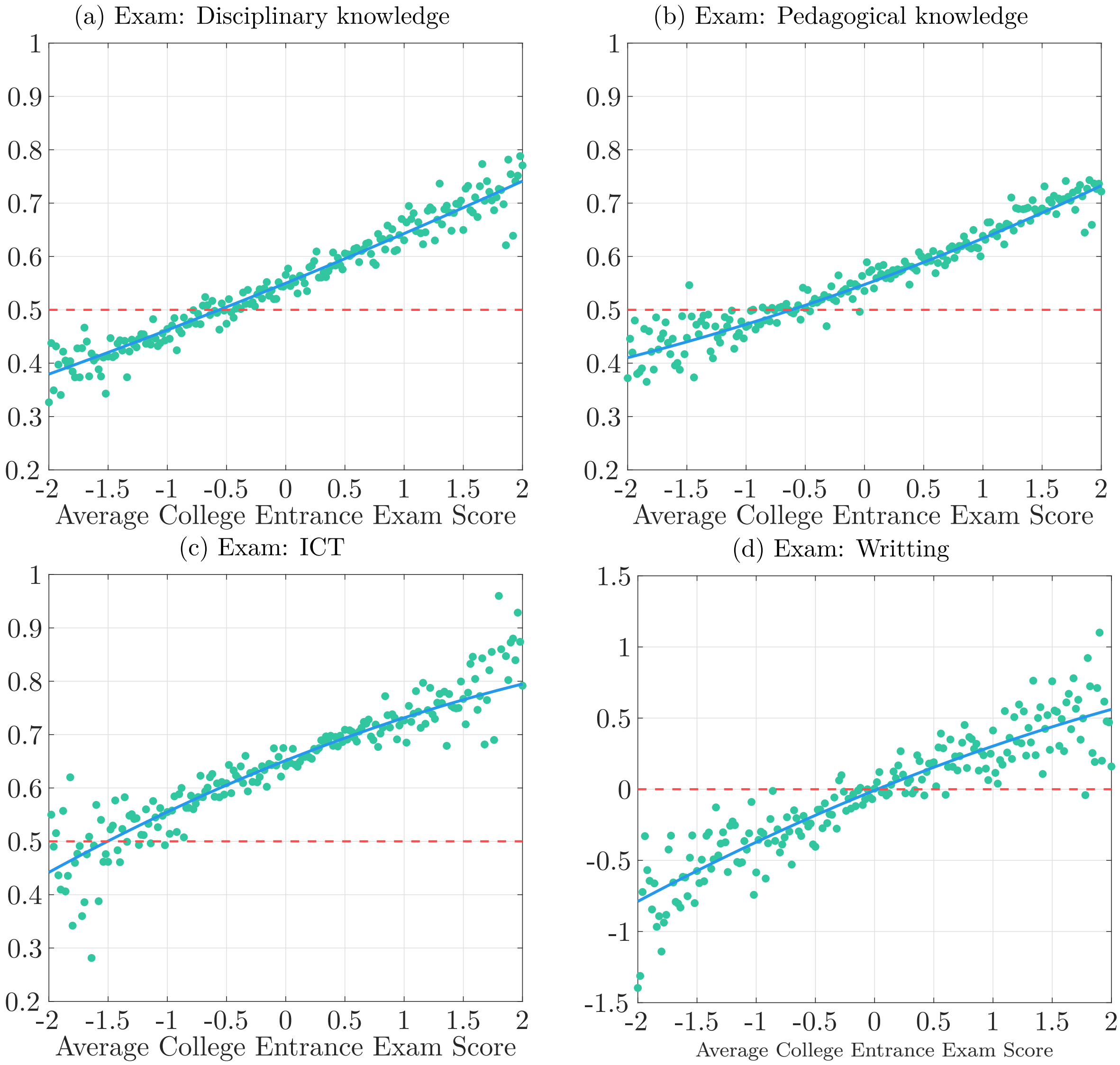
B.1 Teacher Evaluation Description and Implementation

The Teacher's Public Evaluation System[7] (*Evaluación Docente* in Spanish, or ED onwards) is a mandatory assessment for all classroom teachers working in the public sector in Chile.

[6]This cutoffs vary slightly by year, so they should be considered general guidelines and not absolute thresholds.

[7]For more details, see the institutional website www.docentemas.cl.

Figure 2: Exit Exams vs PSU scores



Note: All four figures use information of teachers that took the respective Inicia exit exam test between years 2009 and 2014. PSU scores have a mean of 500 points and a standard deviation (SD) of 100 points, so each plot's x-axis shows ± 2 SD. Solid blue lines correspond to fitted values from local linear regressions using a rectangular kernel with a bandwidth of 10 PSU points, with 95% confidence intervals plotted in gray. The open circles plot the average values of each variable within five points of the PSU score. The dashed red line indicates the cutoff above which the performance in each test is considered acceptable; these cutoffs are .61, .61 and .65 for the Disciplinary, Pedagogical and ICT tests, respectively and -.09 for the Writing test. The cutoffs vary +-slightly over years, so they should be interpreted as proxies. The Figures consider all test-takers with valid scores in the PSU and the Disciplinary test (Figure 2(a), $N = 11,060$), the Pedagogical test (Figure 2(b), $N = 7,447$), the ICT test (Figure 2(c), $N = 5,795$) and the Writing test (Figure 2(d), $N = 9,908$).

The ED declared objective is 'to strengthen the teaching profession and the quality of education'. The assessment is composed by four components, with different weights: (i) a self-evaluation questionnaire (10%); (ii) a third-party reference report, filled by the school principal or supervisor (10%); (iii) one peer review (20%), and a *teacher performance portfolio* (60%). The portfolio component aims to collect direct evidence on teaching skills, pedagogical decisions and classroom practice. It includes two modules. In the first module, teachers plan a class defining its contents and related assessments. They are also asked questions about teaching practices. The second module consists in a videotaped class followed by a questionnaire on the students behavior and understanding, and the teacher's own performance.

The ED assigns a weighted score for each teacher using the components (i) to (iv) above. Then, the score is used to classify each teacher performance in one of four categories: unsatisfactory, basic, competent or outstanding. As opposed to the INICIA exit exam, the ED has consequences associated to performance. Teachers classified in the 'competent' or 'outstanding' categories can opt to receive a monetary bonus. Teachers classified in the unsatisfactory level need to retake the ED. If they remain in the unsatisfactory category after three times, they must leave their schools and can not teach again.

The ED has been implemented gradually since 2004 according to the level at which teachers specialize (pre-school, primary, secondary).[8] Table 5 shows its year-level coverage for ten years 2004 to 2016.

[8] There are also other levels that have been incorporated to the teacher evaluation, like special education and education for adults, but we focus on preschool, primary and secondary levels in our analysis.

Table 5: Teacher Evaluation Implementation by Year and Level Taught

	Level		
Year	Preschool	Primary	Secondary
2004		✓	
2005		✓	✓
2006		✓	✓
2007		✓	✓
2008	✓	✓	✓
2009	✓	✓	✓
2010	✓	✓	✓
2011	✓	✓	✓
2012	✓	✓	✓
2013	✓	✓	✓
2014	✓	✓	✓
2015	✓	✓	✓
2016	✓	✓	✓
2017	✓	✓	✓

Notes: There are also other levels that have been incorporated to the teacher evaluation, like special education and education for adults, but we focus on primary and secondary levels in our analysis.

The ED has carried out more than 174207 assessments for preschool, primary and secondary teachers from 2004 to 2013. Table 6 exhibits the number of evaluations per teacher by year. The system has evaluated 101423 teachers at least once.[9] Approximately half of those teachers have been evaluated twice[10] (∼51K), and a about 35K have been evaluated more than three times.

For purposes of the analysis we will restrict the sample to teachers of primary or secondary education that were evaluated. This sample consist on 78513 teachers from the total of 101K evaluated (%77 of the total sample). Table 7 reports the first ED results per

[9] From the 101K evaluated teachers a fraction has already retired from teaching. To get a sense of the coverage regarding those working currently in the public sector, consider that in year 2016 130K classroom teachers were working in the public sector (in either the preschool, primary or secondary level) and about 101K of them (∼78%) had been evaluated at least once.

[10] All teachers are supposed to be re-evaluated every four years, which the data does not fully support; teachers first classified in the unsatisfactory or basic category should be re-evaluated the year after or two years after the first evaluation respectively.

Table 6: Number of times teachers were evaluated from 2004-2013

Year	N:1	N:2	N:3	N:4	N:5	N:6
2004	1719	0	0	0	0	0
2005	10631	34	0	0	0	0
2006	13931	255	4	0	0	0
2007	10178	208	27	0	0	0
2008	14890	1104	21	0	0	0
2009	8567	5524	25	0	0	0
2010	3873	6422	121	3	0	0
2011	3498	7274	158	9	0	0
2012	3875	10496	693	17	0	0
2013	4343	6447	3818	57	4	0
2014	4993	3536	5118	167	7	0
2015	4620	2828	3889	339	16	0
2016	5707	3229	6118	899	35	1
2017	6667	3657	4757	2080	47	2
All	101423	50744	20456	1518	65	1

Notes: The table above represent the number of tests administered each year by the number of times a teacher was evaluated until each year.

category for all the 78.5K teachers in its first column. Only a 2 percent of the teachers resulted in an 'unsatisfactory' performance; 28% were classified as 'basic', 61% as 'competent' and 9% as 'outstanding'. It also shows the maximum scored achieved by category for some years. The thresholds to be in each category vary by year.

Table 7: Teacher Evaluation Results 2004-2016

Classification	N obs	%	Max: 2004	Max: 2008	Max: 2012	Max: 2016
Outstanding	6875	8.8	3.63	3.59	3.21	3.37
Competent	48130	61.3	3.11	3.25	3	3.15
Basic	22091	28.1	2.64	2.67	2.79	2.9
Usatisfactory	1417	1.8	2	2.26	1.95	2.1
Total:	78513	100	2.84	2.94	2.74	2.88

B.2 Teacher Evaluation Results and PSU Scores

From the sample of first test taken by primary and secondary teachers we examine the correlation between ED and PSU scores. From the 78,513 teachers of primary and secondary education with ED scores about 63K (or 81%) have an available PSU score, while 14974 (or 19%) have not. As we explained in detail in the PSU Section, we collected data on the national college exam (PSU) that teachers took up to 35 years ago (from 1980 onwards). Therefore, we do not have information for the older teachers, many of whom have retired from teaching anyway. On average, the teachers with ED scores but no PSU scores were 61 years old in 2016, and a 44% of them was not teaching during year.

Table 8 shows the teacher evaluation results by availability of PSU scores. Panel A shows the results by the four categories of performance. We also have information on the overall ED score and also the portfolio component score, whose results we present in Panel B of Table 8.

Teachers with PSU scores tend to perform better in the ED. Panel A shows that they fall more in the upper two categories (competent and outstanding) and less in the lower (basic and unsatisfactory). Consistently, teachers with PSU scores also achieve higher ED scores, both overall and in the portfolio component as shown in Panel B. Differences in both Panels are significant at the 1% level.

Figure 5 shows the distributions of the overall ED scores and portfolio scores (Figure 5(a) and Figure 5(b), respectively) for teachers with and without PSU scores. The vertical dashed lines indicate the scores that separate the four categories.[11] Even though differences do not appear distinguishable to the eye, a two-sample Kolmogorov-Smirnov test rejects equality of distribution functions for each score. In any case, given the positive relationship between the ED and PSU scores that we document next, we expect teachers without PSU information to have lower PSU scores.

[11] The cutoffs vary slightly over years, so they should be interpreted as proxies.

Table 8: Teacher Evaluation Results 2004-2013, by PSU Score availability

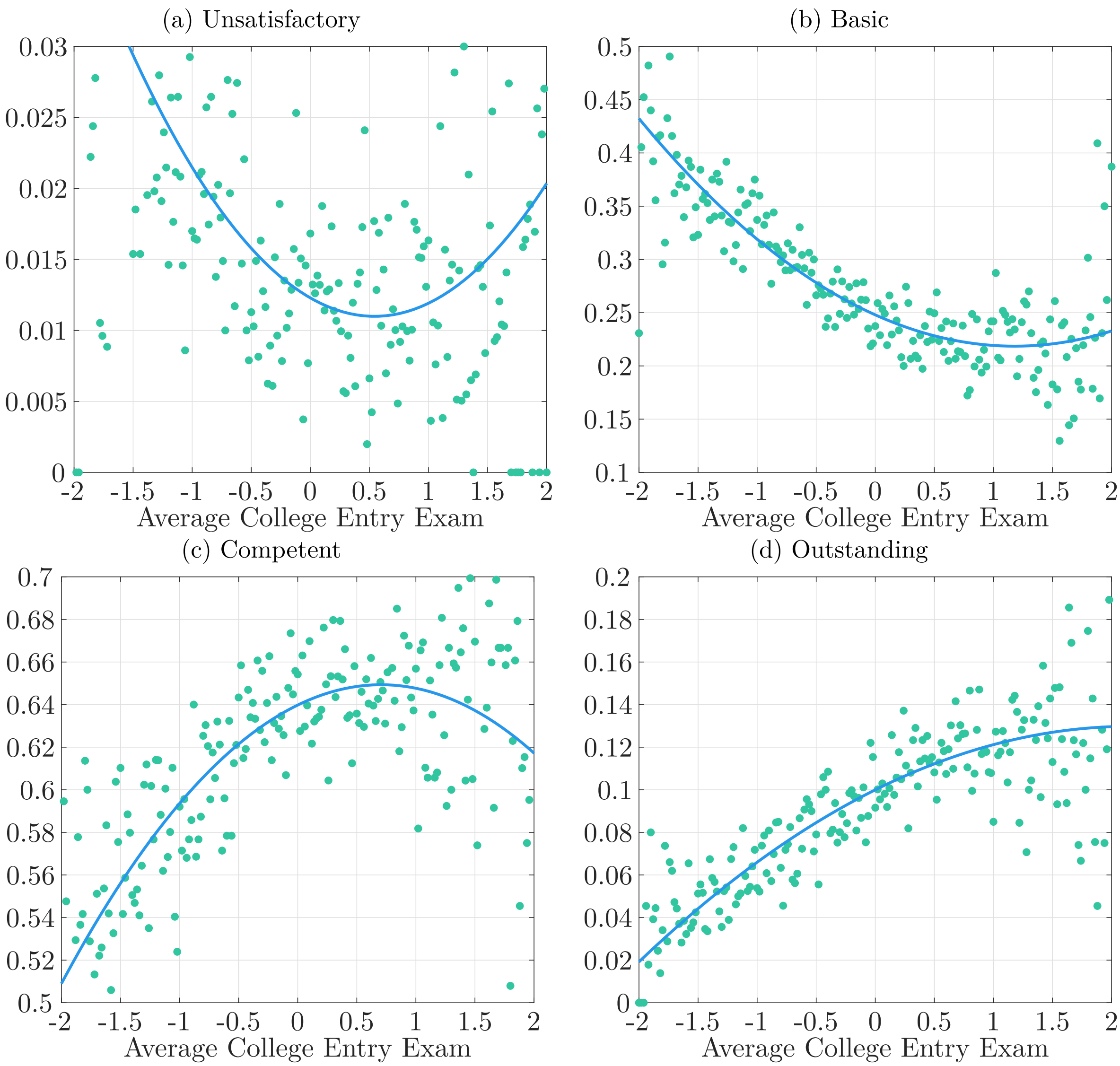
	With PSU Scores		Without PSU Scores		T-Test
Classification	N obs	%	N obs	%	Difference
Outstanding	5931	9.3	944	6.3	3 ***
Competent	39605	62.3	8525	56.9	5.4 ***
Basic	16987	26.7	5104	34.1	-7.4 ***
Unsatisfactory	1016	1.6	401	2.7	-1.1 ***
Total	63539	100	14974	100	
Score	N obs	%	N obs	%	Difference
Overall	2.63	0.28	2.57	0.29	0.06 ***
Portfolio	2.29	0.32	2.2	0.32	0.09 ***

As for the exit exams, our large set of observations allow us to accurately graph the bivariate relation between the respective ED outcomes and the PSU Scores. Figure 3 shows that there is negative relation between being evaluated as unsatisfactory (Figure 3(a)) or basic (Figure 3(b)) and PSU Scores, while there is a positive relation with being classified as competent or outstanding (Figure 3(c) and Figure 3(d), respectively). The positive relation is also clear when examining ED scores both for the overall and portfolio case as shown in Figure 4(a) and Figure 4(b).

In Table 9 we test whether this correlation persists once we control for year fixed effects, and the specialization level of teachers. The regression table reports the coefficients of separate regressions of each teacher evaluation outcome on the PSU score, expressed in terms of standard deviations. The columns add year fixed effects and teacher specialization level fixed effects. Each coefficient should be read as the change in the dependent variable given one standard deviation (SD) of increase in the PSU scores.

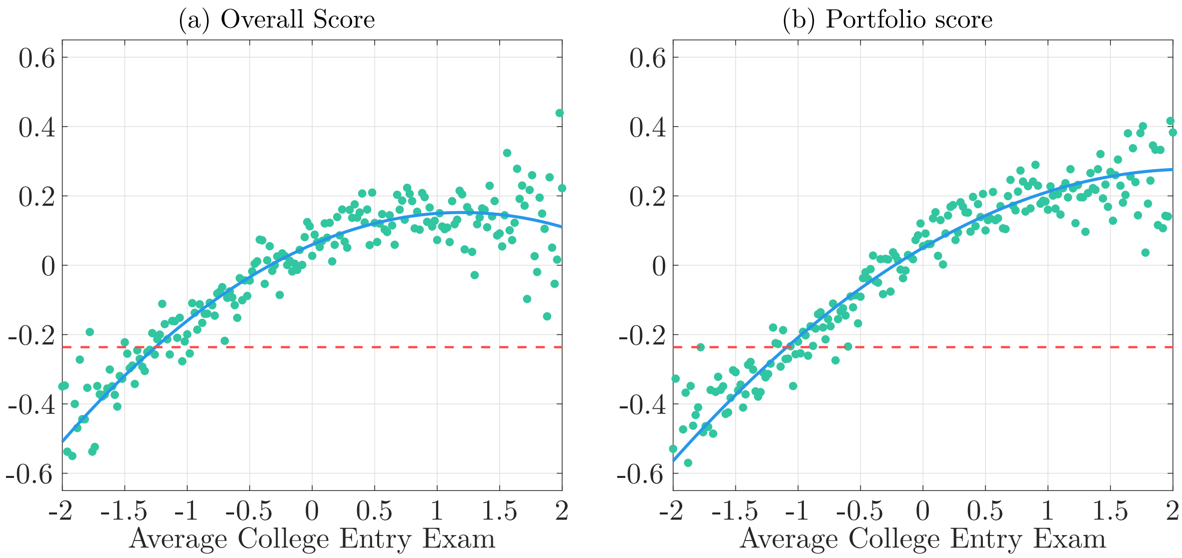
The results show that the coefficients are all significant (at 1% level) and stable across specification for the same outcome (i.e., independent of the controls added in each column). One SD increase in PSU scores is associated to an increase of approximately .7 SD standard deviations in the overall score. For the portfolio score, one SD of increase in the PSU is associated to a .6 standard deviation increase.

Figure 3: Teacher Evaluation Categories vs. PSU Scores



Note: The figures Figure 3(a), Figure 3(b), Figure 3(c) and Figure 3(d) plot the probability of being classified by the Government as unsatisfactory, basic, competent and outstanding respectively. The plots are built with 100 equal-sized bins of the average college entrance exam score and fits estimated lines using all the underlying data. The data consists in students enrolled in years 2004 to 2009 who graduated between 2009 and 2017.

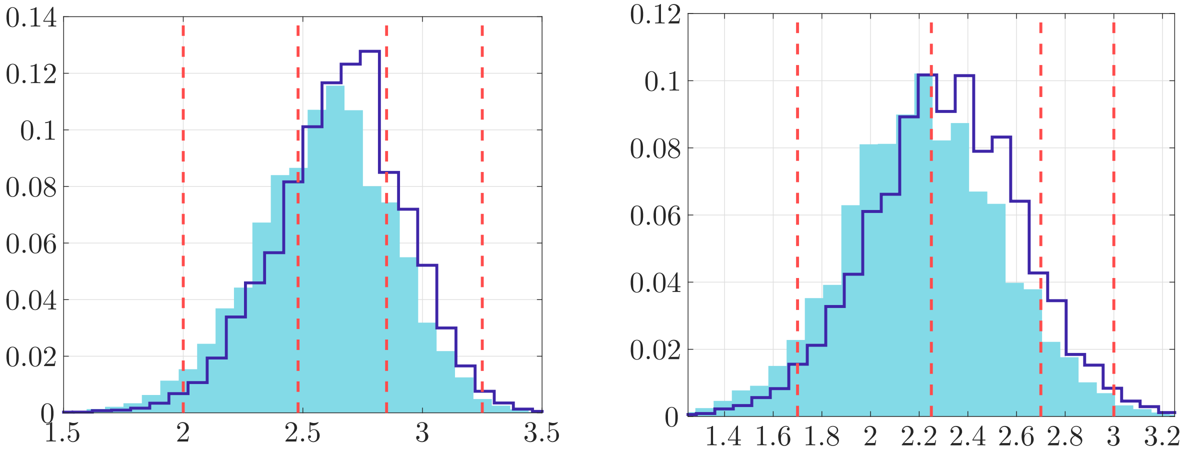
Figure 4: Teacher Evaluation Scores vs. PSU Scores



Note: Both figures use information for teachers working in the public sector that were evaluated at least once between years 2004 and 2016. The dots plot the average values of each variable within five points of the PSU score. The PSU score has a mean of 500 points and a standard deviation (SD) of 100 points, so each Figure plots data up to two SD to the left, and two SD to the right. The vertical axis the overall and portfolio scores of the teacher evaluation as a function of their PSU scores.

Figure 5: Teacher Evaluation Histograms

(a) Overall ED Score (b) Portfolio Score



Note: Both figures use information for teachers working in the public sector that were evaluated at least once between years 2004 and 2016. Figure 5(a) and Figure 5(b) plot the histograms for the overall evaluation score and portfolio evaluation score achieved by teachers respectively. The unshaded histogram with ■ color shows the distribution for teachers without PSU scores, meanwhile the shaded histogram with ■ color plots the distribution for teachers without PSU scores. The vertical dashed lines indicate the scores that separate teachers into four categories of performance (unsatisfactory, basic, competent and outstanding) as explained in the text. The cutoffs vary slightly over years, so they should be interpreted as proxies.

C Teacher Wages

C.1 Teacher Wages Description

The Ministry of Education collects administrative data on all teachers working in Chile each year. Each school of the country reports their number of working teachers disaggregated at the individual level with a set of characteristics of their job, like hours of contract, level taught and the subject they teach. In 2011 the Ministry of Education also asked schools to provide the wages payed to their teachers.

About 111K (88%) have a valid wage.[12]

[12] From the 125K teachers in the private sector, 113K were reported to have positive wages. We trimmed wages below the percentile 1 and above percentile 99, which left us with 111K teachers with valid wages.

Table 9: Regressions of Teacher Evaluation Performance on PSU Scores

	(1)	(2)	(3)
Overall Score, Mean: 0, S.D: 1			
PSU Score	0.801 ***	0.725 ***	0.615 ***
S.E.	(0.042)	(0.041)	(0.041)
PSU Score2	-0.065 ***	-0.059 ***	-0.048 ***
S.E.	(0.004)	(0.104)	(0.001)
N. Obs	[63539]	[63539]	[63539]
Portfolio Score, Mean: 0, S.D: 1			
PSU Score	0.697 ***	0.589 ***	0.477 ***
S.E.	(0.041)	(0.04)	(0.04)
PSU Score2	-0.049 ***	-0.04 ***	-0.031 ***
S.E.	(0.004)	(0.101)	(0.001)
N. Obs	[63539]	[63539]	[63539]
Writing Skills, Mean: 0, S.D: 1			
PSU Score	0.194 ***	0.255 ***	0.536 ***
S.E.	(0.057)	(0.055)	(0.046)
PSU Score2	-0.015 ***	-0.022 ***	-0.049 ***
S.E.	(0.006)	(0.134)	(0.002)
N. Obs	[36771]	[36771]	[36771]
ICT Skills, Mean: 0, S.D: 1			
PSU Score	0.458 ***	0.448 ***	0.628 ***
S.E.	(0.045)	(0.044)	(0.043)
PSU Score2	-0.037 ***	-0.036 ***	-0.055 ***
S.E.	(0.004)	(0.114)	(0.002)
N. Obs	[58523]	[58523]	[58523]
Year F.E.		X	X
Controls			X

Note: Robust standard errors in parentheses. Significance levels: * p0.10, ** p0.05, *** p0.01. The regression table reports the coefficients of 12 separate regressions for different measures of productivity over PSU score, both dependent and independent variables expressed in terms of standard deviations. The columns add year fixed effects, teacher specialization level (primary and secondary levels) fixed effects and a polinumium of order two of experience. Each coefficient should be read as the change in the outcome given one standard deviation (SD) of increase in the PSU scores.

C.2 Teacher Wages and PSU Scores

About 90% from the 111K teachers with valid wages have an available PSU score. The 10% of teachers with no PSU score are older teachers, aged 44 on average, 7 years more than educators with psu available.[13] Table 10 shows that teachers with no PSU score earn higher wages, probably due to seniority. Figure 6 plots the hourly wage histogram by PSU Score availability, which confirms that for most teachers in the upper part of the distribution we do not have PSU scores.

Table 10: Hourly Wage, by PSU Score Availability

With PSU Scores	Mean	Std. Dev.	N	%	p10	p50	p90
No	19.71	6.92	10954.00	10.00	12.75	17.76	29.24
Yes	19.06	6.24	100653.00	90.00	12.97	17.64	27.29
All	19.12	6.31	111607.00	100.00	12.94	17.65	27.54

As opposed to teachers working in the public sector (39%), teachers from the voucher sector (61%) are not subject to any mandatory evaluation. However, while teachers working in public schools benefit from a special labor code, voucher teachers work under the regular, more flexible, labor code. Therefore, their wages should be associated to productivity, reflecting how the labor market values their performance. We use the market wages as a proxy for the voucher teachers quality.

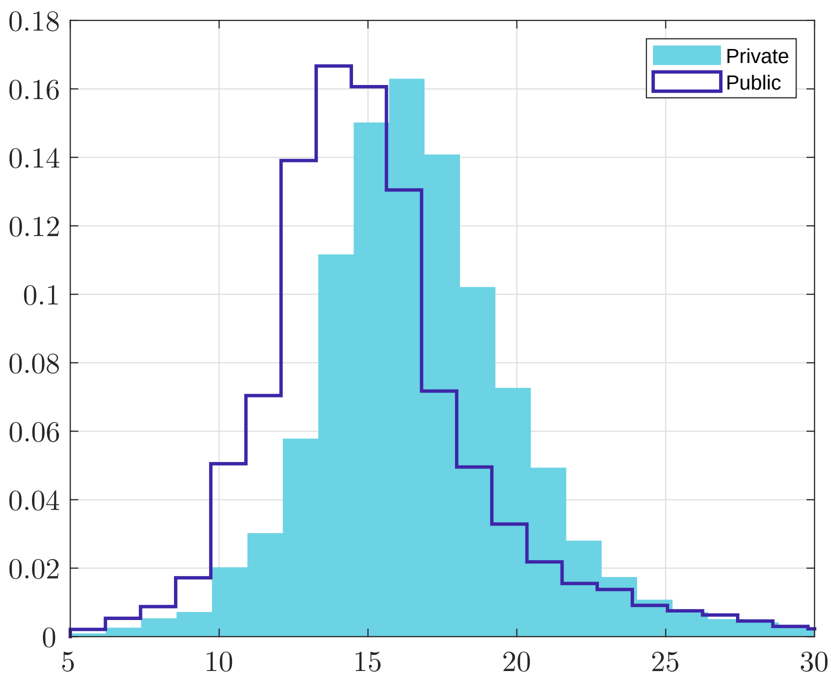
In Table 11 we report the decomposition of teachers in both public and private sector by working hours and, as can be seen, the most productive teachers working in the private sector are the ones that work full time. In contrast, in the public sector the most productive teachers are the ones that work less hours.

[13] As we explained in detail in the PSU Section, we collected data on the national college exam (PSU) that teachers took up to 35 years ago (from 1980 onwards). Therefore, we do not have information for the older teachers.

Table 11: Hourly Wage Summary Statistics

Working Hours	Mean	Std. Dev.	N	%	p10	p50	p90	Corr(PSU)
				Private Sector				
[0 – 30)	18.15	6.11	19752.00	33.00	12.85	16.93	24.48	0.04
[30 – 35)	17.92	4.36	10957.00	19.00	13.33	17.25	23.60	0.07
[35 – 40)	18.07	4.17	11733.00	20.00	13.83	17.38	23.35	0.04
[40 – 45]	18.84	4.69	16711.00	28.00	14.05	18.01	24.99	0.08
All	18.31	5.13	59220.00	100.00	13.47	17.39	24.29	0.06
				Public Sector				
[0 – 30)	19.56	9.00	7202.00	18.00	11.56	16.75	32.12	0.02
[30 – 35)	20.10	6.51	10231.00	26.00	12.61	18.97	29.35	-0.00
[35 – 40)	20.00	6.53	9846.00	25.00	12.61	18.76	29.01	-0.04
[40 – 45]	19.92	6.54	11265.00	29.00	12.55	18.57	28.81	0.03
All	20.19	7.40	39068.00	100.00	12.44	18.42	29.76	0.00

Figure 6: Wage distribution by sector



Note: The unshaded histogram with ■ color shows the distribution of wages for teachers in the public sector, meanwhile the shaded histogram with ■ color plots the distribution for wages of teachers in the private sector.

Moreover, as for previous measures of teacher quality, higher wages are associated to higher PSU scores. For our sample of 59K teachers in private schools with valid wages and PSU Scores, the Spearman correlation for those variables is positive, .06. as shown in Table 11. The correlation is higher for those teachers that work full time at the schools (.08 for the last row). However, this does not occur in the public sector, where the correlation is not clear.

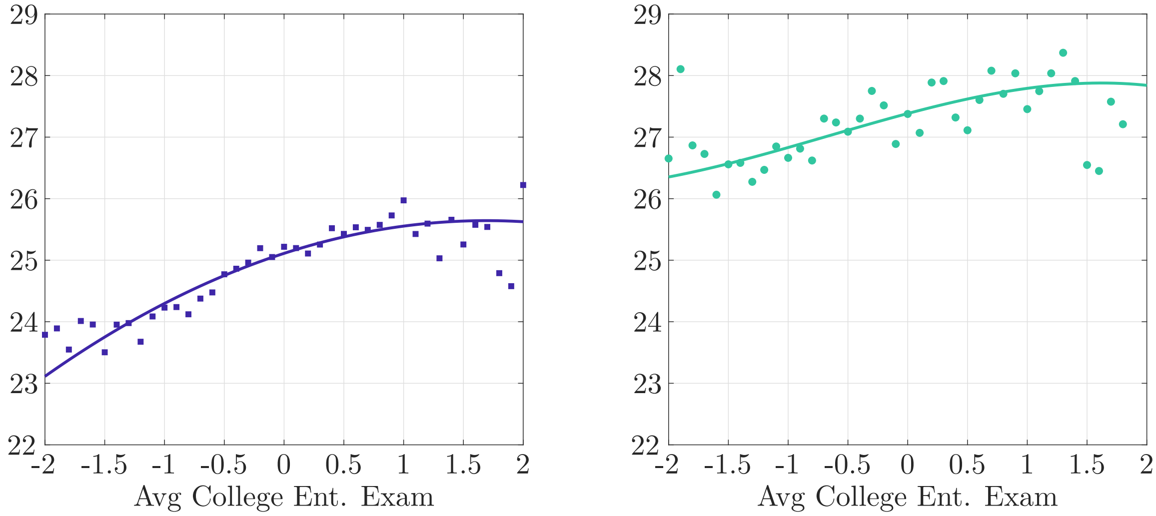
Figure 7 graphs nonparametrically the bivariate relation between wages and the PSU Scores for teachers in the private and public sector. Figure 7(a) shows a concave, non monotonic relation between psu and wages in the private sector, similar to the trend with the evaluacion docente scores. And, Figure 7(b) depict the same relation but for public sector professors. The positive relation exist, however is not as strong as the one reported for the teachers in the private sector.

In Table 9 we test whether this correlation persists once we control for the teachers' age (a quadratic function of age), the specialization level of teachers and the subjects they teach. The regression table reports the coefficients of separate regressions of wages for private and public sector teachers on the PSU score, expressed in terms of standard deviations. The columns add the controls for age, teacher specialization and subject taught. Each coefficient should be read as the change in the dependent variable given one standard deviation (SD) of increase in the PSU scores. We normalized the dependent variables in this case so we can refer to the coeficient as the percentage increase of a standard deviation in the dependent variable.

The results show that the coefficients are all significant (at 1% level). The magnitude decreases as we include controls, nevertheless remains sizable. One SD increase in PSU scores is associated to a 0.4 SD in wages in the private sector and the half the increase (0.24 SD) in the public sector.

Figure 7: Teacher Wages vs. PSU Scores

(a) Private Schools (b) Public Schools



Note: The two figures use information for teachers working in the voucher and public sector since 2011 until 2017, with valid wages and PSU scores. The graphs plot the wages for private sector teachers (Figure 7(a)), and public sector teachers (Figure 7(b)) as a function of their PSU Scores. The open circles the average values of each variable within five points of the PSU score. The PSU score has a mean of 500 points and a standard deviation (SD) of 100 points, so each Figure plots data up to two SD to the left, and two SD to the right.

D Employment: Working in Schools as a function of Exam Scores

D.1 Graduating Cohorts from Education Majors 1995-2013

We gathered records from students who graduated from all education majors in Chile for nine specific years between 1995 and 2013. We then combined them with administrative records on teachers who were working in schools in 2011.

Table 12 shows the number of graduates by year, with information on the fraction working in schools in 2015 and their PAA Scores. On average a 46% of 127K graduates from 1995-2013 were working in schools in 2011, and they have on average a score of 508 and the likelihood of working as teacher years after vary with the graduates' PAA scores, as Figure 8 shows.

Figure 8 shows the fraction of graduates of 2007 working in schools 2, 5, 10 years after as a function their PAA Scores. The inverted 'U' shape of the solid line in suggests that both low and high scored graduates have a lower likelihood of working as a teacher for 5 years or more after. Low PAA scores may not find jobs as teachers, while graduates with high PAA scores may enjoy better job alternatives than working in a school.

Table 12 also shows that the number of students graduating from education majors increased about fourfold (from 3K in 1995 to 13K in 2007). And the likelihood of working in schools in years after does depends on the years after graduation. It fluctuates between 0.29 for those graduated in 1995 and 0.51 for the 2001 graduates, and then goes to 0.41 for the 2013 graduates.

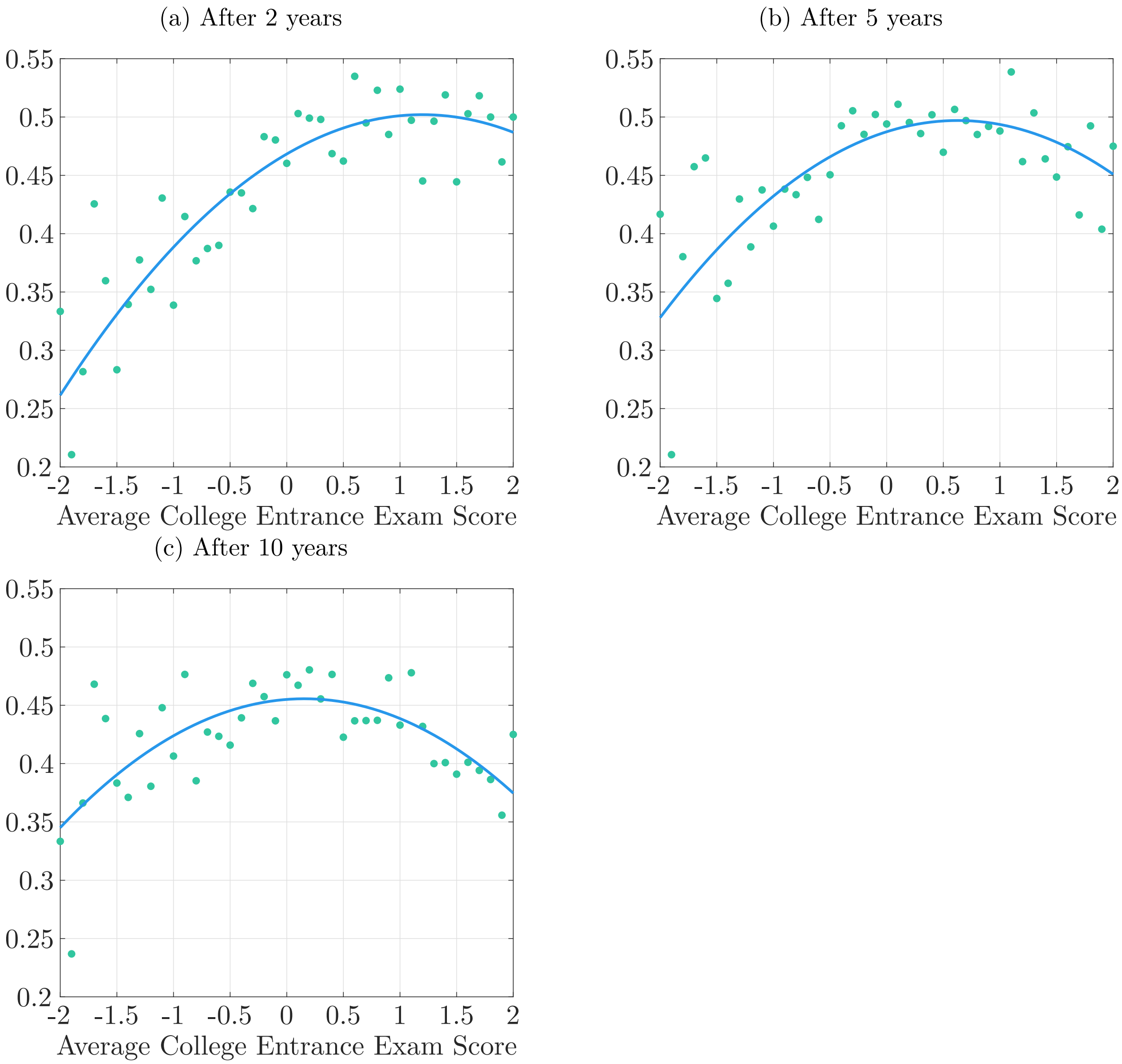
On the other hand, the graduates' PAA scores exhibit a steady tendency to decrease overtime on average. The PAA Scores for each cohort of test-takers have a mean of approx. 500 points and a standard deviation of about 100 points. There fore, the scores presented in the column can be interpreted as PAA Scores decreasing about .37 SD when comparing 1995 graduates with 2013 graduates.

Table 12: Graduates from Education working in Schools in 2011

Grad. Cohort	N	Working in 2015	PAA Score
1995	3062	0.29	539.76
1998	3707	0.35	534.66
2000	3639	0.35	529.19
2001	5051	0.33	519.53
2005	9482	0.39	506.9
2006	10068	0.42	511.65
2007	13403	0.44	507.95
2008	14528	0.45	501.07
2009	15545	0.46	495
2010	11719	0.53	504.53
2011	11836	0.54	502.65
2012	11620	0.54	514.21
2013	13390	0.51	512.52
Total	127050	0.46	508.66

Notes: This table presents information from education major graduates in years 1995, 1998, 2000, 2001 and 2005 to 2009 with valid PAA Scores (92% of the total of graduates). The PAA Scores for each cohort of test-takers has a mean of approx. 500 points and a standard deviation of about 100 points.

Figure 8: Works after graduation

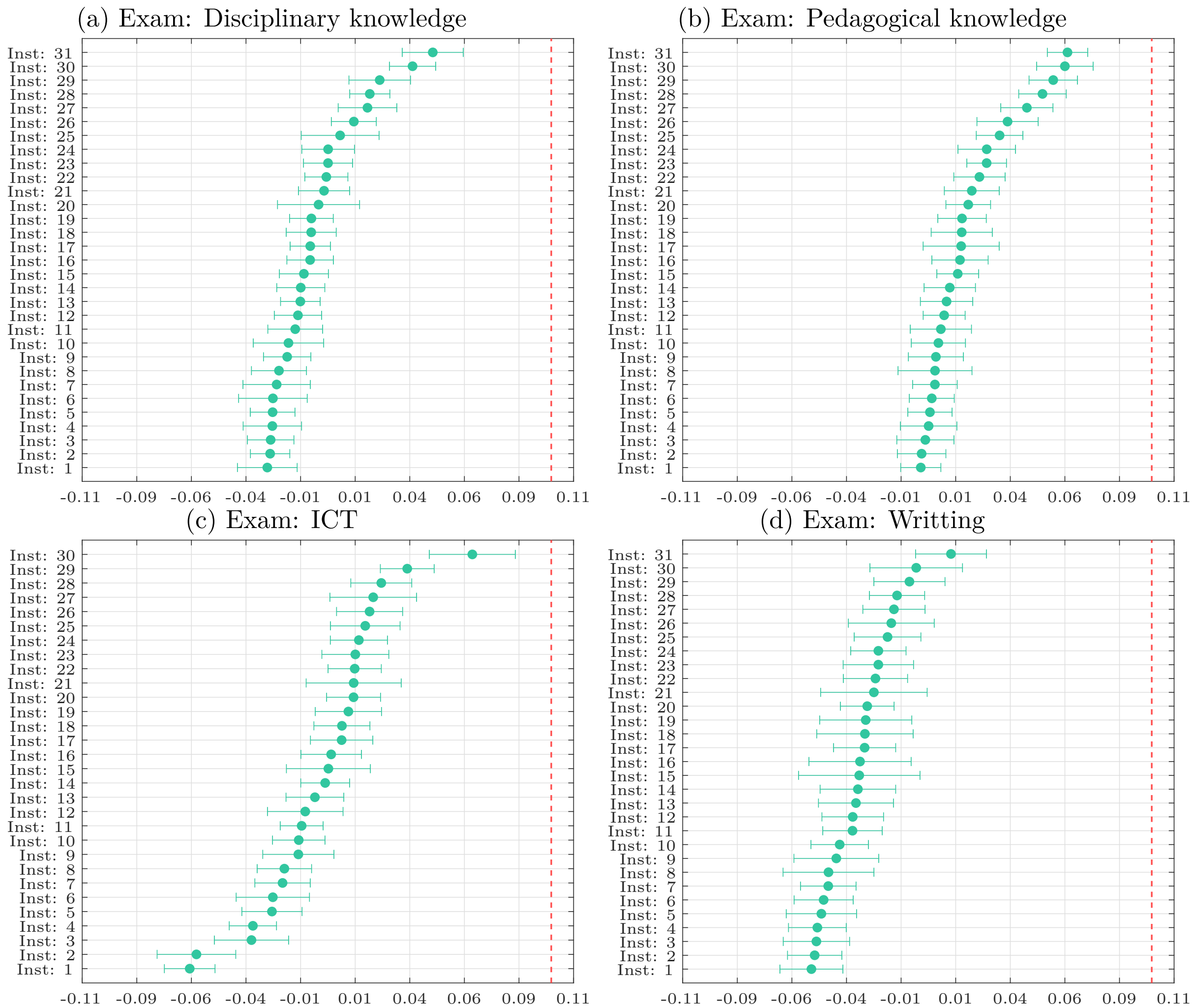


Figures 8(a), 8(b) and 8(c) plot the fraction of teachers employed in schools after 2, 5 and 10 years of graduation. The dots are averages of the outcome variable within 100 equal-sized bins of the average college entrance exam score. The data consists in graduates from teacher colleges in years 1995 to 2017, who are employed (or not) between 2003 to 2018. In the Figures the sample size is $N = 240,549$

E College Value Added

A straightforward way to assess the added value of higher education institutions is by examining more closely their estimated coefficients, which we plot in Figure 9. After controlling for PSU score, at the 5% significance level only three institutions add value to the Disciplinary exit exam score (Figure 9(a)), but the vast majority appears to add no value, in that test or the others. No institution's coefficient is statistically different from zero at the 1% significance level.

Figure 9: Institution Fixed Effects



Note: Each plot shows top 30 institution fixed effects (FE) with at least 100 observations in our sample, with 95% confidence intervals. Coefficients are sorted in descending order.

E.1 Data Analysis Procedures

To evaluate the value-added of schools over teaching performance, we propose to implement a regression discontinuity design over portfolio scores. The technique allows us to compare the portfolio score results for students slightly above the score threshold that allows them to enter a certain teaching school against students that were slightly below the same cutoff and therefore got accepted to the next preferred school. For this we will use the RD sample from Neilson et. al. 2020 and focus only on students that applied to at least one teaching option from the list that they submited to the CRUCH.

In table **??** we show descriptive statistics of this sample. First we list the number of applicants for which we have valid rules and cases in which the student's application were accepted, rejected or waitlisted. The number increased from 38479 in 1997 to 66790 in 2003. In the second column we show the total numbers but only for those listed at least one program in education in his preferences list, the number remains stable from 1977 to 2003 and the share is 32% on average. The share of Male students that applied to education programs decreased from 39.08% in 1977 to 31.99% in 2003. The share of applicants that came from private schools increased from 14% in 1978 to 20% in 1988 and then decreased to 8% in 2003. Regarding students' preferences, for all students that applied to higher education, around 5% listed education as their first option in 1977, this number increased across the years until it reached 13% in 2003. Meanwhile, 22% of students ranked education as they 4th or lower option, this share decreased to 2.4% in 2003. The average number of options listed by students that included education in their options were 1.8.

In table **??** we show the average score for different sections of the exam year by year for students applying to education. The scores for mathematics and verbal increase from 1977 to 1988 and then decrease for the years after 2000, meanwhile the scores for the optional exams from column 3 to 7 remain stable across all the years.

Table 13: Descriptive Preferences

	Total Apps	Education Apps	Share Edu	Share Male	Share Private	Edu First	Edu Second	Edu Third	Edu Other	Average Apps
1977	38479	13678	35.55	39.08	0.00	4.79	4.03	3.87	22.64	1.52
1978	43227	15898	36.78	39.14	13.71	4.45	4.32	4.43	23.58	1.58
1979	43164	16258	37.67	40.20	17.96	8.56	5.91	5.55	17.66	1.55
1980	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1981	40982	16856	41.13	40.39	18.01	11.35	6.67	5.88	17.22	1.80
1982	44112	17475	39.62	40.29	17.22	11.35	5.89	4.79	17.58	1.91
1983	49488	22978	46.43	43.84	17.52	12.64	7.03	5.70	21.06	2.01
1984	50237	25354	50.47	41.85	17.37	17.09	8.21	6.35	18.83	2.31
1985	48433	21747	44.90	40.95	17.66	16.00	7.76	5.76	15.38	2.33
1986	43812	13114	29.93	39.87	23.57	10.46	4.90	3.91	10.67	1.93
1987	40288	8836	21.93	35.74	18.68	8.43	3.85	2.77	6.88	1.67
1988	35851	6784	18.92	35.19	20.37	6.85	3.67	2.64	5.77	1.51
1989 - 1999	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2000	56294	9302	16.52	28.67	13.90	7.13	3.69	2.39	3.31	1.22
2001	63618	13260	20.84	30.98	10.92	10.44	4.86	2.49	3.06	1.51
2002	66081	14107	21.35	30.42	9.84	11.87	4.71	2.23	2.54	1.54
2003	66790	14780	22.13	31.99	8.26	13.01	4.68	2.08	2.36	1.52

Note: The shares on the first two columns (Male and Private) are based on the total students from Final Sample in table **??**. The third column indicates the share of students that listed Education in any of his preferences. And the last for columns indicate the share of students that put education as the first, second, third and other choice, the sum of the last four columns is equal to the fourth column.

Table 14: Descriptive Statistics

	PAA: Verbal	PAA: Math	PCE: History	PCE: Biology	PCE: CCSS	PCE: Physics	PCE: Math
1977	533.68	541.72	NaN	519.17	525.90	495.49	500.14
1978	545.73	551.12	NaN	522.21	534.79	498.55	506.61
1979	559.85	548.82	NaN	525.86	527.51	506.44	506.39
1981	550.88	572.01	NaN	541.36	553.76	NaN	517.73
1982	546.14	547.24	NaN	517.72	555.05	500.76	505.33
1983	557.03	575.08	NaN	528.44	553.55	504.81	506.84
1984	571.50	572.75	NaN	534.70	554.51	509.40	513.54
1985	568.15	581.65	552.10	531.22	558.08	503.70	520.47
1986	595.25	599.85	572.75	556.74	574.48	508.07	526.24
1987	582.48	587.11	563.67	527.43	559.17	497.53	501.66
1988	588.62	585.00	568.29	532.45	553.41	479.06	493.18
2000	576.47	532.67	556.50	503.11	536.75	468.29	488.96
2001	564.16	533.80	546.48	506.26	540.82	471.50	500.34
2002	565.71	546.76	559.70	508.01	544.29	469.83	504.52
2003	574.46	552.52	557.32	504.89	538.67	463.43	501.99

Note: .

E.1.1 Results

Table 15: Institution Causal Effects over Portfolio Score (Close RD, with Controls)

	Portfolio Exam			Works as Teacher			Works in High VA		
	Coefficient	SE	N Obs	Coefficient	SE	N Obs	Coefficient	SE	N Obs
Inst 1	-0.015	0.096	183	0.056	0.051	1577	0.055	0.04	1577
Inst 2	-0.021	0.035	1112	0.042**	0.019	9559	0.036**	0.016	9559
Inst 3	-0.082***	0.031	1214	0.039*	0.022	7151	0.013	0.017	7151
Inst 4	0.065	0.072	224	0.073*	0.041	2359	0.053*	0.032	2359
Inst 5	-0.09	0.091	227	-0.004	0.051	1839	-0.02	0.04	1839
Inst 6	-0.033	0.084	238	0.033	0.052	1386	0.063	0.042	1386
Inst 7	-0.032	0.143	64	0.088	0.089	405	-0.04	0.056	405
Inst 8	-0.099	0.114	109	0.068	0.071	800	0.117**	0.052	800
Inst 9	0.071	0.056	445	0.013	0.027	4863	-0.013	0.023	4863
Inst 10	-0.021	0.078	188	-0.049	0.052	1297	-0.023	0.04	1297
Inst 11	-0.144	0.206	61	0.064	0.088	438	-0.041	0.059	438
Inst 12	-0.006	0.1	100	-0.111	0.077	518	-0.022	0.057	518
Inst 13	0.036	0.049	558	0.03	0.031	3353	-0.015	0.025	3353
Inst 14	-0.001	0.045	560	0.005	0.028	4494	-0.011	0.023	4494
Inst 15	-0.001	0.107	102	0.022	0.09	445	0.045	0.063	445
Inst 16	-0.035	0.055	444	-0.078**	0.038	2281	-0.004	0.03	2281
Inst 17	-0.111	0.098	100	0.179**	0.085	478	0.093	0.079	478
Inst 18	0.074	0.072	281	-0.034	0.048	1597	0.029	0.04	1597
Inst 19	-0.015	0.054	424	-0.001	0.042	2051	0.028	0.032	2051
Inst 20	0.016	0.32	31	0.317*	0.183	125	-0.083	0.135	125
Inst 21	0.059	0.052	588	0.015	0.038	2856	0.018	0.029	2856
Inst 22	-0.196*	0.107	73	0.035	0.076	524	0.017	0.068	524
Inst 23	-0.06	0.095	85	0.145*	0.08	464	0.082	0.063	464
Inst 24	0	0.102	84	0.071	0.084	495	-0.05	0.068	495

Note: .

Table 16: Institution Causal Effects over Portfolio Score (Very Close RD, with Controls)

	Portfolio Exam			Works as Teacher			Works in High VA		
	Coefficient	SE	N Obs	Coefficient	SE	N Obs	Coefficient	SE	N Obs
Inst 1	0.03	0.1	183	0.046	0.049	1773	0.037	0.038	1773
Inst 2	-0.034	0.036	1112	0.048**	0.019	10020	0.035**	0.015	10020
Inst 3	-0.076**	0.032	1214	0.046**	0.022	7485	0.02	0.017	7485
Inst 4	0.064	0.075	224	0.07*	0.04	2493	0.048	0.031	2493
Inst 5	-0.112	0.094	227	0.017	0.049	2057	-0.027	0.037	2057
Inst 6	-0.023	0.087	238	0.018	0.053	1443	0.045	0.041	1443
Inst 7	-0.022	0.15	64	-0.001	0.084	459	-0.056	0.052	459
Inst 8	-0.117	0.105	109	0.092	0.066	891	0.127**	0.05	891
Inst 9	0.041	0.059	445	0.007	0.027	5052	-0.012	0.022	5052
Inst 10	-0.051	0.08	188	-0.064	0.051	1349	-0.023	0.039	1349
Inst 11	-0.188	0.187	61	0.015	0.084	466	-0.045	0.056	466
Inst 12	0.014	0.096	100	-0.132*	0.075	576	-0.025	0.053	576
Inst 13	0.037	0.049	558	0.033	0.032	3485	-0.014	0.024	3485
Inst 14	0	0.047	560	0.004	0.028	4659	-0.015	0.022	4659
Inst 15	-0.008	0.103	102	0.063	0.084	467	0.075	0.059	467
Inst 16	-0.053	0.056	444	-0.089**	0.038	2350	0.006	0.029	2350
Inst 17	-0.161*	0.096	100	0.187**	0.084	480	0.137*	0.077	480
Inst 18	0.075	0.076	281	0.018	0.048	1684	0.058	0.039	1684
Inst 19	0.002	0.053	424	0.002	0.041	2203	0.025	0.031	2203
Inst 20	0.23	0.221	31	0.142	0.175	131	-0.111	0.128	131
Inst 21	0.07	0.053	588	0.028	0.036	3101	0.026	0.027	3101
Inst 22	-0.18	0.121	73	0.07	0.075	530	0.05	0.067	530
Inst 23	-0.119	0.096	85	0.128	0.078	471	0.069	0.06	471
Inst 24	0.027	0.112	84	0.124	0.084	499	-0.038	0.067	499

Note: .

Table 17: Institution Causal Effects over Portfolio Score (Close RD, no Controls)

	Portfolio Exam			Works as Teacher			Works in High VA		
	Coefficient	SE	N Obs	Coefficient	SE	N Obs	Coefficient	SE	N Obs
Inst 1	0.056	0.123	108	-0.005	0.064	1000	0.075	0.051	1000
Inst 2	-0.006	0.043	709	0.056**	0.024	6093	0.041**	0.02	6093
Inst 3	-0.099**	0.039	804	0.033	0.027	4785	0.012	0.022	4785
Inst 4	0.108	0.1	138	0.087*	0.051	1485	0.08**	0.04	1485
Inst 5	-0.164	0.126	129	0.026	0.066	1078	0.014	0.051	1078
Inst 6	0.014	0.119	144	0.026	0.068	885	0.004	0.054	885
Inst 7	-0.141	0.202	44	0.156	0.112	259	0.044	0.073	259
Inst 8	-0.114	0.122	66	0.049	0.091	486	0.068	0.067	486
Inst 9	0.081	0.072	287	-0.034	0.034	3144	-0.053*	0.029	3144
Inst 10	-0.001	0.101	123	-0.027	0.065	826	-0.028	0.051	826
Inst 11	-0.262	0.33	42	0.049	0.113	277	-0.014	0.074	277
Inst 12	-0.113	0.132	62	-0.074	0.095	335	-0.03	0.073	335
Inst 13	0.041	0.059	375	-0.002	0.04	2216	-0.005	0.032	2216
Inst 14	-0.043	0.058	391	-0.015	0.035	3095	-0.018	0.029	3095
Inst 15	-0.042	0.143	73	0.129	0.114	310	0.065	0.081	310
Inst 16	-0.038	0.063	291	-0.074	0.047	1551	0.02	0.037	1551
Inst 17	-0.181	0.134	66	0.177	0.11	319	0.142	0.106	319
Inst 18	0.093	0.097	176	-0.072	0.061	970	0.002	0.051	970
Inst 19	-0.032	0.069	272	0.032	0.053	1312	0.047	0.041	1312
Inst 20	0.213	0.512	21	0.503**	0.246	77	0.049	0.176	77
Inst 21	0.094	0.067	354	0.051	0.048	1782	0.029	0.036	1782
Inst 22	-0.093	0.139	48	0.1	0.095	352	0.041	0.085	352
Inst 23	-0.091	0.13	67	0.061	0.098	336	0.05	0.076	336
Inst 24	-0.155	0.125	62	0.075	0.106	325	-0.051	0.087	325

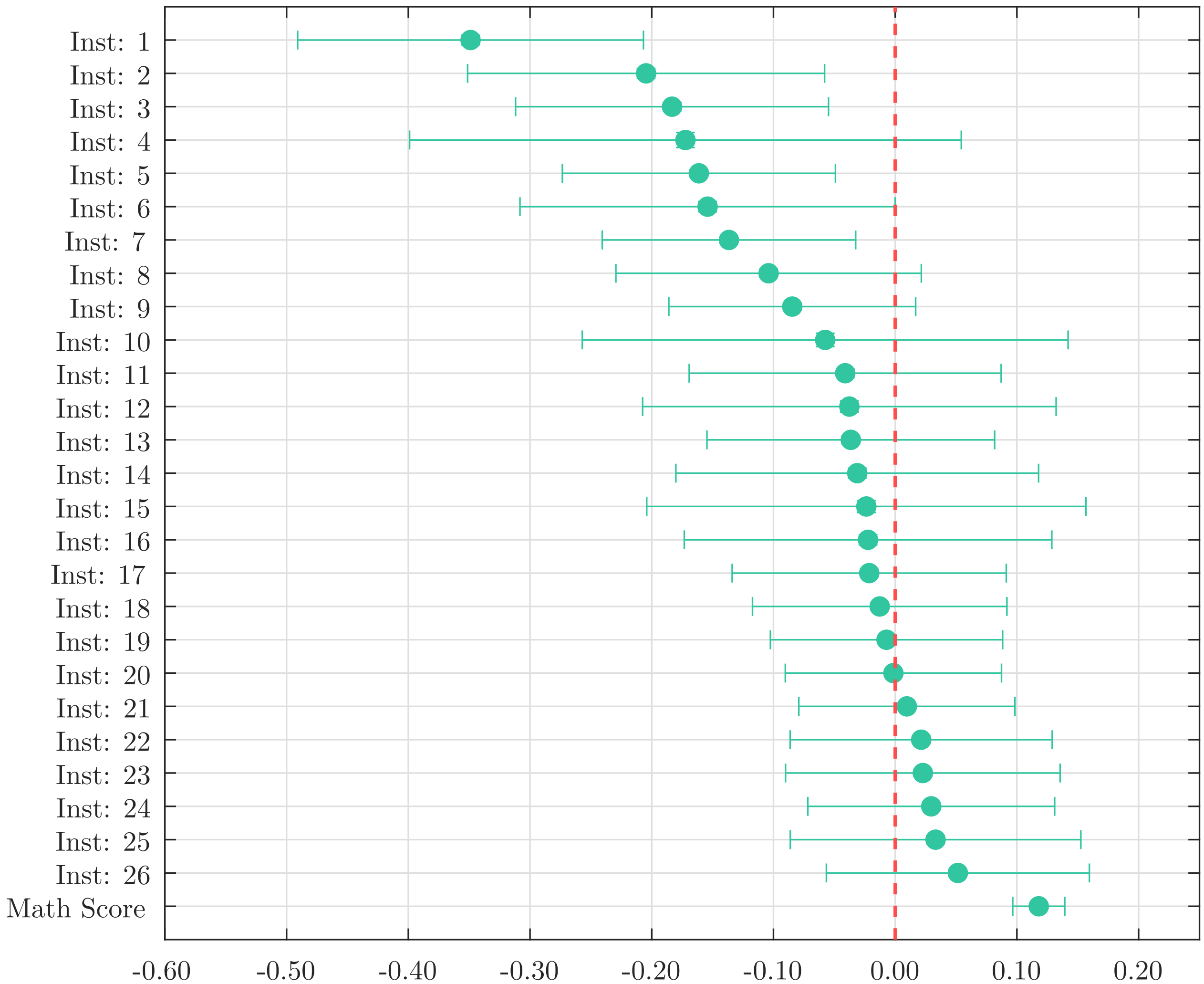
Note: .

Table 18: Institution Causal Effects over Portfolio Score (Very Close RD, no Controls)

	Portfolio Exam			Works as Teacher			Works in High VA		
	Coefficient	SE	N Obs	Coefficient	SE	N Obs	Coefficient	SE	N Obs
Inst 1	0.056	0.121	108	-0.011	0.061	1110	0.055	0.048	1110
Inst 2	-0.022	0.044	709	0.056**	0.024	6365	0.037*	0.019	6365
Inst 3	-0.096**	0.04	804	0.045*	0.027	5005	0.022	0.021	5005
Inst 4	0.099	0.103	138	0.09*	0.05	1564	0.069*	0.038	1564
Inst 5	-0.205*	0.123	129	0.052	0.063	1197	0.003	0.048	1197
Inst 6	0.051	0.119	144	0.002	0.068	910	-0.019	0.053	910
Inst 7	-0.079	0.196	44	0.021	0.106	298	-0.003	0.066	298
Inst 8	-0.116	0.114	66	0.049	0.084	534	0.065	0.063	534
Inst 9	0.04	0.076	287	-0.039	0.034	3278	-0.053*	0.028	3278
Inst 10	-0.091	0.1	123	-0.03	0.064	852	-0.015	0.049	852
Inst 11	-0.184	0.265	42	0.029	0.106	297	-0.015	0.072	297
Inst 12	-0.114	0.119	62	-0.132	0.093	372	-0.026	0.068	372
Inst 13	0.043	0.06	375	-0.003	0.04	2292	-0.01	0.031	2292
Inst 14	-0.062	0.06	391	-0.021	0.035	3204	-0.032	0.028	3204
Inst 15	-0.035	0.132	73	0.154	0.105	325	0.092	0.074	325
Inst 16	-0.038	0.063	291	-0.069	0.046	1604	0.034	0.036	1604
Inst 17	-0.253*	0.127	66	0.198*	0.105	321	0.201**	0.101	321
Inst 18	0.101	0.103	176	0.005	0.06	1015	0.039	0.049	1015
Inst 19	-0.011	0.07	272	0.005	0.052	1402	0.032	0.038	1402
Inst 20	0.132	0.244	21	0.287	0.233	83	-0.01	0.162	83
Inst 21	0.111	0.07	354	0.063	0.046	1927	0.032	0.034	1927
Inst 22	-0.151	0.152	48	0.106	0.094	357	0.067	0.084	357
Inst 23	-0.186	0.13	67	0.045	0.095	340	0.032	0.074	340
Inst 24	-0.099	0.132	62	0.109	0.108	328	-0.033	0.086	328

Note: .

Figure 10: Program Value Added over Portfolio Scores - RD coefficients



Note: The plot shows top 26 institution treshold crossing effects for students that got accepted into each education program versus the rest of programs of education. Standard errors are computed at the 95% confidence.

E.1.2 Robustness

Table 19: Testing Manipulation (15 pts bandwidth)

	I:1	I:2	I:3	I:4	I:5	I:6
α_1	-.0001	-.0003***	0	-.0009**	-.002***	-.0003
	(.0001)	(.0001)	(.0001)	(.0003)	(.0003)	(.0002)
α_0	.0016***	.0015***	.0016***	.0026***	.0029***	.002***
	(.0001)	(.0001)	(.0001)	(.0003)	(.0003)	(.0002)
Mean	.0018	.0014	.0015	.0022	.0025	.0019
N Obs.	1109	6345	4993	1559	1191	908
	I:7	I:8	I:9	I:10	I:11	I:12
α_1	.0001	0	0	.0011	.0002	0
	(.0008)	(.0004)	(.0002)	(.0007)	(.0003)	(.0005)
α_0	.0051***	.0029***	.0012***	.0024***	.0035***	.0039***
	(.0006)	(.0003)	(.0001)	(.0002)	(.0001)	(.0003)
Mean	.0049	.0029	.0012	.0028	.0037	.0037
N Obs.	295	532	3258	851	296	372
	I:13	I:14	I:15	I:16	I:17	I:18
α_1	-.0001	.0001	-.0011**	.0003*	.0003	-.0001
	(.0002)	(.0002)	(.0005)	(.0002)	(.0004)	(.0001)
α_0	.0018***	.0021***	.0048***	.0013***	.0035***	.0015***
	(.0001)	(.0001)	(.0004)	(.0001)	(.0002)	(.0001)
Mean	.0016	.0021	.0042	.0013	.0037	.0014
N Obs.	2289	3193	322	1603	321	1015
	I:19	I:20	I:21	I:22	I:23	I:24
α_1	0	.0027	-.0007***	-.0009*	.0006	-.0001
	(.0002)	(.0028)	(.0002)	(.0004)	(.0006)	(.0006)
α_0	.0021***	.0103***	.0023***	.0044***	.0038***	.0038***
	(.0001)	(.0019)	(.0002)	(.0004)	(.0004)	(.0005)
Mean	.002	.0147	.0021	.0035	.0042	.0037
N Obs.	1395	83	1921	357	340	328

Note: .

Table 20: Testing Manipulation (10 pts bandwidth)

	I:1	I:2	I:3	I:4	I:5	I:6
α_1	-.0001	-.0006***	.0002	-.0011*	-.003***	-.0009**
	(.0003)	(.0002)	(.0003)	(.0006)	(.0005)	(.0004)
α_0	.0024***	.0024***	.0023***	.0038***	.0045***	.0032***
	(.0002)	(.0001)	(.0002)	(.0005)	(.0005)	(.0004)
Mean	.0027	.002	.0022	.0032	.0038	.0028
N Obs.	720	4338	3490	1078	777	620
	I:7	I:8	I:9	I:10	I:11	I:12
α_1	.0007	-.0006	.0001	.0023	.0001	.0006
	(.0012)	(.0006)	(.0003)	(.0014)	(.0007)	(.0009)
α_0	.0066***	.0051***	.0017***	.0035***	.0051***	.0054***
	(.0008)	(.0005)	(.0002)	(.0004)	(.0003)	(.0006)
Mean	.0076	.0043	.0017	.0042	.0054	.0052
N Obs.	199	351	2228	564	204	270
	I:13	I:14	I:15	I:16	I:17	I:18
α_1	.0001	0	-.0029***	.0005*	.0004	0
	(.0003)	(.0003)	(.0009)	(.0003)	(.0006)	(.0002)
α_0	.0024***	.0032***	.007***	.0019***	.0051***	.0021***
	(.0002)	(.0002)	(.0006)	(.0002)	(.0004)	(.0002)
Mean	.0023	.003	.0059	.0019	.0052	.002
N Obs.	1594	2177	229	1101	224	696
	I:19	I:20	I:21	I:22	I:23	I:24
α_1	.0003	.005	-.0008**	-.001	.0008	-.0004
	(.0004)	(.0036)	(.0004)	(.0007)	(.001)	(.001)
α_0	.0028***	.0173***	.0032***	.0061***	.0051***	.0059***
	(.0003)	(.0011)	(.0003)	(.0005)	(.0007)	(.0008)
Mean	.0029	.021	.003	.0053	.0055	.0053
N Obs.	961	53	1309	249	257	227

Note: .

Table 21: Testing Manipulation (5 pts bandwidth)

	I:1	I:2	I:3	I:4	I:5	I:6
α_1	-.0003	-.001*	-.0001	-.0023	-.0074***	-.0031**
	(.0008)	(.0006)	(.0008)	(.0018)	(.0017)	(.0013)
α_0	.005***	.0047***	.0051***	.0085***	.0108***	.0074***
	(.0006)	(.0004)	(.0006)	(.0013)	(.0015)	(.0012)
Mean	.0052	.0039	.0042	.0062	.0081	.0055
N Obs.	353	2256	1789	535	368	302
	I:7	I:8	I:9	I:10	I:11	I:12
α_1	.0064**	-.0022*	.0004	.0056	.0019	.0002
	(.0028)	(.0013)	(.001)	(.004)	(.0018)	(.0017)
α_0	.0118***	.0093***	.0034***	.008***	.0104***	.0104***
	(.0016)	(.0011)	(.0004)	(.0007)	(.0008)	(.0008)
Mean	.014	.0083	.0033	.0083	.0101	.0088
N Obs.	98	193	1126	293	105	164
	I:13	I:14	I:15	I:16	I:17	I:18
α_1	0	-.0001	-.0057***	.0011**	.0031	0
	(.0008)	(.001)	(.0018)	(.0005)	(.002)	(.0006)
α_0	.0051***	.0069***	.0145***	.0035***	.0105***	.0039***
	(.0004)	(.0006)	(.0015)	(.0003)	(.0011)	(.0004)
Mean	.0046	.006	.0111	.0035	.0108	.0038
N Obs.	809	1103	121	616	105	381
	I:19	I:20	I:21	I:22	I:23	I:24
α_1	.0005	-.0086	-.0012	-.0001	.0018	-.0027
	(.001)	(.0067)	(.0011)	(.0017)	(.0025)	(.002)
α_0	.0056***	.0455***	.0062***	.0113***	.0081***	.0123***
	(.0007)	(0)	(.0007)	(.0012)	(.0017)	(.0015)
Mean	.0058	.0496	.006	.0101	.0099	.0116
N Obs.	478	22	652	128	149	110

Note: .

Table 22: Test Covariate Smoothness

	Female	Capital	Public	Private	TookPF
Bandwidth: 15 points					
	Female	Capital	Public	Private	TookPF
α_1	-.0036	-.0094	.0158*	-.0103	.0185**
	(.0096)	(.0108)	(.0082)	(.0078)	(.0079)
α_0	.7128***	.3906***	.8385***	.1468***	.132***
	(.0067)	(.0069)	(.0057)	(.0053)	(.0048)
Mean	.71	.4	.84	.15	.14
N Obs.	34876	34876	34876	34876	34876
Bandwidth: 10 points					
	Female	Capital	Public	Private	TookPF
α_1	-.0021	-.0078	.0189*	-.0132	.0122
	(.0115)	(.0128)	(.0098)	(.0093)	(.0097)
α_0	.708***	.3882***	.8382***	.1476***	.1328***
	(.0081)	(.0082)	(.0069)	(.0064)	(.0059)
Mean	.71	.39	.84	.15	.14
N Obs.	23916	23916	23916	23916	23916
Bandwidth: 5 points					
	Female	Capital	Public	Private	TookPF
α_1	-.0024	.0112	.0167	-.0103	-.0012
	(.0163)	(.0177)	(.0132)	(.0126)	(.0135)
α_0	.7064***	.3724***	.8399***	.1464***	.1442***
	(.0113)	(.0115)	(.0092)	(.0085)	(.008)
Mean	.71	.39	.84	.14	.14
N Obs.	12256	12256	12256	12256	12256

Note: .

F Government Policy: Recruiting

F.1 Description of the program

The *Beca de Vocación Profesor* (BVP) was designed to attract high-scoring applicants to enroll at teacher colleges. The scholarship is offered in two ways:

- Type 1: Offered to students that enter a pedagogical career as freshman students with PSU exam taken the previous year.
- Type 2: Offered to senior college students that are looking to follow a program of pedagogical formation.

F.2 Benefits of BVP

Scholarship type 1 - pedagogy

- A scholarship to cover enrollment tuition fees for students that scored over 600 on average between the mathematics and verbal exam in PSU.
- A scholarship to cover enrollment tuition fees plus 80 000 pesos (or 100 dollars) each month for students that scored over 700 on average between the mathematics and verbal exam in PSU.
- A scholarship to cover enrollment tuition fees plus 80 000 pesos (or 100 dollars) each month and a semester in a education program in a foreign country for students that scored over 600 on average between the mathematics and verbal exam in PSU.

Scholarship type 2 - degree

- Offer a scholarship for the tuition fee of the last year of the program and the pedagogical year for students that look for a pedagogical career and obtained more than 600 points in PSU.
  - For students with PSU scores higher than 600 points, a scholarship for tuition and enrollment.
  - For students with PSU scores higher than 700 points, a scholarship for tuition and enrollment and 80 000 pesos (or 100 dollars) monthly.
- Bachelors degrees other than pedagogy do not receive BVP financial aid.

F.3 BVP Requirements

- No socioeconomic requirements.
- Chilean citizenship.
- Applied to pedagogical careers through the *Council of Chancellors of Chilean Universities* (CRUCH) admission system or be admitted into a regular pedagogical career in accredited institutions or in the process of accreditation according to the *Comisión Nacional de Acreditación* (CNA).
- Enter an eligible major as freshmen in the academic year 2011, with PSU taken in December 2010, independent from the graduation year from high school. Students who were enrolled in teacher colleges previously are not eligible.
- Obtain at least 600 points in PSU (Weighted average of 50% mathematics and 50% language) or obtain the Academic Excellence Scholarship (BEA) and a PSU score over 580.
- Enrollment in an institution and accredited career for at least two years and with cutoff score higher than 500 points on PSU as declared by the university.

F.4 Institution Requirements

- The career has to be accredited for at least two years.
- The admission cutoff score declared by the career for the *Oferta Académica 2011* has to be of 500 points at least (50% Language and 50% Mathematics),
- Only a 15% of students can be accepted by special admission according to previously defined conditions by the Ministry of Education (including students with supernumerary vacant or Academic Excellence Scholarship).
- Only regular pedagogical careers are eligible for this program. Distance and other special programs are not eligible for thes BVP.

F.5 Procedures

There are four stages to consider: application, pre-selection of beneficiaries, selection of beneficiaries, and appealing process.

1. Application. During October, applicants must complete a form at the website www.becavocaciondeprofesor.cl or www.fuas.cl. They provide academic, personal and socioeconomic information.

2. Pre-selection. The information provided by the applicants is validated with administrative data from the Ministry of Education and all applicants that comply with the assignment requirements (scores higher than 600 points, etc) enter a list of pre-selected beneficiaries. This list is published in the website www.becavocaciondeprofesor.cl. In addition, the same website publishes which institutions/majors are eligible for the benefit.

3. Selection. Students can check whether they are in the pre-selected list. Once they know their college entrance exam scores, they decide whether to enroll at eligible institutions/majors. If they do, then they enter the list of selected beneficiaries. Institutions need to send by May 31 their list of enrolled students, which the MINEDUC uses to start the payment process to the respective institution/major.

4. Appealing. Since 2012 there is an appealing process for applicants who did not make it to the selection list. They must upload supporting documents to the website www.becasycreditos.cl, and follow the respective instructions.

F.6 Descriptive Statistics

A total of 250,758 students took the college entrance exam in December 2010, aiming to start classes when the academic year starts, in March 2011. All of these test-takers are potentially eligible for the BVP if they achieve scores above the policy cutoffs.

Test-takers complete a survey providing information on their gender, date of birth, household income bracket and parental schooling among other characteristics. We combine this data with the scores information at the individual level, which we merge with administrative records of higher education enrollment coming from the MINEDUC. The enrollment records have information for the population of students enrolled in higher education institutions in the country.

In Table 23 we show the descriptive statistics for all test-takers, in three panels with information on scores, demographics, and higher education enrollment. The scores have a mean of about 500 points each.

Test takers are on average 19 years old at the moment of the test, and about half of them are girls. Their parents have on average slightly more than 11 years of completed schooling, and about 40% lives in the capital city. All this figures are consistent with data coming from national surveys (CASEN 2016) and censuses. About 55%, 35% and 10% graduated from voucher, public and private high schools, which again are consistent with population figures on enrollment in the country (MINEDUC 2018).

The last panel shows the fraction of test takers who enroll in higher education. A 63% of them enroll at any institution, 44% enrolls at colleges and half of that enrolls at the CRUCH universities. An 8% enrolls at any teacher college and a 5% enrolls at teacher colleges that were BVP eligible.

Table 24 shows mean characteristics for the same variables in Table 23, for four groups near the BVP policy cutoffs. The first group consists on test-takers with scores in the 480 to 520 range, i.e., 20 points around the BVP policy threshold of 500 points. The next three groups are constructed similarly, for test-takers with 20 points around the 600, 700 and 720 BVP policy cutoffs.

The data shows that variables correlated with scores, like income or parental schooling, increase by range of the scores. The fraction of students enrolled in higher education, college and CRUCH universities also increases with each score range, while enrollment at teacher colleges falls. Enrollment at eligible teacher colleges is negligible near the 500 cutoff, consistent with the design of the BVP policy.

Table 23: Descriptive Statistics for all Test-Takers

Variable	(1) Observations	(2) Mean	(3) Std. Deviation	(4) Min	(5) Max
Scores					
College Exam Score	250,758	501.06	102.34	178	850
Math Score	250,758	501.07	111.27	150	850
Language Score	250,758	501.04	108.34	150	850
Takes History Test	250,758	0.62	0.49	0	1
History Score	154,790	500.41	109.55	150	850
Takes Science Test	250,758	0.56	0.50	0	1
Science Score	139,783	500.52	109.47	150	850
High School GPA Score	248,807	535.81	99.88	208	826
Demographics					
Female	250,758	0.52	0.50	0	1
Age	250,758	19.38	3.17	15	78
Income (1-12 bracket)	250,758	3.40	2.88	1	12
Private Health Insurance	250,758	0.21	0.40	0	1
Father Schooling (years)	215,105	11.45	3.77	0	17
Mother Schooling (years)	233,044	11.30	3.57	0	17
Capital City	248,462	0.40	0.49	0	1
Public High School	248,462	0.35	0.48	0	1
Private High School	248,462	0.10	0.30	0	1
Voucher High School	248,462	0.55	0.50	0	1
Enrollment					
Enroll Higher Education	250,758	0.63	0.48	0	1
Enroll College	250,758	0.44	0.50	0	1
Enroll CRUCH	250,758	0.21	0.41	0	1
Enroll Any Teacher College	250,758	0.08	0.28	0	1
Enroll Eligible Teacher College	250,758	0.03	0.18	0	1

Notes: Table 23 shows descriptive statistics for the 250,758 students took the college entrance exam in December 2010. The college entrance exam score is the math-language average score; the history and science tests are optional exams. The High School GPA Score has valid data for 99.2% of the test-takers (248,807 of 250,758). The Age corresponds to the age at the moment of the test. The variables of parental schooling have missing information due to both non-response and test-takers not knowing the answer. Capital City indicates whether the test-taker lives in the capital of the country at the moment of the test, while the variables Public, Private and Voucher High School indicate the type of high school from which the test-takers graduated. These last four variables have a response rate of 99.1%. The enrollment variables come from population records collected by the Ministry of Education. Enroll in Higher Education takes value one if the test-taker enrolled at any institute or university. Enroll College is equal to one if the test-taker enrolled at any college; enroll CRUCH does the same if the test taker enrolled at universities belonging to the *Consejo de Rectores*. Enroll at any teacher college (TC) takes value one if test taker enrolled in any education major in the country, and Enroll Eligible TC does the same for enrollment at eligible teacher colleges.

Table 24: Mean Characteristics Near the BVP Policy Cutoffs

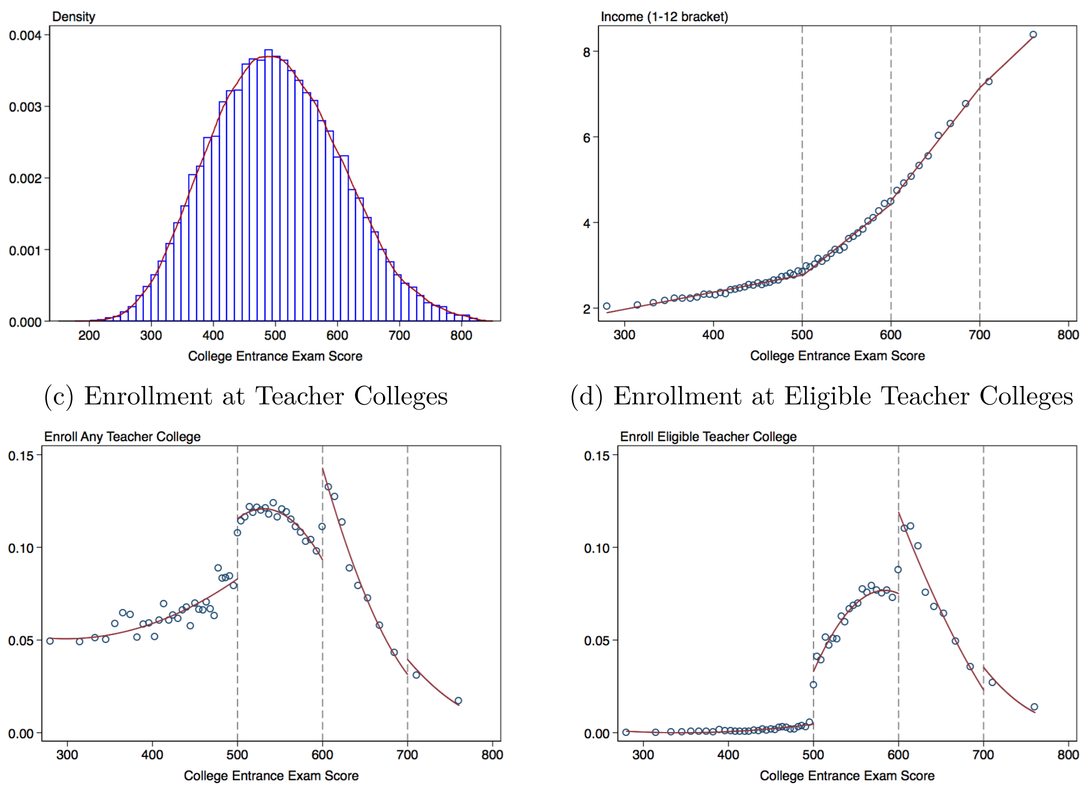
Variable	Score Range of the College Entrance Exam [480,520]	[580,620]	[680,720]	[700,740]
Scores				
College Exam Score	500	599	698	718
Math Score	499	599	700	728
Language Score	500	599	696	708
Takes History Test	0.64	0.52	0.47	0.44
History Score	510	603	676	682
Takes Science Test	0.54	0.70	0.77	0.80
Science Score	469	572	666	685
High School GPA Score	521	587	668	682
Demographics				
Female	0.54	0.50	0.43	0.40
Age	19.26	19.20	19.04	18.96
Income (1-12 bracket)	2.90	4.54	7.02	7.46
Private Health Insurance	0.15	0.34	0.59	0.63
Father Schooling (years)	11.15	13.02	14.75	15.03
Mother Schooling (years)	11.07	12.78	14.36	14.59
Capital City	0.37	0.44	0.57	0.58
Public High School	0.35	0.25	0.19	0.19
Private High School	0.04	0.18	0.46	0.51
Voucher High School	0.61	0.57	0.35	0.31
Enrollment				
Enroll Higher Education	0.66	0.81	0.90	0.92
Enroll College	0.46	0.78	0.90	0.92
Enroll CRUCH	0.14	0.46	0.69	0.73
Enroll Any Teacher College	0.10	0.11	0.04	0.03
Enroll Eligible Teacher College	0.02	0.09	0.03	0.02
Observations	37,589	23,932	6,210	4,602

Notes: Table 24 shows mean characteristics for the same variables in Table 23, for four groups near the BVP policy cutoffs. The first group consists on test-takers with scores in the 480 to 520 range, i.e., 20 points around the BVP policy threshold of 500 points. The next three groups are constructed similarly, for test-takers with 20 points around the 600, 700 and 720 BVP policy cutoffs. The number of observations correspond to those with valid scores in the respective score range.

F.7 Results

Our main results show that the policy attracted higher scoring test-takers to teacher colleges. Figure 11 summarizes the first results coming from the RD design. Figure 11(a) and Figure 11(b) are robustness tests, showing no manipulation of the running variable (the college entrance exam score) and that other covariates, such as household income behave smoothly near the policy thresholds.Figure 11(c) and Figure 11(d) illustrate effects on enrollment at teacher colleges (TC).

Figure 11: Main Results



(a) Score Distribution

(b) Household Income

(c) Enrollment at Teacher Colleges

(d) Enrollment at Eligible Teacher Colleges

Note: Figure 11(a) plots the distribution of scores for all test takers. Figure 11(b), Figure 11(c) and Figure 11(d) plot the mean of the y-axis variable within bins of scores, and fit estimated lines using all the underlying data.

Figure 11(c) and Figure 11(d) are suggestive of effects at the 500 and 600 points and a smaller increase at 700 points, for both all teacher colleges and eligible teacher colleges. In

?? we show the point estimates computed using an optimal bandwidth for each cutoff.

Effects at the 500 cutoff are of 3.3 percentage points (pp) over a mean of 8.6 pp , while the increase at the 600 cutoff is 3.7 pp over a mean of 9.5 pp. At 700 points effects are of 2.5 pp. smaller and negligible at 720 and with the opposite sign. demand incentives, scholarships for high scoring test-takers. supply incentives, participating teacher colleges cannot enroll low scoring students

Table 25: Enrollment at Any Teacher College

	(1)	(2)	(3)	(4)
$\widehat{\alpha_1}$	0.033***	0.037***	0.024*	-0.011
	(0.004)	(0.007)	(0.009)	(0.008)
Mean Below Cutoff	0.086	0.095	0.025	0.032
Effect Size	.364	.373	1.085	-
Optimal Bandwidth	48.8	35.6	26.6	34.9
Cutoff	500	600	700	720
Observations	87463	42418	8538	8210

	(1)	(2)	(3)	(4)
RD_Estimate	0.032***	0.035***	0.025**	-0.010
	(0.004)	(0.007)	(0.009)	(0.008)
Mean Below Cutoff	0.076	0.004	-0.116	-0.119
Effect Size	.398	7.355	-.243	.087
Optimal Bandwidth	48.3	34.3	26.3	34.5
Cutoff	500	600	700	720
Observations	86457	40559	8423	8210

	(1)	(2)	(3)	(4)
RD_Estimate	0.032***	0.035***	0.025**	-0.010
	(0.004)	(0.007)	(0.009)	(0.008)
Mean Just Below Cutoff	.086	.095	.025	.032
Optimal Bandwidth	48.3	34.3	26.3	34.5
Cutoff Value	500	600	700	720
Effective Observations	86,457	40,559	8,423	8,210
All Observations	250,758	250,758	250,758	250,758

Table 26: Enrollment at Eligible Teacher Colleges

	(1)	(2)	(3)	(4)
$\widehat{\alpha_1}$	0.033***	0.031***	0.022*	-0.010
	(0.002)	(0.006)	(0.008)	(0.007)
Mean Below Cutoff	0.005	0.073	0.022	0.027
Effect Size	6.896	.384	1.088	-
Optimal Bandwidth	41.9	31.3	29.1	34.3
Cutoff	500	600	700	720
Observations	75825	36955	9596	8034

	(1)	(2)	(3)	(4)
RD_Estimate	0.033***	0.029***	0.022**	-0.008
	(0.002)	(0.006)	(0.008)	(0.007)
Mean Below Cutoff	-0.034	-0.041	-0.118	-0.106
Effect Size	-.965	-.648	-.212	.073
Optimal Bandwidth	41.8	30.7	28.4	33.3
Cutoff	500	600	700	720
Observations	75825	36437	9178	7719

	(1)	(2)	(3)	(4)
RD_Estimate	0.033***	0.029***	0.022**	-0.008
	(0.002)	(0.006)	(0.008)	(0.007)
Mean Just Below Cutoff	.005	.073	.022	.027
Optimal Bandwidth	41.8	30.7	28.4	33.3
Cutoff Value	500	600	700	720
Effective Observations	75,825	36,437	9,178	7,719
All Observations	250,758	250,758	250,758	250,758

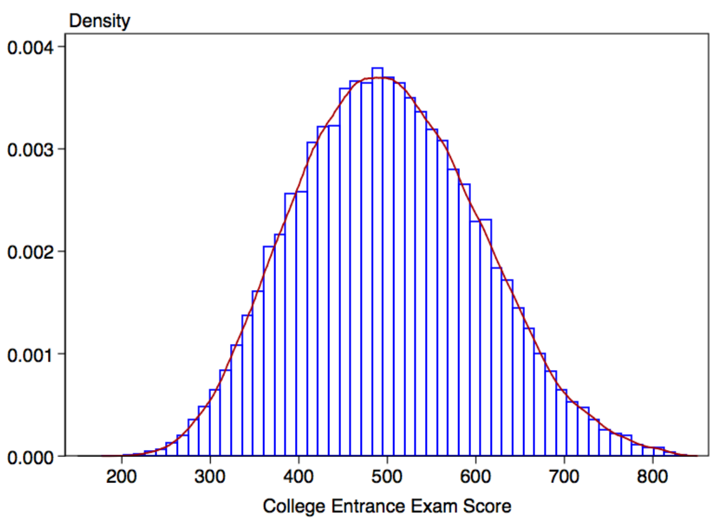
F.8 Robustness

F.8.1 Density of Running Variable

In this section we examine whether there is manipulation of the college entrance exams near the cutoffs in our data. In the centralized admission system in Chile scores are administered by a specialized agency (DEMRE), and test-takers do not know how to convert their performance in a score when they are taking the exam. Their raw score is a function of good and bad answers, and their final score is computed after standardizing raw scores taking into account all test-takers in the country.

In Figure 12 we plot the distribution of the college entrance exam score for all test takers. By construction, its a smooth bell-shaped distribution, showing no bunching at particular points of the support of the scores.

Figure 12: College Entrance Exam Density



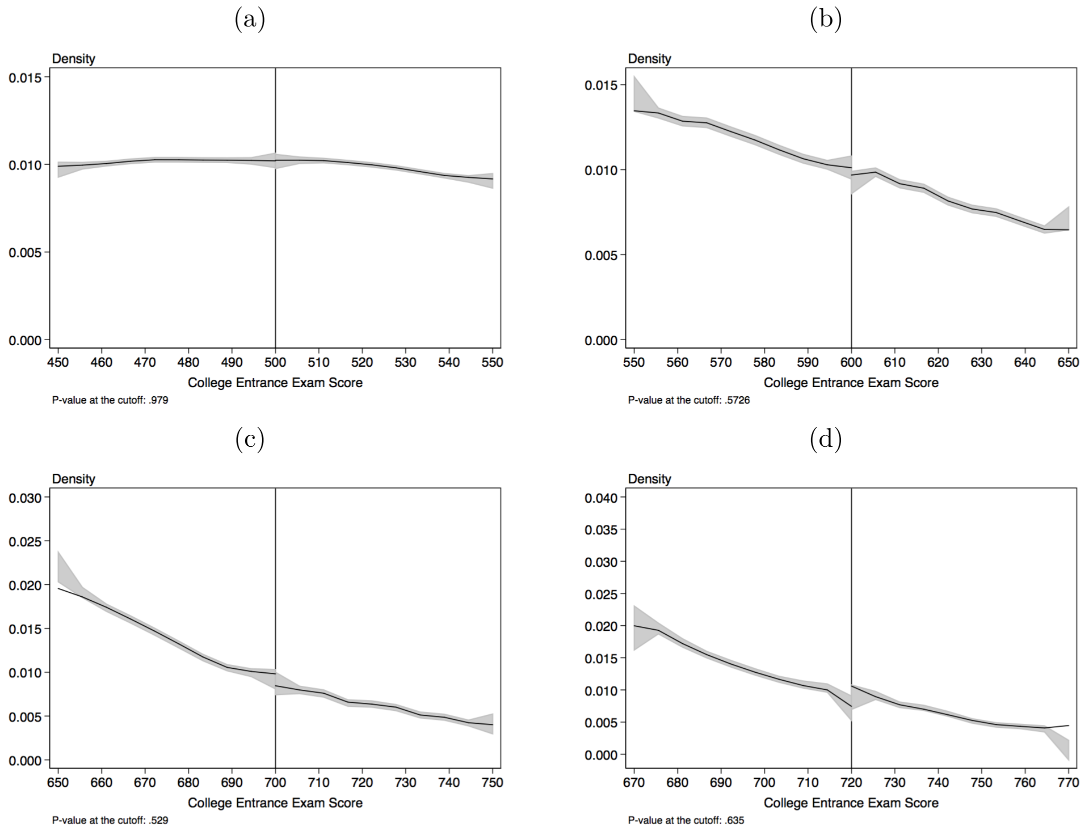
Note: The graphs in Figure 13 plot the mean of the y-axis variable within day of birth, and fit estimated lines using all the underlying data, allowing for different slopes on each side of the cutoff. Each day of birth contains about 2K observations. **??** shows the results from the estimation of equation (**??**) for these outcomes.

However, for the sake of the argument, let's suppose, for example, that some institutions could try to game the system by administratively manipulating their applicants' test scores when reporting to the Ministry of Education that those applicants are eligible for the scholarship. If that was the case, we would observe bunching of scores just above say, the 600 points threshold.

We test for manipulation using a nonparametric test **?** of discontinuity in the density of students with scores in the vicinity of the BVP cutoffs of 500, 600, 700 and 720 points. ble. **??** provides a graphical representation of the continuity in density test approach, plotting the density of observations by scores in our data.

At the bottom of each graph, we provide the p-value associated to the manipulation test. In all cases, a high p-value, which indicates that there is no statistical evidence of systematic manipulation of the running variable. This plot is consistent with the results from the formal test from **?**, as the density estimates above and below the the cutoff (the two intercepts in the figure) are very near each other.

Figure 13: Density Tests



Note: The graphs in Figure 13 plot the mean of the y-axis variable within day of birth, and fit estimated lines using all the underlying data, allowing for different slopes on each side of the cutoff. Each day of birth contains about 2K observations. **??** shows the results from the estimation of equation (**??**) for these outcomes.

In addition to the nonparametric test by **?** we also test parametrically whether the density changes at the cutoff in Table 27. Columns (1) to (3) in panales 1 to 4 show the coefficient $\widehat{\alpha_1}$ estimated from the equation (**??**) for test takers using the density of observations per score bins as dependent variable. The columns vary the points near the cutoff used to run our regressions. The results are again consistent with both the graphical representation of the data and the nonparametric test, indicating no statistical evidence of systematic manipulation.

Table 27: Testing Manipulation

	(1)	(2)	(3)	(4)	(5)	(6)
$\widehat{\alpha_1}$	-0.002	-0.002	-0.003	0.001	0.001	0.000
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Bandwith	20	10	5	20	10	5
Cutoff	500	500	500	600	600	600
Observations	37,432	18,383	8,699	23,814	11,914	5,909

	(1)	(2)	(3)	(4)	(5)	(6)
$\widehat{\alpha_1}$	-0.002	-0.003	-0.004*	-0.003	-0.007	-0.012
	(0.002)	(0.002)	(0.002)	(0.004)	(0.006)	(0.009)
Bandwith	20	10	5	20	10	5
Cutoff	700	700	700	720	720	720
Observations	6,156	3,076	1,523	4,598	2,320	1,154

Notes: Table 27 show the coefficient $\widehat{\alpha_1}$ estimated from the equation (**??**) for test takers using the density of observations per score bins as dependent variable. Robust standard errors (in parentheses) are clustered by score bins.

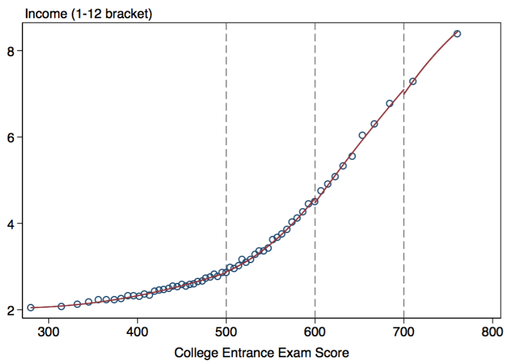
F.8.2 Covariates Smoothness

Our research design mimics a local experiment where test takers are exogenously allocated to receive a scholarship to study at teacher colleges. In this section we show that there are no other changes in our observable covariates occurring at the score threshold that could confound our analysis. **??** shows the results of estimating equation (**??**) using each covariate in **??** as dependent variable.

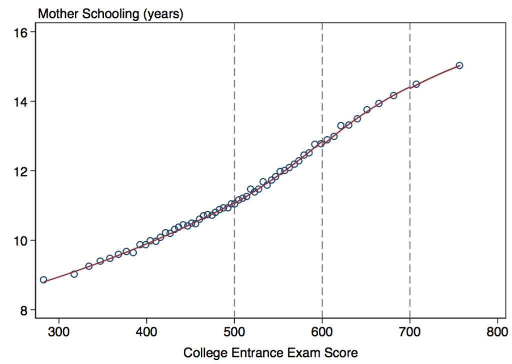
We complement these results with a graphical illustration for every covariate in Figure 14, which provide further evidence of a smooth behavior at the test score cutoff.

Figure 14: Covariates Smoothness

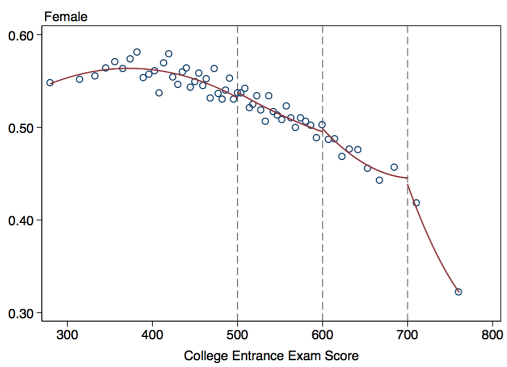
(a) Household Income



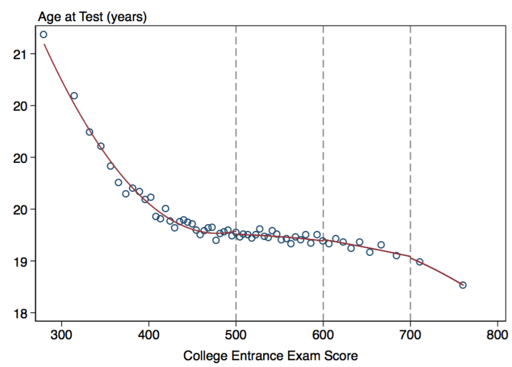
(b) Mother's Schooling



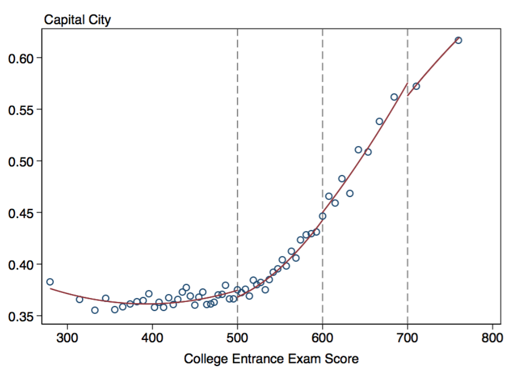
(c) Girl



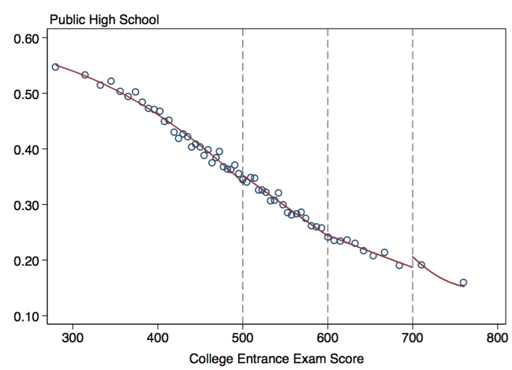
(d) Age



(e) Capital City



(f) Public School



Note: The graphs in **??** plot the mean of the y-axis variable within bins of scores, and fit estimated lines using all the underlying data, allowing for different slopes on each side of the cutoff. The y-axis variables are described in **??**.

Table 28: Covariates Smoothness

	(1) Girl	(2) Age	(3) Income	(4) Capital	(5) Capital Region	(6) Public School	(7) Private School	(8) Voucher School	(9) Father's Schooling	(10) Mother's Schooling
$\widehat{\alpha_1}$	-0.040	-0.002	0.016	0.008	-0.023	-0.007	-0.002	-0.001	0.003	-0.124
	(0.074)	(0.002)	(0.020)	(0.072)	(0.043)	(0.010)	(0.008)	(0.004)	(0.008)	(0.081)
Bandwidth	20	20	20	20	20	20	20	20	20	20
Cutoff	500	500	500	500	500	500	500	500	500	500
Observations	35,577	37,432	37,432	37,432	37,432	37,036	37,036	37,036	37,036	32,657
	(1) Girl	(2) Age	(3) Income	(4) Capital	(5) Capital Region	(6) Public School	(7) Private School	(8) Voucher School	(9) Father's Schooling	(10) Mother's Schooling
$\widehat{\alpha_1}$	0.001	0.008	-0.025	-0.072	0.019	0.005	-0.018	0.013	-0.134	-0.212*
	(0.001)	(0.015)	(0.061)	(0.113)	(0.012)	(0.010)	(0.012)	(0.014)	(0.093)	(0.107)
Bandwidth	20	20	20	20	20	20	20	20	20	20
Cutoff	600	600	600	600	600	600	600	600	600	600
Observations	23,814	23,814	23,814	23,814	23,660	23,660	23,660	23,660	20,909	22,401
	(1) Girl	(2) Age	(3) Income	(4) Capital	(5) Capital Region	(6) Public School	(7) Private School	(8) Voucher School	(9) Father's Schooling	(10) Mother's Schooling
$\widehat{\alpha_1}$	-0.002	-0.018	0.018	0.290	0.017	0.033*	-0.003	-0.030	0.054	-0.063
	(0.002)	(0.040)	(0.105)	(0.213)	(0.018)	(0.020)	(0.031)	(0.022)	(0.136)	(0.106)
Bandwidth	20	20	20	20	20	20	20	20	20	20
Cutoff	700	700	700	700	700	700	700	700	700	700
Observations	6,156	6,156	6,156	6,156	6,131	6,131	6,131	6,131	5,284	5,519
	(1) Girl	(2) Age	(3) Income	(4) Capital	(5) Capital Region	(6) Public School	(7) Private School	(8) Voucher School	(9) Father's Schooling	(10) Mother's Schooling
$\widehat{\alpha_1}$	-0.003	0.043	-0.159	0.318	-0.019	-0.000	0.037	-0.037	0.040	-0.000
	(0.004)	(0.050)	(0.116)	(0.220)	(0.026)	(0.024)	(0.028)	(0.027)	(0.185)	(0.143)
Bandwidth	20	20	20	20	20	20	20	20	20	20
Cutoff	720	720	720	720	720	720	720	720	720	720
Observations	4,598	4,598	4,598	4,598	4,574	4,574	4,574	4,574	3,931	4,109

Notes: Table 28 show the coefficient $\widehat{\alpha_1}$ estimated from the equation (**??**) for test takers using the density of observations per score bins as dependent variable. Robust standard errors (in parentheses) are clustered by score bins.

F.8.3 Bandwidth Selection

Table 29: Enrollment Estimates, 50 points near the cutoffs

	(1) Enroll Higher Ed	(2) Enroll College	(3) Enroll CRUCH	(4) Enroll Any Teacher College	(5) Enroll Elig. Teacher College
$\widehat{\alpha_1}$	0.022**	0.031***	0.035***	0.034***	0.041***
	(0.009)	(0.009)	(0.006)	(0.005)	(0.002)
$\widehat{\alpha_0}$	0.653***	0.465***	0.122***	0.085***	0.007***
	(0.007)	(0.008)	(0.003)	(0.003)	(0.001)
Effect Size	.033	.066	.287	.395	5.605
Bandwidth	50	50	50	50	50
Cutoff	500	500	500	500	500
Observations	89,520	89,520	89,520	89,520	89,520

	(1) Enroll Higher Ed	(2) Enroll College	(3) Enroll CRUCH	(4) Enroll Any Teacher College	(5) Enroll Elig. Teacher College
$\widehat{\alpha_1}$	0.023***	0.020***	-0.015	0.043***	0.043***
	(0.006)	(0.007)	(0.010)	(0.006)	(0.007)
$\widehat{\alpha_0}$	0.805***	0.779***	0.469***	0.095***	0.095***
	(0.004)	(0.005)	(0.007)	(0.003)	(0.004)
Effect Size	.029	.025	-.032	.45	.451
Bandwidth	50	50	50	50	50
Cutoff	600	600	600	600	600
Observations	58,989	58,989	58,989	58,989	58,989

	(1) Enroll Higher Ed	(2) Enroll College	(3) Enroll CRUCH	(4) Enroll Any Teacher College	(5) Enroll Elig. Teacher College
$\widehat{\alpha_1}$	0.011	0.011	-0.024	0.013*	0.018*
	(0.009)	(0.008)	(0.015)	(0.007)	(0.011)
$\widehat{\alpha_0}$	0.891***	0.889***	0.704***	0.030***	0.099***
	(0.007)	(0.007)	(0.010)	(0.004)	(0.006)
Effect Size	.013	.013	-.033	.443	.183
Bandwidth	50	50	50	50	50
Cutoff	700	700	700	700	700
Observations	17,523	17,523	17,523	17,523	17,523

	(1) Enroll Higher Ed	(2) Enroll College	(3) Enroll CRUCH	(4) Enroll Any Teacher College	(5) Enroll Elig. Teacher College
$\widehat{\alpha_1}$	0.006	0.006	0.007	-0.008	-0.020*
	(0.008)	(0.008)	(0.019)	(0.006)	(0.011)
$\widehat{\alpha_0}$	0.923***	0.921***	0.724***	0.028***	0.105***
	(0.005)	(0.005)	(0.013)	(0.005)	(0.007)
Effect Size	.007	.007	.009	-.287	-.191
Bandwidth	50	50	50	50	50
Cutoff	720	720	720	720	720
Observations	12,332	12,332	12,332	12,332	12,332

Notes: Table 30 show the coefficient $\widehat{\alpha_1}$ estimated from the equation (**??**) for test takers using the density

Table 30: Enrollment Estimates, 20 points near the cutoffs

	(1) Enroll Higher Ed	(2) Enroll College	(3) Enroll CRUCH	(4) Enroll Any Teacher College	(5) Enroll Elig. Teacher College
$\widehat{\alpha_1}$	0.037***	0.051***	0.040***	0.037***	0.040***
	(0.011)	(0.012)	(0.009)	(0.008)	(0.004)
$\widehat{\alpha_0}$	0.626***	0.432***	0.116***	0.080***	0.009***
	(0.007)	(0.007)	(0.004)	(0.004)	(0.001)
Effect Size	.06	.118	.345	.456	4.424
Bandwidth	20	20	20	20	20
Cutoff	500	500	500	500	500
Observations	37,432	37,432	37,432	37,432	37,432

	(1) Enroll Higher Ed	(2) Enroll College	(3) Enroll CRUCH	(4) Enroll Any Teacher College	(5) Enroll Elig. Teacher College
$\widehat{\alpha_1}$	0.021**	0.016	-0.037**	0.030***	0.027**
	(0.010)	(0.011)	(0.015)	(0.010)	(0.011)
$\widehat{\alpha_0}$	0.805***	0.779***	0.472***	0.098***	0.097***
	(0.006)	(0.008)	(0.011)	(0.005)	(0.007)
Effect Size	.026	.02	-.077	.311	.279
Bandwidth	20	20	20	20	20
Cutoff	600	600	600	600	600
Observations	23,814	23,814	23,814	23,814	23,814

	(1) Enroll Higher Ed	(2) Enroll College	(3) Enroll CRUCH	(4) Enroll Any Teacher College	(5) Enroll Elig. Teacher College
$\widehat{\alpha_1}$	-0.005	0.000	-0.025	0.024**	0.013
	(0.011)	(0.012)	(0.022)	(0.011)	(0.017)
$\widehat{\alpha_0}$	0.907***	0.901***	0.710***	0.024***	0.088***
	(0.009)	(0.009)	(0.013)	(0.005)	(0.009)
Effect Size	-.006	0	-.035	.985	.143
Bandwidth	20	20	20	20	20
Cutoff	700	700	700	700	700
Observations	6,156	6,156	6,156	6,156	6,156

	(1) Enroll Higher Ed	(2) Enroll College	(3) Enroll CRUCH	(4) Enroll Any Teacher College	(5) Enroll Elig. Teacher College
$\widehat{\alpha_1}$	0.008	0.007	0.005	-0.002	-0.027
	(0.011)	(0.011)	(0.034)	(0.010)	(0.018)
$\widehat{\alpha_0}$	0.928***	0.926***	0.723***	0.024***	0.117***
	(0.007)	(0.008)	(0.027)	(0.008)	(0.012)
Effect Size	.009	.008	.007	-.102	-.23
Bandwidth	20	20	20	20	20
Cutoff	720	720	720	720	720
Observations	4,598	4,598	4,598	4,598	4,598

Notes: Table 30 show the coefficient $\widehat{\alpha_1}$ estimated from the equation (**??**) for test takers using the density of observations per score bins as dependent variable. Robust standard errors (in parentheses) are clustered by score bins.

Table 31: Enrollment Estimates, 10 points near the cutoffs

	(1) Enroll Higher Ed	(2) Enroll College	(3) Enroll CRUCH	(4) Enroll Any Teacher College	(5) Enroll Elig. Teacher College
$\widehat{\alpha_1}$	0.034**	0.046**	0.045***	0.039***	0.039***
	(0.014)	(0.019)	(0.012)	(0.011)	(0.005)
$\widehat{\alpha_0}$	0.621***	0.435***	0.109***	0.082***	0.010***
	(0.007)	(0.010)	(0.004)	(0.006)	(0.002)
Effect Size	.055	.106	.41	.478	3.814
Bandwidth	10	10	10	10	10
Cutoff	500	500	500	500	500
Observations	18,383	18,383	18,383	18,383	18,383

	(1) Enroll Higher Ed	(2) Enroll College	(3) Enroll CRUCH	(4) Enroll Any Teacher College	(5) Enroll Elig. Teacher College
$\widehat{\alpha_1}$	0.007	0.009	-0.060***	0.030**	0.024
	(0.013)	(0.014)	(0.020)	(0.014)	(0.016)
$\widehat{\alpha_0}$	0.814***	0.780***	0.485***	0.095***	0.095***
	(0.008)	(0.010)	(0.013)	(0.007)	(0.010)
Effect Size	.009	.012	-.123	.313	.254
Bandwidth	10	10	10	10	10
Cutoff	600	600	600	600	600
Observations	11,914	11,914	11,914	11,914	11,914

	(1) Enroll Higher Ed	(2) Enroll College	(3) Enroll CRUCH	(4) Enroll Any Teacher College	(5) Enroll Elig. Teacher College
$\widehat{\alpha_1}$	-0.000	0.002	0.002	0.024	0.005
	(0.013)	(0.014)	(0.029)	(0.017)	(0.023)
$\widehat{\alpha_0}$	0.922***	0.920***	0.713***	0.027***	0.096***
	(0.011)	(0.012)	(0.019)	(0.008)	(0.012)
Effect Size	0	.003	.003	.9	.054
Bandwidth	10	10	10	10	10
Cutoff	700	700	700	700	700
Observations	3,076	3,076	3,076	3,076	3,076

	(1) Enroll Higher Ed	(2) Enroll College	(3) Enroll CRUCH	(4) Enroll Any Teacher College	(5) Enroll Elig. Teacher College
$\widehat{\alpha_1}$	0.005	0.004	0.018	-0.019	-0.032
	(0.016)	(0.016)	(0.057)	(0.013)	(0.026)
$\widehat{\alpha_0}$	0.938***	0.936***	0.718***	0.038***	0.123***
	(0.010)	(0.011)	(0.051)	(0.010)	(0.021)
Effect Size	.006	.005	.025	-.506	-.259
Bandwidth	10	10	10	10	10
Cutoff	720	720	720	720	720
Observations	2,320	2,320	2,320	2,320	2,320

Notes: Table 31 show the coefficient $\widehat{\alpha_1}$ estimated from the equation (**??**) for test takers using the density of observations per score bins as dependent variable. Robust standard errors (in parentheses) are clustered by score bins.

Table 32: Enrollment Estimates, 5 points near the cutoffs

	(1) Enroll Higher Ed	(2) Enroll College	(3) Enroll CRUCH	(4) Enroll Any Teacher College	(5) Enroll Elig. Teacher College
$\widehat{\alpha_1}$	0.026	0.036	0.064***	0.047***	0.041***
	(0.019)	(0.027)	(0.016)	(0.015)	(0.006)
$\widehat{\alpha_0}$	0.624***	0.446***	0.102***	0.080***	0.009***
	(0.011)	(0.012)	(0.004)	(0.007)	(0.002)
Effect Size	.042	.082	.63	.587	4.61
Bandwidth	5	5	5	5	5
Cutoff	500	500	500	500	500
Observations	8,699	8,699	8,699	8,699	8,699

	(1) Enroll Higher Ed	(2) Enroll College	(3) Enroll CRUCH	(4) Enroll Any Teacher College	(5) Enroll Elig. Teacher College
$\widehat{\alpha_1}$	0.023	0.018	-0.058**	0.015	0.008
	(0.016)	(0.017)	(0.023)	(0.017)	(0.021)
$\widehat{\alpha_0}$	0.815***	0.788***	0.497***	0.094***	0.101***
	(0.008)	(0.012)	(0.014)	(0.008)	(0.014)
Effect Size	.028	.023	-.117	.164	.08
Bandwidth	5	5	5	5	5
Cutoff	600	600	600	600	600
Observations	5,909	5,909	5,909	5,909	5,909

	(1) Enroll Higher Ed	(2) Enroll College	(3) Enroll CRUCH	(4) Enroll Any Teacher College	(5) Enroll Elig. Teacher College
$\widehat{\alpha_1}$	0.014	0.019	0.091***	0.022	-0.000
	(0.016)	(0.018)	(0.018)	(0.032)	(0.040)
$\widehat{\alpha_0}$	0.927***	0.922***	0.663***	0.029**	0.094***
	(0.011)	(0.013)	(0.009)	(0.013)	(0.019)
Effect Size	.015	.021	.137	.778	-.002
Bandwidth	5	5	5	5	5
Cutoff	700	700	700	700	700
Observations	1,523	1,523	1,523	1,523	1,523

	(1) Enroll Higher Ed	(2) Enroll College	(3) Enroll CRUCH	(4) Enroll Any Teacher College	(5) Enroll Elig. Teacher College
$\widehat{\alpha_1}$	-0.009	-0.006	0.028	0.010	0.030
	(0.030)	(0.031)	(0.098)	(0.022)	(0.037)
$\widehat{\alpha_0}$	0.924***	0.920***	0.665***	0.031	0.084**
	(0.027)	(0.028)	(0.091)	(0.021)	(0.035)
Effect Size	-.01	-.007	.043	.325	.358
Bandwidth	5	5	5	5	5
Cutoff	720	720	720	720	720
Observations	1,154	1,154	1,154	1,154	1,154

Notes: Table 32 show the coefficient $\widehat{\alpha_1}$ estimated from the equation (**??**) for test takers using the density of observations per score bins as dependent variable. Robust standard errors (in parentheses) are clustered by score bins.

G Government Policy: Screening

The Chilean Government enacted the Law 20,903, which creates a new system of professional development for teachers in the country. The Law includes guidelines for the recruitment, development and retention of teachers.

G.1 Admission to Teacher Colleges

The Law included new conditions for teacher colleges' admissions, making use of the PSU score distribution and High School GPA ranking, defined as follows:

The PSU score for a given year is equivalent to the average score in Mathematics and Language, considering all test takers on a particular year. The Agency in charge of the PSU test, DEMRE will deliver the necessary statistics regarding the PSU scores and a certificates documenting that the score for each applicant.

The High School GPA ranking is computed by the Ministry of Education, MINEDUC, taking into consideration the high school GPA of all students in the same cohort in the respective high school. Through the website Ayuda Mineduc (Mineduc Help), the Ministry will provide a document certifying the applicants' High School ranking.

The requirements for admissions in teacher colleges are designed to be gradually stricter over time:

- For the admission process in years 2017 to 2022, applicants to teacher colleges have to satisfy **at least one** of the following requirements:
  - Achieved a PSU score that is at least as high as the 50th percentile of the distribution (500 points in the average score between mathematics and language).
  - Achieved High School GPA in the top 30% of performance.
- For the admission process in years 2023 to 2025, applicants to teacher colleges have to satisfy **at least one** of the following requirements:
  - Achieved a PSU score that is at least as high as the 60th percentile of the distribution (525 points in the average score between mathematics and language).
  - Achieved High School GPA in the top 20% of performance.
  - Achieved a PSU score that is at least as high as the 50th percentile of the distribution (500 points) **and** a High School GPA in the top 40% of performance.

- For the admission process in year 2026 and onwards, applicants to teacher colleges have to satisfy **at least one** of the following requirements:
  - Achieved a PSU score that is at least as high as the 70th percentile of the distribution (550 points in the average score between mathematics and language).
  - Achieved High School GPA in the top 10% of performance.
  - Achieved a PSU score that is at least as high as the 50th percentile (500 points) of the distribution **and** a High School GPA in the top 30% of performance.

All of the conditions stated above are designed as **minimal** requirements for admission to teacher colleges. Each institution is allowed to consider stricter conditions, define number of vacancies or slots and application mechanisms. However, all the requirements must be informed before the beginning of admission process, each year.

H Machine Learning Rule

Table 33: Feature Contribution to Model Accuracy (AUC)

	Graduation	Works 7 years	Works in high VA
Model 1: Only PSU scores (Training sample: 65491)			
Logistic Regression	54.17%	64.53%	65.06%
Random Forest	57.33%	64.93%	65.25%
Bagging Regressors	58.65%	60.77%	60.33%
Adaboost Classifier	59.05%	65.02%	65.24%
Gradient Boosting	65.11%	65.14%	65.35%
Model 2: PSU scores and Transcripts (Training sample: 52716)			
Logistic Regression	53.99%	65.51%	65.96%
Random Forest	56.64%	65.69%	65.91%
Bagging Regressors	57.46%	63.68%	63.20%
Adaboost Classifier	58.66%	65.84%	65.90%
Gradient Boosting	64.41%	66.05%	66.17%
Model 1: PSU scores, Transcripts and SES (Training sample: 24778)			
Logistic Regression	60.91%	65.38%	66.23%
Random Forest	63.54%	65.19%	65.80%
Bagging Regressors	64.17%	63.80%	63.44%
Adaboost Classifier	62.57%	64.89%	65.31%
Gradient Boosting	64.66%	65.51%	66.13%

Note: The table shows the area under the curve estimated for different machine learning algorithms (Logistic Regression, Random Forest, Bagging Regressors, Adaboost Classifier, and Gradient Boosting) over three different versions (PSU, PSU + Transcripts, PSU + Transcritps + SES) and using different outcome variables (Graduation, Working in Schools after 7 years and Working in High Value added schools). Our estimates are results of a Grid Search over a high dimensional grid of hyperparameters whose combination was crossvalidated using 6 different sub samples from the training sample.

Table 34: Feature Contribution to Model Accuracy (AUC)

	Sample 3	Sample 2	Sample 1
High Value Added School			
PSU	64.47%	65.43%	65.69%
T+PSU	64.40%	66.11%	
NSE+T+PSU	65.45%		
Works after 7 years			
PSU	63.56%	65.78%	65.67%
T+PSU	64.37%	66.30%	
NSE+T+PSU	64.86%		
Graduates after 6 years			
PSU	60.73%	59.22%	59.51%
T+PSU	58.18%	57.59%	
NSE+T+PSU	63.94%		
Nobs Train	24778	52716	65491
Nobs Test	2754	5858	11558

Note: The table shows the area under the curve for the Gradien Boosting Machine estimated for three different models (PSU, PSU + Transcripts, PSU + Transcritps + SES) using different outcome variables (Graduation, Working in Schools after 7 years and Working in High Value added schools). Our estimates are results of a Grid Search over a high dimensional grid of hyperparameters whose combination was crossvalidated using 6 different samples.

Table 35: Machine learning contribution to screening performance

Graduation			Works after 6 years			High Value Added School		
P. 2017	P. 2020	P. 2023	P. 2017	P. 2020	P. 2023	P. 2017	P. 2020	P. 2023
Model: PSU Sample: 65491 Outcome: High VA								
10.91%	18.33%	27.3%	-5.17%	-7.78%	-9.11%	-8.03%	-11.15%	-14.22%
Model: PSU Sample: 65491 Outcome: High VA								
1.07%	2.76%	3.9%	3.53%	4.46%	7.37%	3.23%	4.31%	7.15%
Model: PSU Sample: 65491 Outcome: High VA								
1.08%	2.27%	2.15%	2.99%	4.22%	6.71%	3.12%	4.81%	7.05%
Model: PSU & Transcripts Sample: 52716 Outcome: High VA								
11.72%	18.7%	31.98%	-6.87%	-8.55%	-10.11%	-9.29%	-13.58%	-14.42%
Model: PSU & Transcripts Sample: 52716 Outcome: High VA								
1.46%	0.64%	2.58%	4.29%	6.16%	6.4%	2.93%	4.62%	5.7%
Model: PSU & Transcripts Sample: 52716 Outcome: High VA								
1.15%	0.09%	1.23%	3.65%	4.27%	6.0%	2.12%	3.96%	5.7%
Model: PSU & Transcripts & SES Sample: 24778 Outcome: High VA								
18.23%	27.93%	35.7%	-6.75%	-6.22%	-9.09%	-11.81%	-10.98%	-16.71%
Model: PSU & Transcripts & SES Sample: 24778 Outcome: High VA								
6.52%	5.34%	6.3%	4.92%	7.43%	8.92%	2.95%	5.49%	6.33%
Model: PSU & Transcripts & SES Sample: 24778 Outcome: High VA								
4.18%	7.19%	1.05%	3.43%	6.89%	8.06%	2.43%	5.69%	9.62%

Note: The table computes the contribution of the Machine Learning rule to the performance of the screening policy with different samples. All of the models use a Gradient Boosting Machine algorithm. We test nine combinations of the model where we use different three diferent versions of input features (PSU, PSU + Transcripts, PSU + Transcripts + SES) and three different outcomes (Timely graduation, Work in schools after 7 years, Working in High Value Added) to predict performance. Finally, we use the probabilities estimated in the test sample and compare to the performance of the government policy.

Table 36: Machine learning contribution to screening performance holding sample constant

Graduation			Works after 6 years			High Value Added School		
P. 2017	P. 2020	P. 2023	P. 2017	P. 2020	P. 2023	P. 2017	P. 2020	P. 2023
Model: PSU Sample: 24778 Outcome: High VA								
19.73%	26.49%	32.28%	-4.0%	-7.3%	-9.09%	-10.24%	-14.02%	-16.71%
Model: PSU Sample: 24778 Outcome: High VA								
4.79%	4.11%	3.59%	3.72%	5.14%	7.19%	1.68%	3.86%	5.98%
Model: PSU Sample: 24778 Outcome: High VA								
4.35%	7.19%	2.19%	3.2%	4.19%	7.37%	1.04%	3.86%	6.66%
Model: PSU & Transcripts Sample: 24778 Outcome: High VA								
14.88%	22.79%	28.61%	-6.41%	-8.24%	-9.95%	-12.5%	-14.84%	-17.97%
Model: PSU & Transcripts Sample: 24778 Outcome: High VA								
3.34%	4.93%	0.52%	4.35%	6.35%	7.03%	2.26%	4.67%	5.06%
Model: PSU & Transcripts Sample: 24778 Outcome: High VA								
5.02%	5.13%	-0.52%	4.58%	5.81%	5.66%	1.74%	5.69%	4.56%
Model: PSU & Transcripts & SES Sample: 24778 Outcome: High VA								
18.23%	27.93%	35.7%	-6.75%	-6.22%	-9.09%	-11.81%	-10.98%	-16.71%
Model: PSU & Transcripts & SES Sample: 24778 Outcome: High VA								
6.52%	5.34%	6.3%	4.92%	7.43%	8.92%	2.95%	5.49%	6.33%
Model: PSU & Transcripts & SES Sample: 24778 Outcome: High VA								
4.18%	7.19%	1.05%	3.43%	6.89%	8.06%	2.43%	5.69%	9.62%

Note: The table computes the contribution of the Machine Learning rule to the performance of the screening policy holding the sample constant at 24778 observations. All of the models use a Gradient Boosting Machine algorithm. We test nine combinations of the model where we use different three diferent versions of input features (PSU, PSU + Transcripts, PSU + Transcripts + SES) and three different outcomes (Timely graduation, Work in schools after 7 years, Working in High Value Added) to predict performance. Finally, we use the probabilities estimated in the test sample and compare to the performance of the government policy.