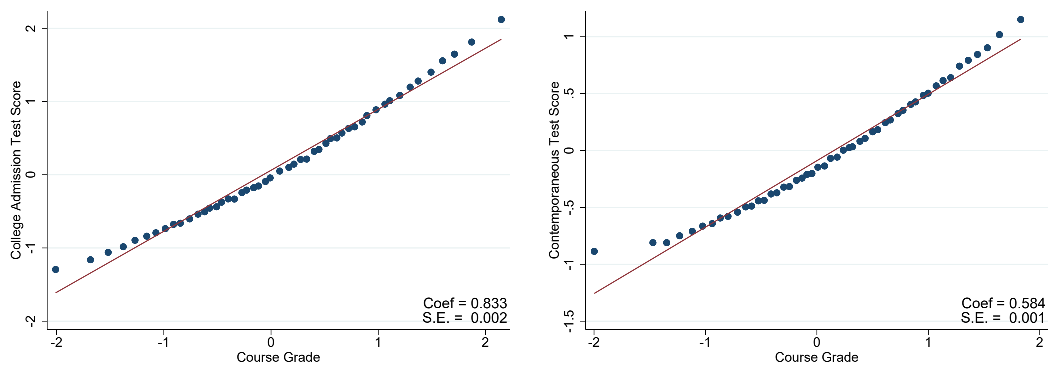
ONLINE APPENDIX

Managers' Productivity and Recruitment in the Public Sector
By Pablo Muñoz and Mounu Prem

A Additional Figures and Tables

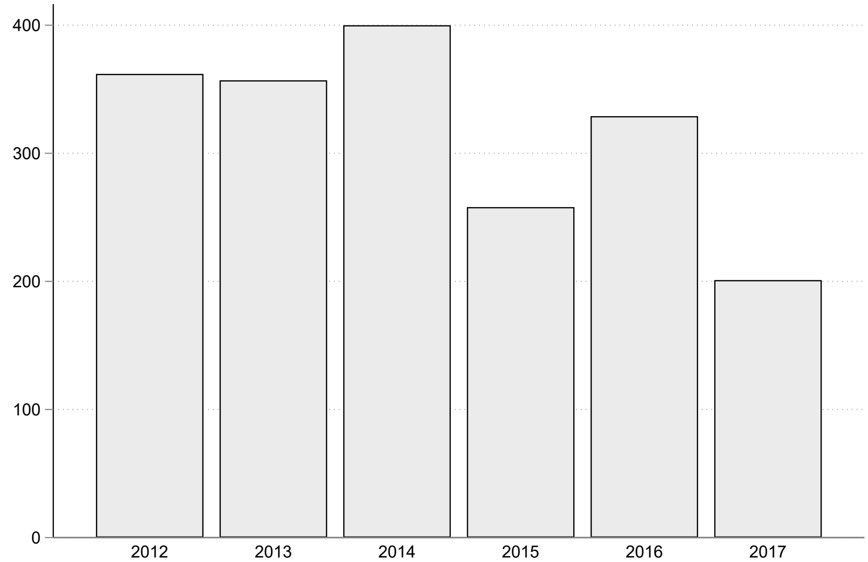
Figure A.1: Course Grades and Test Scores



A. PSU Test Score **B.** SIMCE Test Score

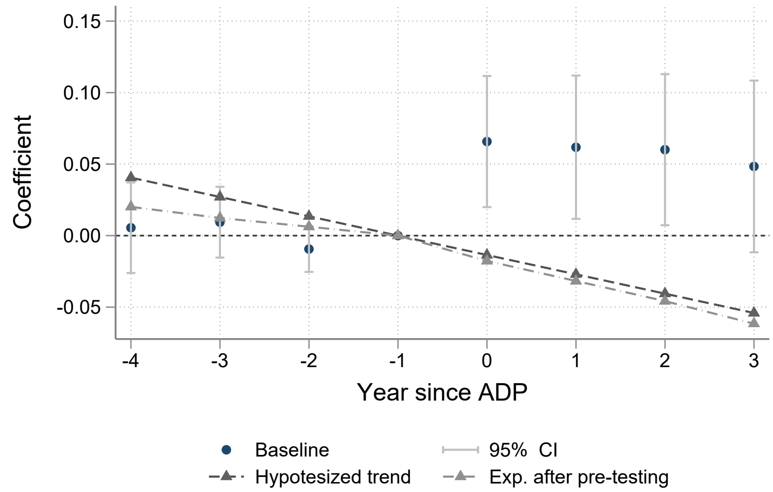
Notes: Panel A considers a sample of 132,585 students accepted into college and for whom we can compute college admission scores from the *Prueba de Selection Universitaria* (PSU), in the 2017 process. The college admission score is an institution-major-specific weighted average of applicants' high-school course grades and entrance exam scores. Panel B considers a sample of 1,061,231 students for whom we observe test scores from the *Sistema de Medición de la Calidad de la Educación* (SIMCE) and course grades contemporaneously for Math and Spanish between 2011 and 2016. We report the coefficient and robust standard error from a linear regression of test scores on course grades.

Figure A.2: Number of Newly Elected Principals by ADP, per Year

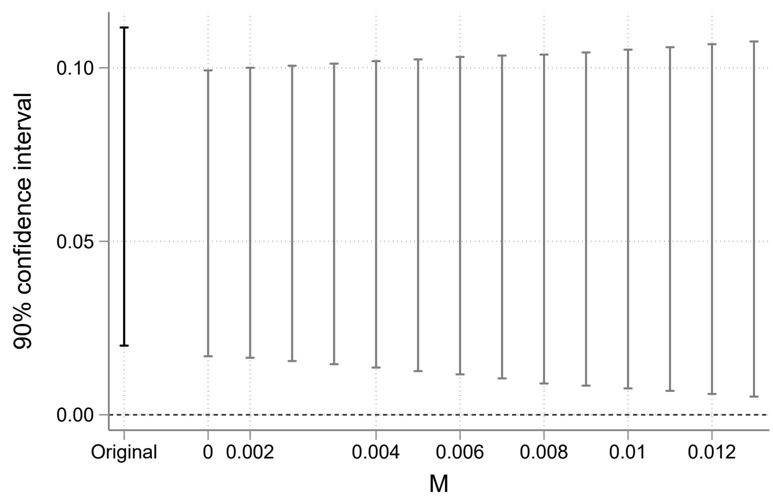


Notes: This figure shows the number of schools that elected a principal through the new ADP selection system for the first time, by year.

Figure A.3: Principal Selection and Principal Effectiveness: Parallel trends violations



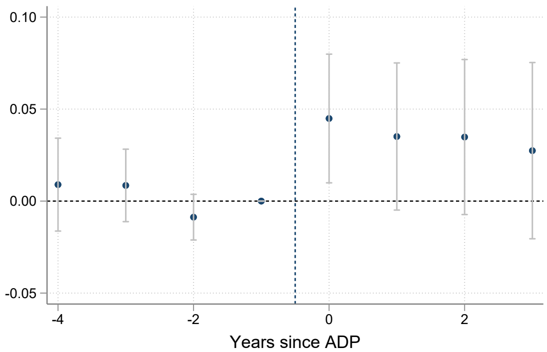
A. Pre-trend



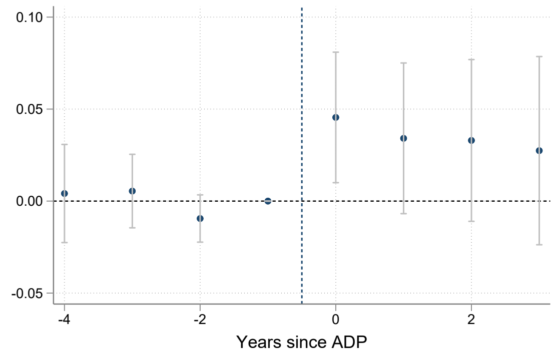
B. Violations from parallel trends

Notes: This figure presents two exercises related to the parallel trends assumption. In Panel A, we present the baseline dynamic figure, but we add the pre-trend that has a 80% power of being detected given the precision of the estimates in the pre-period and the adjusted pre-trend that takes into account the pre-testing bias that arises from the fact that the analysis shown is conditional on passing a pre-test (Roth, 2022). In Panel B, we follow Rambachan and Roth (2021) and estimate the confidence set at 90% for our parameter of interest, allowing for linear and non-linear deviations from the parallel trends assumption. We estimate the confidence set for the coefficient in the year that there was a change in the school principal (year=0). In the case of non-linear deviations, we allow the change in trend from consecutive periods (M) to be as large as the size of the pre-trend that has a 80% power of being detected given the precision of the estimates in the pre-period (Roth, 2022), which is 0.013. In Figure A.3 Panel B, we present the results.

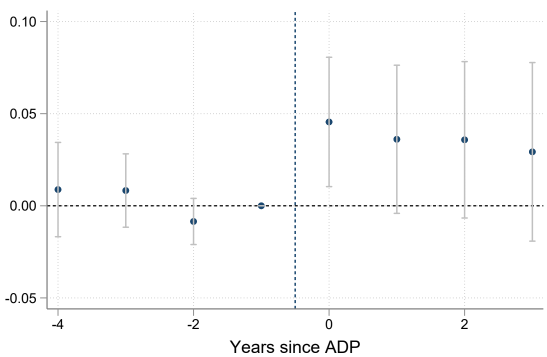
Figure A.4: Principal Selection and Principal Effectiveness: Weighted least squares



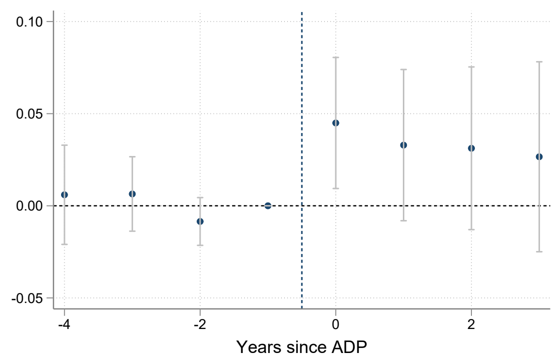
A. Baseline



B. Adds municipal controls



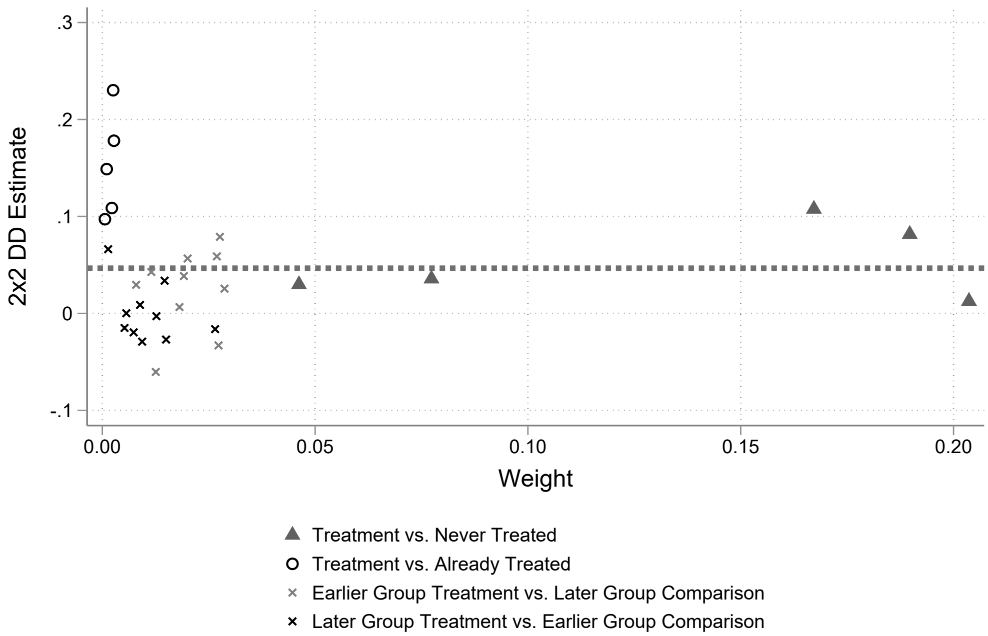
C. Adds school controls



D. Adds both set of controls

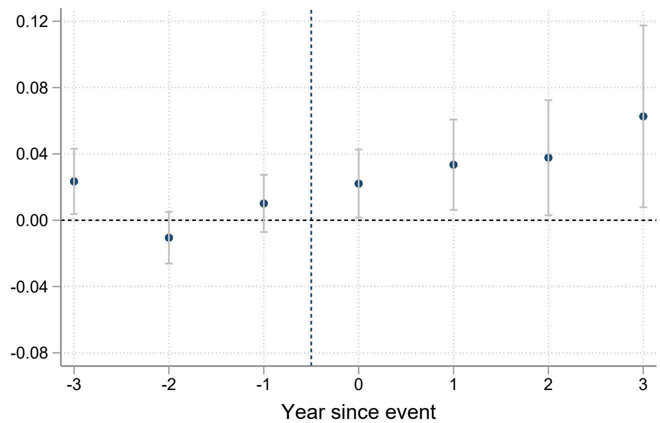
Notes: This figure presents the dynamic version of our *staggered difference-in-differences*. We use as dependent variable the un-adjusted principal fixed effects and we the regressions by the inverse of the standard deviation of the estimate. Panel A presents the dynamic version of the staggered difference-in-differences model suggested by De Chaisemartin and d'Haultfoeuille (2020). Panel B presents the dynamic version of the staggered difference-in-differences suggested by Callaway and Sant'Anna (2020). All panels include confidence intervals at the 95%. In panel B, we cannot reject the null hypothesis of all the coefficients being equal to zero at conventional levels (in the pre-period). The p-value of this test is > 0.09 in Panel A.

Figure A.5: Goodman-Bacon (2021) Decomposition

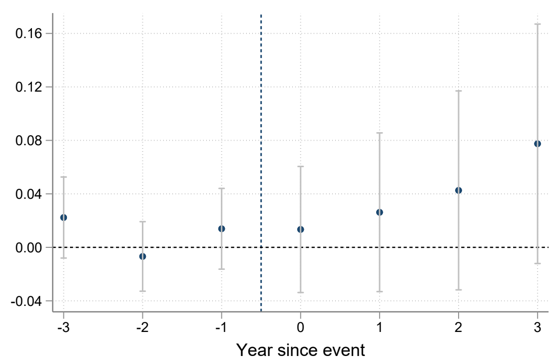


Notes: This figure presents the decomposition of the two-way fixed effect estimator suggested by Goodman-Bacon (2021).

Figure A.6: Principal Selection and Principal Effectiveness within Public Schools



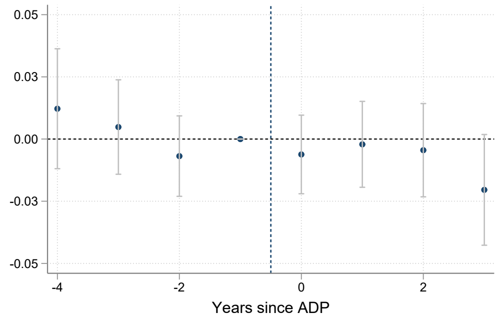
A. De Chaisemartin and d'Haultfoeuille (2020)



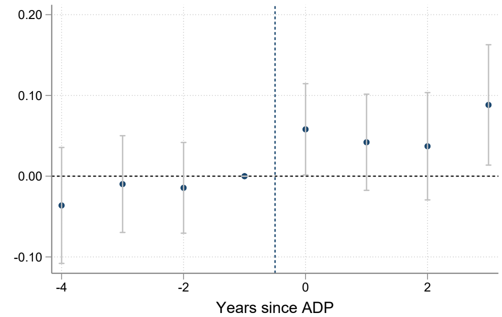
B. Callaway and Sant'Anna (2020)

Notes: This figure presents the dynamic version of our *staggered difference-in-differences* approach in the sample of public schools. Panel A presents the dynamic version of the staggered difference-in-differences model suggested by De Chaisemartin and d'Haultfoeuille (2020). Panel B presents the dynamic version of the staggered difference-in-differences suggested by Callaway and Sant'Anna (2020). All panels include confidence intervals at the 95%. In panel B, we cannot reject the null hypothesis of all the coefficients being equal to zero at conventional levels (in the pre-period). The p-value of this test is > 0.09 in Panel A.

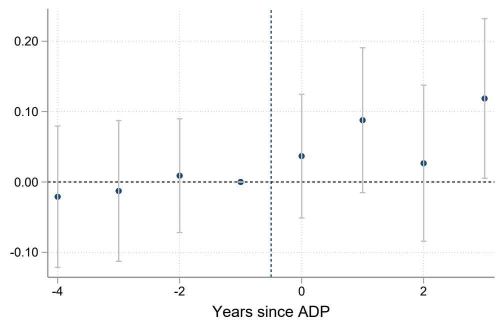
Figure A.7: Principal Selection and Test Scores



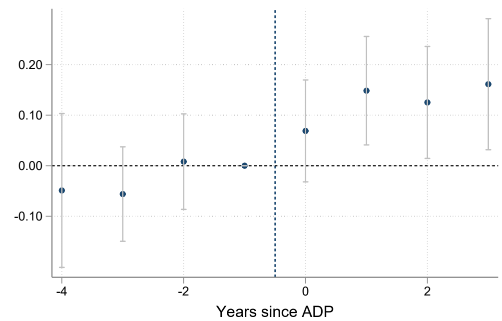
A. Corr GPA-Test scores (SIMCE)



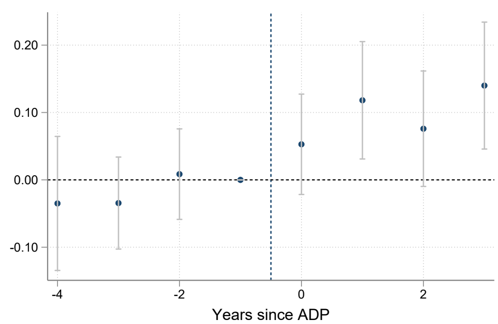
B. Test scores (SIMCE)



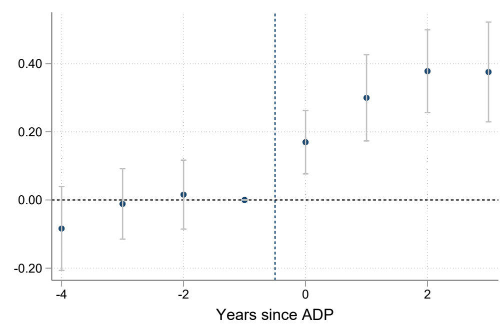
C. Math (PSU)



D. Spanish (PSU)



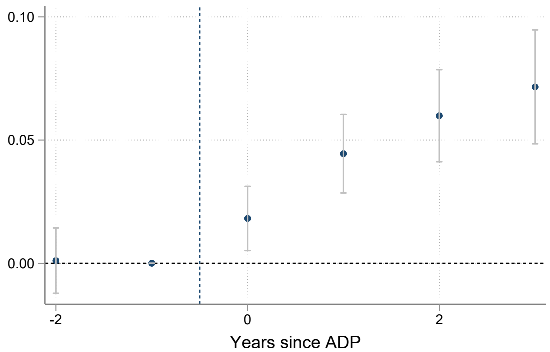
E. Average Math and Spanish (PSU)



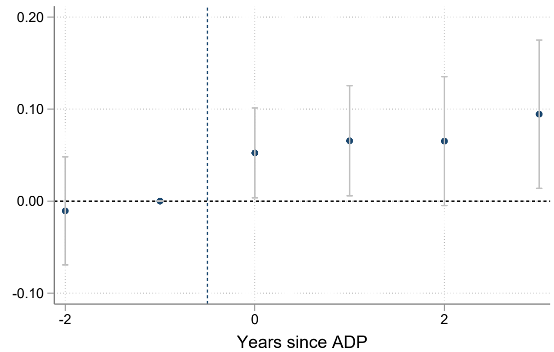
F. Application score (PSU)

Notes: This figure shows the impact of appointments under the new selection system on college admission scores. The figure plots the point estimates and 95% confidence intervals estimated from equation (3). Panel A uses as a dependent variable the within school-year level correlation between standardized test scores in 4th grade and GPA at the student level. Panel B shows the results for the average SIMCE standardized test scores. Panels C and D show the impact on the mandatory exams of Math and Spanish, while Panel E shows the impact of the average between Math and Spanish. Panel F plots the impact on the composite score used for admissions. This score is a weighted average of entry exam scores and course grades, with weights defined by each degree (institution-major pair). We consider the weights of the most preferred degree of a student (as revealed by her preferences in the application process) to construct this score. All panels include school and year fixed effects and also controls by school and municipality characteristics during the pre-reform period (measured in 2010), interacted with year dummies.

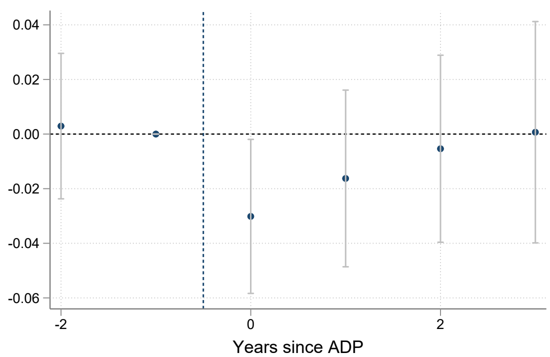
Figure A.8: Principal Selection, GPA, and School Attendance



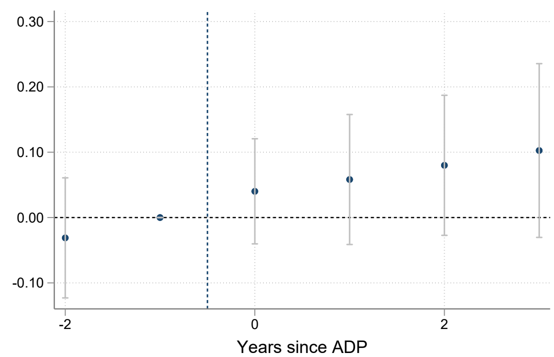
A. GPA: 1 to 8th grade



B. GPA: 9 to 12th grade



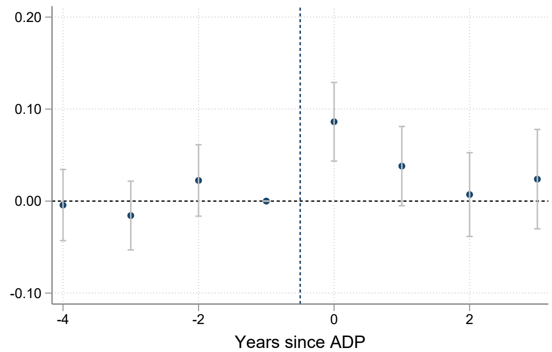
C. Attendance: 1 to 8th grade



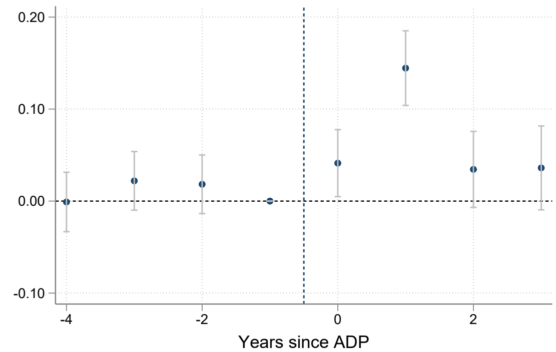
D. Attendance: 9 to 12th grade

Notes: This figure shows the impact of appointments under the new selection system on students' course grades and yearly attendance. The figure plots the point estimates and 95% confidence intervals estimated from equation (3). Panel A presents the results for standardized average GPA for grades 1 to 8, while Panel B shows the results for grades 9 to 12. Panels C and D present the results using students' yearly attendance as the dependent variable, for 1 to 8th and for 9th to 10th grade, respectively. All panels include school and year fixed effects.

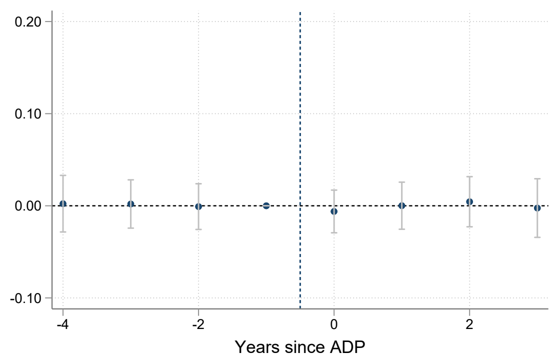
Figure A.9: Principal Selection and School Staff



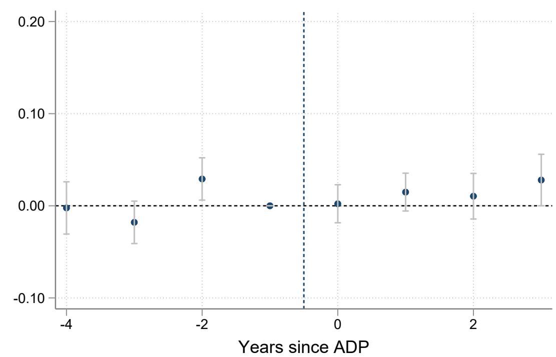
A. Firing: Principal's Team



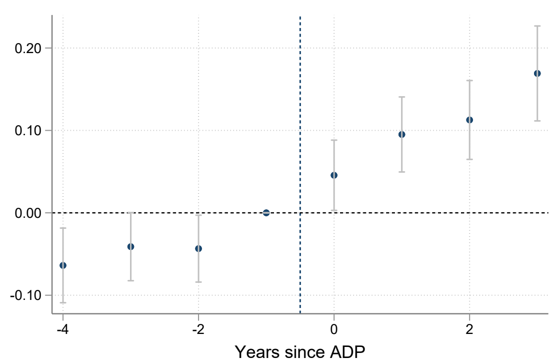
B. Hiring: Principal's Team



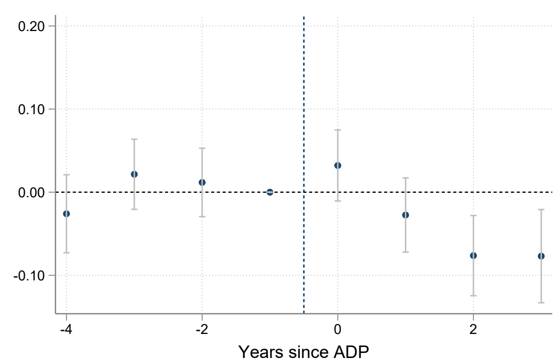
C. Firing: All Teachers



D. Hiring: All Teachers



E. Firing: Teachers with Bad Evaluation



F. Hiring: Teachers with Bad Evaluation

Notes: This figure plots the point estimates and 95% confidence intervals estimated from equation (3). The dependent variable is a dummy that takes the value one if there were any teachers fired (Panels A, C, and E) or any teachers hired (Panels B, D, and F). In Panels A and B, this dummy is based on the principal's team (deputy director, inspector general, and the chief technician), while in Panels C and D, is based on all the teachers' body, while in Panels E and F is based on teachers with poor performance according to teachers evaluations. All panels include school and year fixed effects and also controls by school and municipality characteristics during the pre-reform period (measured in 2010), interacted with year dummies.

Table A.1: Descriptive Statistics in Different Samples

	Full Sample		**Δ Teacher=1**		**LCS=1**	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
	(1)	(2)	(3)	(4)	(5)	(6)
Primary (2-8)	0.8	0.4	0.8	0.4	0.9	0.3
Secondary (9-11)	0.2	0.4	0.2	0.4	0.1	0.3
Subject = Math	0.5	0.5	0.5	0.5	0.5	0.5
Course Grade	5.7	0.6	5.7	0.6	5.7	0.6
% Attendance	91.7	3.3	91.7	3.3	91.8	3.1
% Rural School	0.1	0.3	0.1	0.3	0.1	0.3
% Public School	0.4	0.5	0.4	0.5	0.3	0.5
School Size	794.3	596.6	825.9	618.7	847.4	624.3
Sample Size	12,709,601		9,120,261		7,735,653	

Notes: This table presents descriptive statistics of students in three different samples. "Full Sample" includes all students in our dataset after excluding preschools, adults' schools, and special education schools. We also exclude classes that had more than one teacher per year and eliminate the bottom and top one percent of classroom size outliers. "ΔTeacher = 1" corresponds to the restricted sample of students for whom the teacher, in a given subject, changed between t and $t + 1$. Finally, "LCS" includes all students within the largest connected set of teachers and principals.

Table A.2: Manager Effectiveness and Observable Characteristics

	Principal Effectiveness $\hat{\theta}_p$			
	All		Public	Private
	(1)	(2)	(3)	(4)
Age	0.010	0.010	0.029	-0.000
	(0.003)	(0.003)	(0.008)	(0.004)
Age^2	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Female	0.050	0.049	0.065	0.042
	(0.009)	(0.010)	(0.015)	(0.013)
Perm. Contract	0.019	0.023	-0.000	0.042
	(0.017)	(0.019)	(0.028)	(0.025)
Hours Contract	-0.004	-0.003	-0.008	-0.003
	(0.001)	(0.001)	(0.005)	(0.001)
College Degree	-0.015	-0.033	0.021	-0.097
	(0.015)	(0.016)	(0.023)	(0.022)
Ever Teacher		0.019	-0.005	0.027
		(0.012)	(0.017)	(0.016)
Ever Admin. Supp. Worker		-0.014	-0.040	-0.016
		(0.012)	(0.017)	(0.018)
Ever Admin Worker		-0.034	0.038	-0.027
		(0.033)	(0.055)	(0.040)
Observations	42,022	35,343	15,833	19,503

Notes: This table presents the correlation between the principal effectiveness estimated from equation (1) and principal characteristics. These characteristics include age, gender, experience, type and hours of contract, and indicators for holding a college degree, and for their experience in previous "schooling type" of positions. All specifications include year and municipality fixed effects. Robust standard errors are presented in parentheses.

Table A.3: Teachers' Survey Responses

% Teachers highly agreeing that the principal:	$\hat{\beta}$ (1)	Standard error (2)	Mean Dep Var (3)	Obs (4)	Placebo p-value (5)	RW p-value (6)
Does a good job	0.023	(0.004)	0.460	5,351	0.000	0.001
Can be trusted	0.016	(0.004)	0.521	5,349	0.000	0.001
Makes good decisions	0.024	(0.003)	0.459	6,386	0.000	0.001
Is effective	0.023	(0.004)	0.448	6,382	0.000	0.001
Is good at communicating	0.022	(0.005)	0.529	5,355	0.000	0.001
Engages teachers	0.028	(0.004)	0.444	6,367	0.000	0.001
Engages parents	0.028	(0.003)	0.464	6,386	0.000	0.001
Knows teacher needs	0.027	(0.004)	0.439	6,389	0.000	0.001
Knows student needs	0.029	(0.005)	0.502	5,351	0.000	0.001
Includes teachers	0.025	(0.004)	0.469	7,230	0.000	0.001
Promotes good work climate	0.022	(0.004)	0.525	5,273	0.000	0.001

Notes: To construct this table, we first create an indicator variable at the survey respondent level, which takes a value of one if the survey respondent "highly agrees" with the statement. In cases when the survey had 5 or 4 options, we always use the highest number to create the dummy. Then, we take the average across respondents at the school-year level and assign this to a principal. Columns 1 and 2 report the estimated coefficients and bootstrapped standard errors from a regression on the fraction of the teaching staff highly agreeing with a given statement and our measure of principal effectiveness. To gauge effect sizes, we report the mean of the dependent variable in column 3. Column 5 reports the results from a permutation test for which we randomly reshuffled principal fixed effects 1,000 times. The p-value of the test is calculated as the proportion of sampled permutations s where the value of $\hat{\beta}_s$s was greater than or equal to our estimate $\hat{\beta}$. Finally, column 6 presents p-values adjusted for multiple hypothesis testing using the step-down procedure of Romano and Wolf (2005).

Table A.4: Principals' Effectiveness and the Management of Schools

		Parents' Complaints						Teachers' Turnover		
	Sorting Index	Z-score	Accidents	Infrastructure	Teachers' Absenteeism	Bullying Discrimination	Denied Enrollment	All	High-VA	Low-VA
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Principal Effectiveness	-0.068	-0.024	-0.001	-0.002	-0.002	-0.019	-0.007	-0.002	-0.012	0.011
	(0.069)	(0.007)	(0.002)	(0.001)	(0.001)	(0.006)	(0.003)	(0.001)	(0.001)	(0.001)
	[0.415]	[0.009]	[0.584]	[0.415]	[0.415]	[0.009]	[0.039]	[0.297]	[0.009]	[0.009]
Observations	13,803	10,225	10,225	10,225	10,225	10,225	10,225	42,279	42,279	42,279
Mean Dep Var	5.019	-6.89e-09	0.0442	0.0272	0.0214	0.340	0.114	0.118	0.0571	0.0605
R-squared	0.094	0.099	0.098	0.060	0.067	0.102	0.071	0.106	0.079	0.075
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows the results from a set of regressions of different outcome variables on principal effectiveness. The sorting Index is defined à la Kremer and Maskin (1996) and reflects the amount of variation in classrooms' average course grades that comes from variation between instead of within classrooms. Parents' complaints refer to the number of complaints per 100 students issued by parents for different causes related to the management of the schools. Teacher turnover corresponds to the share of teachers who will leave the school the next year. All regressions include year and municipality fixed effects. Bootstrapped standard errors (100 replications) are clustered at the school principal level. In square brackets, we present p-values that control for the false discovery rate in related groups of outcomes following Romano and Wolf (2005).

Table A.5: Principals' Effectiveness and School Finance

		Expenditures					Income				
	Expenditures to Income Ratio	Log Total Expenditure	Personnel (%)	Learning (%)	Operations (%)	Other (%)	Log Total Income	Subsidies (%)	Self-Revenues (%)	Initial Budget (%)	Other (%)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Principal Effectiveness	-0.202	-0.012	-0.007	-0.030	-0.046	-0.027	0.003	-0.004	-0.066	-0.024	0.021
	(0.195)	(0.016)	(0.014)	(0.024)	(0.020)	(0.026)	(0.014)	(0.014)	(0.033)	(0.036)	(0.015)
Observations	4,122	4,114	4,080	4,006	4,052	4,054	4,115	4,127	2,839	3,342	4,121
R-squared	0.050	0.218	0.214	0.208	0.177	0.160	0.202	0.197	0.239	0.162	0.165
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep Var	0.60	19.97	19.73	16.68	17.23	17.18	20.13	19.83	17.80	17.61	17.14

Notes: This table shows the results from a set of regressions of different measures of school finance on the standardized measure of principal effectiveness discussed in section 3. The data on school finance is only for the year 2016. All the regressions include municipality fixed effects. Robust standard errors are presented in parentheses.

Table A.6: Principal Selection and Principal Effectiveness: Weighted least squares

	(1)	(2)	(3)	(4)
(a) Principal Turnover × ADP	0.046	0.049	0.048	0.047
	(0.024)	(0.025)	(0.024)	(0.025)
(b) Principal Turnover	-0.021	-0.023	-0.022	-0.023
	(0.018)	(0.019)	(0.018)	(0.019)
Observations	30,721	30,721	30,721	30,721
R-squared	0.925	0.925	0.925	0.925
Year FE	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes
School controls	No	Yes	No	Yes
Municipality controls	No	No	Yes	Yes
# of Schools	4,934	4,934	4,934	4,934
pvalue a + b = 0	0.112	0.098	0.111	0.130

Notes: This table presents the effects of the new selection system (ADP) on the un-adjusted standardized measure of principal effectiveness discussed in section 3. All regressions are weighted by the inverse of the standard deviation of each estimated fixed effect. "ADP" is a dummy that takes the value one after the first time a school selects a principal using the new recruitment system. "Principal Turnover" is a dummy that takes the value one after the first time a school selects a new principal (after 2012). Columns 1 to 4 estimate the regressions described by equation (2). Robust standard errors clustered at the school level in parenthesis.

Table A.7: Principal Selection and Principal Effectiveness by Rurality

	Rural	Urban	Rural	Urban
	(1)	(2)	(3)	(4)
Panel A: Based on school status				
(a) Principal Turnover × ADP	0.033	0.049	0.106	0.023
	(0.071)	(0.024)	(0.035)	(0.018)
(b) Principal Turnover	0.069	-0.033		
	(0.065)	(0.019)		
Observations	7,057	23,631	5,763	8,387
# of Schools	1,201	3,825	997	1,450
R-squared	0.933	0.931	0.926	0.927
Mean Dep Var	0.190	-0.051	0.266	0.015
p-value a + b = 0	0.003	0.316		
p-value rural-urban	0.798	0.798	0.057	0.057
Panel B: Based on municipality characteristic				
(a) Principal Turnover × ADP	0.123	0.049	0.058	0.039
	(0.064)	(0.025)	(0.036)	(0.019)
(b) Principal Turnover	-0.075	-0.017		
	(0.056)	(0.019)		
Observations	6,300	24,305	4,395	9,745
# of Schools	1,029	3,905	749	1,639
R-squared	0.934	0.929	0.927	0.924
Mean Dep Var	0.270	-0.0628	0.299	0.0364
p-value a + b = 0	0.179	0.049		
p-value rural-urban	0.331	0.331	0.642	0.642
Year FE	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes

Notes: This table presents the effects of the new selection system (ADP) on principal effectiveness by rurality of the school and municipality. "ADP" is a dummy that takes the value one after the first time a school selects a principal using the new recruitment system. "Principal Turnover" is a dummy that takes the value one after the first time a school selects a new principal (after 2012). All columns estimate the regressions described by equation (2). Columns 1 and 2 present the results for the comparison between public and private schools, while columns 3 and 4 show the results only within the public sector. Columns 1 and 3 (2 and 4) present the results for rural (urban) areas. Panel A uses the definition of rural/urban school from the ministry of education, while Panel B separates municipalities into rural/urban using the median of the empirical distribution of the share of rural population in municipality based on the 2017 Census. Robust standard errors clustered at the school level in parenthesis.

Table A.8: School and Municipality Characteristics, by ADP Adoption

	Never ADP	Ever ADP	Difference	Early ADP	Late ADP	Difference	Private Schools
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: School characteristics							
Total Enrollment	100.330	454.654	354.325	489.115	433.365	-55.750	466.429
	(186.517)	(351.519)	(0.000)	(352.286)	(349.506)	(0.001)	(482.246)
Δ Total Enrollment	-4.870	-14.698	-9.827	-16.734	-13.428	3.306	-0.564
	(21.439)	(45.673)	(0.000)	(51.932)	(41.269)	(0.136)	(44.728)
Rural School	0.838	0.256	-0.582	0.224	0.276	0.051	.235
	(0.369)	(0.437)	(0.000)	(0.418)	(0.447)	(0.015)	(0.423)
Income per student	19.631	6.455	-13.175	6.471	6.446	-0.025	8.094
	(25.551)	(1.859)	(0.000)	(1.798)	(1.896)	(0.784)	(11.538)
Δ Income per student	0.595	-0.055	-0.650	-0.099	-0.028	0.072	-.064
	(15.703)	(1.173)	(0.080)	(1.739)	(0.590)	(0.209)	(6.139)
Share of disadvantaged students	0.702	0.508	-0.193	0.493	0.518	0.026	0.410
	(0.220)	(0.225)	(0.000)	(0.240)	(0.214)	(0.018)	(0.283)
Δ Share of disadvantaged students	-0.037	-0.014	0.023	-0.010	-0.016	-0.006	-0.004
	(0.175)	(0.114)	(0.000)	(0.121)	(0.108)	(0.249)	(0.110)
4rd grade test scores (Spanish)	255.565	255.610	0.045	256.303	255.203	-1.100	267.873
	(30.287)	(22.027)	(0.961)	(21.944)	(22.077)	(0.344)	(26.146)
Δ 4rd grade test scores (Spanish)	-1.590	-2.637	-1.047	-1.792	-3.140	-1.348	-4.876
	(33.854)	(22.127)	(0.308)	(21.959)	(22.223)	(0.251)	(21.871)
4rd grade test scores (Math)	238.975	245.459	6.484	245.793	245.262	-0.531	256.763
	(33.034)	(24.510)	(0.000)	(24.177)	(24.715)	(0.682)	(256.76)
Δ 4rd grade test scores (Math)	9.237	8.194	-1.043	8.549	7.982	-0.567	4.134
	(34.392)	(23.280)	(0.324)	(22.527)	(23.727)	(0.646)	(22.926)
Graduation test score (Spanish)	414.049	436.469	22.420	439.418	434.231	-5.188	513.328
	(43.775)	(57.979)	(0.000)	(55.863)	(59.561)	(0.380)	(76.776)
Δ Graduation test score (Spanish)	-4.119	-4.079	0.040	-3.011	-4.896	-1.885	-2.173
	(29.886)	(25.310)	(0.989)	(21.255)	(28.042)	(0.475)	(25.191)
Graduation test score (Math)	418.480	441.136	22.657	441.347	440.977	-0.370	516.988
	(41.965)	(55.204)	(0.000)	(54.580)	(55.794)	(0.948)	(80.753)
ΔGraduation test score (Math)	-7.798	-3.738	4.061	-3.929	-3.591	0.337	-3.281
	(31.307)	(25.322)	(0.155)	(21.378)	(28.017)	(0.898)	(24.495)
Panel B: Municipality characteristics							
Share of households in poverty	0.124	0.082	-0.042	0.082	0.082	-0.000	0.0933
	(0.075)	(0.056)	(0.000)	(0.057)	(0.055)	(0.952)	(0.076)
Income per capita	1.699	2.151	0.453	2.223	2.107	-0.115	2.386
	(0.489)	(1.115)	(0.000)	(1.400)	(0.892)	(0.033)	(1.652)
Unemployment rate	0.080	0.080	0.001	0.083	0.079	-0.004	0.083
	(0.047)	(0.047)	(0.626)	(0.050)	(0.045)	(0.064)	(0.042)
Average years of schooling	8.974	9.998	1.024	9.930	10.041	0.110	10.286
	(1.124)	(1.315)	(0.000)	(1.385)	(1.269)	(0.083)	(1.567)
Observations	3,029	1,820	4,849	695	1,125	1,820	3,782

Notes: This table presents the differences between public schools that have selected principals under the ADP system and schools that have not. It also shows the differences between early (2012-13) adopters and late (post-2014) adopters of the ADP selection system. All characteristics are measured in 2010 (pre-reform). Δ represents the first difference of the predetermined (pre-reform) school characteristic. Columns 1 and 2 present the statistics for ADP and non-ADP, while column 3 presents the difference and the p-value of the difference (in parenthesis). Columns 4 and 5 present the statistics for early and late adopters, while column 6 presents the difference between both and the p-value of the difference. Finally, column 7 presents summary statistics for all private schools.

Table A.9: Principal Selection and Principal Effectiveness: Placebos

	Public Schools (pre-reform)		Private Schools (post-reform)	
	(1)	(2)	(3)	(4)
Principal Turnover	0.006	0.021	-0.032	-0.026
	(.095)	(0.110)	(.019)	(0.036)
Observations	5,308	5,308	17,502	17,502
# of Schools	1,668	1,668	2,802	2,802
Year FE	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes

Notes: This table presents the estimate from our placebo exercise looking at the impact of non-ADP principal turnovers on the standardized measure of principal effectiveness discussed in section 3. In columns 1 and 2, "Principal turnover" is a dummy that takes the value one after a principal turnover in a public school in the period 2009-2010 (pre-ADP reform). The number of schools that had a principal turnover in 2009 or 2010 is 292. In columns 3 and 4, "Principal turnover" is a dummy that takes the value one after the first time a private school selects a new principal (after 2012). Columns 1 and 3 show the estimates from the model suggested by De Chaisemartin and d'Haultfoeuille (2020), while columns 2 and 4 show the estimates from the model suggested by Callaway and Sant'Anna (2020). Robust standard errors clustered at the school level in parentheses.

Table A.10: Principal Selection and Principal Effectiveness: Controlling for school markets

		Based on a distance of:			
Market:	Municipality	4kms	5kms	6kms	7kms
	(1)	(2)	(3)	(4)	(5)
(a) Principal Turnover × ADP	0.058	0.047	0.050	0.057	0.057
	(0.023)	(0.025)	(0.024)	(0.024)	(0.024)
(b) Principal Turnover	-0.027	-0.025	-0.023	-0.027	-0.024
	(0.018)	(0.019)	(0.019)	(0.018)	(0.018)
Observations	30,440	26,914	27,944	28,886	29,564
# of Schools	4,908	4,347	4,505	4,654	4,756
R-squared	0.937	0.937	0.936	0.935	0.933
Year FE	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes
School Market-Year	Yes	Yes	Yes	Yes	Yes
# of markets	295	276	220	178	127
p-value a + b = 0	0.053	0.184	0.096	0.058	0.034

Notes: This table presents the effects of the new selection system (ADP) on the standardized measure of principal effectiveness discussed in section 3. "ADP" is a dummy that takes the value one after the first time a school selects a principal under the ADP system. "Principal Turnover" is a dummy that takes the value one after the first time a school selects a new principal (after 2012). All columns estimate the regressions described by equation (2), but adds a school market fixed effect interacted with year fixed effects. In column 1, we define a school market based on the municipality in which the school is located. In columns 2 to 4, we follow Cuesta et al. (2020) and Neilson (2019) and define a market based on the distance between schools. In Chile, the students' average distance to chosen schools is 2kms and the 90th percentile of such distribution is 5kms (Cuesta et al., 2020). Based on this, we define a market based on the schools that are close to each other. We use four different diameters (k) to define a market 4, 5, 6, and 7kms. We implement this by creating a symmetric adjacency matrix (A), where the element $A(i, j)$ takes the value one if i and j are less than k kms away from each other. Then, we construct the set of "connected components" of the matrix, where a component is defined as a set of schools where one can always find a "path" that connects two pairs of schools. Robust standard errors clustered at the school level in parenthesis.

Table A.11: Principal Selection, GPA, and School Attendance

	GPA		Attendance	
Grades:	1 to 8	9 to 12	1 to 8	9 to 12
	(1)	(2)	(3)	(4)
(a) Principal Turnover × ADP	0.046	0.072	-0.016	0.108
	(0.007)	(0.025)	(0.011)	(0.046)
(b) Principal Turnover	-0.018	-0.015	-0.001	-0.039
	(0.005)	(0.010)	(0.008)	(0.011)
Observations	25,178	9,459	25,178	9,458
R-squared	0.832	0.801	0.786	0.782
Year FE	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes
Mean dep var	0.079	-0.178	0.073	-0.026
p-value a + b = 0	0.000	0.017	0.054	0.117

Notes: This table presents the effects of the new selection system (ADP) on average GPA and average attendance rate. "ADP" is a dummy that takes the value one after the first time a school selects a principal using the new recruitment system. "Principal Turnover" is a dummy that takes the value one after the first time a school selects a new principal (after 2012). All columns estimate the regressions described by equation (2). In columns 1 and 2 we consider standardized course grades as the dependent variable and in columns 3 and 4 we consider standardized yearly attendance as the dependent variable. Robust standard errors clustered at the school level in parenthesis.

Table A.12: Principal Selection, Students' Sorting, and Teachers' Surveys

	Sorting Index	Does a good job	Can be trusted	Makes good decisions	Is effective	Is good at communicating	Is a good manager	Engages teachers	Engages parents	Knows teacher needs	Knows student needs	Includes teachers	Promotes good work climate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(a) Principal Turnover × ADP	-0.003	0.052	0.071	0.020	0.017	0.086	-0.004	0.048	0.011	0.009	0.088	-0.016	0.027
	(0.002)	(0.043)	(0.044)	(0.022)	(0.022)	(0.043)	(0.022)	(0.021)	(0.022)	(0.022)	(0.043)	(0.017)	(0.044)
(b)Principal Turnover	0.001	-0.045	-0.036	0.019	0.036	-0.061	0.021	-0.010	0.006	0.032	-0.052	0.032	-0.024
	(0.001)	(0.031)	(0.031)	(0.015)	(0.015)	(0.032)	(0.015)	(0.015)	(0.014)	(0.015)	(0.031)	(0.012)	(0.032)
Observations	12,547	11,658	11,604	16,808	15,969	11,674	16,522	16,364	16,177	16,542	11,551	21,142	10,520
R-squared	0.560	0.511	0.485	0.456	0.479	0.496	0.476	0.460	0.466	0.468	0.516	0.425	0.513
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep var	0.050	0.472	0.533	0.467	0.454	0.540	0.488	0.452	0.471	0.444	0.511	0.470	0.532
p-value a + b = 0	0.309	0.831	0.289	0.030	0.003	0.426	0.332	0.030	0.352	0.019	0.246	0.244	0.922

Notes: This table presents the effects of the new selection system (ADP) on answers from the teachers' surveys. "ADP" is a dummy that takes the value one after the first time a school selects a principal under the ADP system. "Principal Turnover" is a dummy that takes the value one after the first time a school selects a new principal (after 2012). All columns estimate the regressions described by equation (2). The dependent variable in column 1 corresponds to the sorting index discussed in section 3. To construct the dependent variables in columns 2-13, we first create an indicator variable at the survey respondent level, which takes a value of one if the survey respondent is "highly agree" with the statement. In cases when the survey had 5 or 4 options, we always use the highest number to create the dummy. Then, we take the average across respondents at the school-year level and assign this to a principal. Robust standard errors clustered at the school level in parenthesis.

Table A.13: Principal Selection and Principals' Characteristics

	Ever teacher	Ever administrative	Ever principal	Ever in private sector	Age	Female	College degree
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Baseline							
(a) Principal Turnover × ADP	-0.009	-0.012	-0.029	0.022	-0.254	0.020	0.059
	(0.025)	(0.007)	(0.015)	(0.010)	(0.458)	(0.023)	(0.016)
(b) Principal Turnover	0.121	-0.007	0.011	0.001	-4.278	-0.003	0.023
	(0.019)	(0.007)	(0.014)	(0.009)	(0.384)	(0.017)	(0.011)
R-squared	0.778	0.476	0.550	0.919	0.728	0.752	0.742
p-value a + b = 0	0.000	0.000	0.0158	0.000	0.000	0.261	0.000
Panel B: Adds county-year fixed effects							
(a) Principal Turnover × ADP	-0.006	-0.010	-0.025	0.027	-0.098	0.011	0.065
	(0.025)	(0.007)	(0.015)	(0.011)	(0.472)	(0.024)	(0.016)
(b) Principal Turnover	0.124	-0.010	0.005	0.002	-4.643	0.005	0.024
	(0.019)	(0.007)	(0.014)	(0.009)	(0.392)	(0.017)	(0.011)
R-squared	0.800	0.520	0.589	0.925	0.750	0.772	0.763
p-value a + b = 0	0.000	0.000	0.0199	0.000	0.000	0.331	0.000
Panel C: Adds school market-year fixed effects							
(a) Principal Turnover × ADP	-0.009	-0.014	-0.033	0.025	-0.272	0.018	0.059
	(0.025)	(0.007)	(0.015)	(0.011)	(0.467)	(0.023)	(0.016)
(b) Principal Turnover	0.118	-0.008	0.011	0.000	-4.348	0.002	0.022
	(0.019)	(0.007)	(0.014)	(0.009)	(0.388)	(0.017)	(0.011)
R-squared	0.782	0.490	0.563	0.921	0.736	0.762	0.753
p-value a + b = 0	0.000	0.000	0.005	0.000	0.000	0.206	0.000
Observations	24,734	29,564	24,734	24,734	29,564	29,564	29,564
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep var	0.356	0.978	0.932	0.117	54.36	0.547	0.878

Notes: This table presents the effects of the new selection system (ADP) on principals' characteristics. "ADP" is a dummy that takes the value one after the first time a school selects a principal using the new recruitment system. "Principal Turnover" is a dummy that takes the value one after the first time a school selects a new principal (after 2012). All columns estimate the regressions described by equation (2). In Panel B, we define a school market based on the municipality in which the school is located and add an interaction between county and year fixed effects. In Panel C, we consider the market definition using the 7km distance. Robust standard errors clustered at the school level in parenthesis.

Table A.14: Principals' Characteristics by ADP Status

	Public Schools			
	Not ADP	ADP	Difference	Private Schools
	(1)	(2)	(3)	(4)
Panel A: Ever worked				
As teacher	0.541	0.441	-0.099	0.431
	(0.498)	(0.497)	(0.000)	(0.495)
As administrative before	0.950	0.964	0.014	0.937
	(0.217)	(0.185)	(0.034)	(0.244)
In private sector	0.009	0.035	0.025	0.229
	(0.096)	(0.183)	(0.000)	(0.420)
Panel B: Principal characteristics				
College degree	0.839	0.901	0.063	0.893
	(0.368)	(0.299)	(0.000)	(0.309)
Age	57.214	55.780	-1.434	54.294
	(8.762)	(8.940)	(0.000)	(11.979)
Female	0.490	0.489	-0.001	0.615
	(0.500)	(0.500)	(0.962)	(0.486)
Observations	2,057	1,770	3,827	4,434

Notes: This table compares the characteristics of public schools' principals who have been appointed under the ADP system and those who have not. Columns 1 and 2 present the average and standard deviation of different characteristics, and column 3 presents the difference among these two groups and its p-value (in parentheses). Finally, column 4 presents the average and standard deviation for school principals at private schools.

Table A.15: Market Level Treatments and Principals' Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Ever teacher	Ever administrative	Ever principal	Ever in private sector	Age	Female	College degree
Panel A: County							
Post ADP turnover in the market	-0.003	0.000	0.008	0.001	-0.401	-0.008	0.005
	(0.011)	(0.004)	(0.008)	(0.005)	(0.255)	(0.010)	(0.006)
R-squared	0.787	0.530	0.589	0.931	0.759	0.787	0.795
Panel B: School market							
Post ADP turnover in the market	-0.013	0.006	0.003	0.011	-0.684	0.000	0.016
	(0.019)	(0.008)	(0.014)	(0.008)	(0.450)	(0.021)	(0.008)
R-squared	0.787	0.530	0.589	0.931	0.759	0.787	0.795
Observations	14,298	16,945	14,298	14,298	16,945	16,945	16,945
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep var	0.320	0.972	0.916	0.191	53.50	0.607	0.886

Notes: This table presents the effects of the new selection system (ADP) on the relevant school market on principals' characteristics in private schools. "Post ADP turnover in the market" is a dummy that takes the value one after the first time a school selects a principal under the ADP system in a given market. The sample includes only private schools. In Panel A, the market is defined as the municipality in which the school is located, while in Panel B, the market is defined using the 7km distance. Robust standard errors clustered at the school level in parenthesis.

Table A.16: Characteristics of Origin and Destination Schools of ADP principals

	School of Origin	School of Destination	Mean Difference
	(1)	(2)	(3)
Panel A: School characteristics			
Monthly principal wage (1,000 USD)	2.572	2.575	0.003
	(0.887)	(0.607)	(0.029)
Monthly school wage (1,000 USD)	1.323	1.350	0.027
	(0.270)	(0.243)	(0.009)
Share of disadvantaged students	34.922	62.684	27.762
	(23.404)	(16.754)	(0.717)
Average test scores	247.614	247.668	0.054
	(19.246)	(19.315)	(0.765)
Total enrollment	458.944	432.487	-26.457
	(351.421)	(321.265)	(11.863)
Income per student	8.048	10.696	2.648
	(3.942)	(3.179)	(0.126)
Rural school	0.247	0.221	-0.026
	(0.413)	(0.411)	(0.015)
Panel B: Municipality characteristics			
Share of households in poverty	0.073	0.033	-0.040
	(0.054)	(0.019)	(0.002)
Income per capita	2.358	3.489	1.131
	(1.244)	(1.830)	(0.075)
Unemployment rate	0.081	0.079	-0.002
	(0.044)	(0.026)	(0.002)
Average years of schooling	10.126	10.833	0.707
	(1.354)	(1.278)	(0.068)
Observations	1,611	1,611	3,222

Notes: This table compares the school of origin and destination of principals elected by the new ADP selection system. Columns 1 and 2 present the average and standard deviation of different characteristics of the schools and the municipalities where schools are located. Column 3 presents the mean difference between these two groups and the standard deviation of the difference (in parentheses).

B Additional Specification Checks

In our setting, principal fixed effects would identify the causal effect of principals on students under a *strict exogeneity* or *selection on observables* assumption, i.e., conditional on observable characteristics and teacher fixed effects, the correlation between the assignment of students to principals and other determinants of students achievement is innocuous. Although this identification assumption is ultimately untestable —what Holland (1986) called "the fundamental problem of causal inference"— we can leverage our data to implement some of the validation exercises proposed in the literature.

First, in the spirit of Chetty et al. (2014) and the omnibus test in Angrist et al. (2017), we present quasi-experimental evidence from an analog to the ideal experiment of random principal assignment to schools. This design exploits principal turnover for identification, thus it rests on the identification assumption that principal turnover within a school is uncorrelated with student and school characteristics.[39] We begin with event studies looking at the evolution of course grades around the events of entry and exit of low and high value-added principals (Figure B.1). For this exercise, we restrict the sample to the subset of principals who switched schools between 2011 and 2016 (the period for which course grade data is available), and who belong to the top or bottom 25% of the principal effectiveness distribution. Let year 0 denote the school year that a principal enters or exits a school and define all other school years relative to that year (e.g., if the principal enters in 2013, year 2011 is -2 and year 2015 is +2). We define an entry event as the arrival of a principal whose effectiveness is either in the top or bottom quartile of the distribution of principal effectiveness, and we define exit events analogously. The series in Figure B.1 plots school-year means of standardized course grades in the two years before and after a low value-added principal exits the school. As in Chetty et al. (2014), we do not condition on any other covariates in this figure: each point simply shows average course grades for different years within a school. Consistent with the idea that our estimates of principal effectiveness are forecast unbiased, the null hypothesis that the observed impact on mean gains equals the increase in principal effectiveness cannot be rejected. In all but the last panel (D), the change in course grade gains is significantly different from 0 with p-values < 0.01 and is not significantly different from what one would forecast based on the change in mean principal effectiveness. These event studies show that student achievement changes sharply across time *as predicted* by the change in principal effectiveness, when high or low value-added

[39] Although untestable, this assumption is plausible insofar as teachers and students are unlikely to immediately switch to a different school because the principal changed.

principals enter or exit a school.

Second, two-way fixed effects specifications are simple and tractable. Nevertheless, when used for estimating worker and firm fixed effects, these specifications are prone to be criticized (see Card et al., 2018 for a discussion).[40] Since our model also considers additive teacher and principal effects, one might be worried about the bias in our measure of principal effectiveness. We address this issue in the spirit of Card et al. (2013) and plot the mean course grades of the students taught by teacher j before and after the teacher started working under a new principal p. For this, we first residualize course grades using all controls in our main specification (including lagged course grades), but excluding teachers' and principals' fixed effects. Figure B.2 presents these profiles. We see that teachers who moved from working under a principal with students in the lowest (1st) quartile of course grades to working under a principal with students in the highest (4th) quartile experienced a large average gain in their students' course grade, while those who moved in the opposite direction experienced large loses. Moving within a quartile group, by comparison, is associated with relatively small changes in residualized course grades. Moreover, although we do not condition on holding teacher-principal relationships for at least 2 years, the trends prior and after moving are very similar across groups, and the mean change in course grades for teachers who move in opposite directions between quartile groups (e.g, from quartile 1 to quartile 2, versus from quartile 2 to quartile 1) are of similar magnitude and uniformly of opposite sign. While not perfect, this figure is consistent with the symmetry implications of the additive two-way fixed effects model with exogenous mobility.

Third, to assuage concerns related to student sorting we follow Jackson et al. (2022) and show that conditional on our school level controls, predicted outcomes based on individual characteristics are unrelated to our measure of principal effectiveness. Specifically, we focus on students who took the SIMCE national exams at some point, for whom we have the following attributes: family income category (low, medium, high), parents' education (college graduate or not), and parents' ethnicity. Then, we predict course grades based on a linear regression of course grades on students' attributes, grade and year. Figure B.3, Panel A, shows a binned scatterplot of the actual course grades against the predicted course grades. The predicted outcomes tracks actual outcomes very well. We then examine if principal effectiveness is correlated with predicted course grades in a regression model with only the time-varying and across-time averaged school characteristics used in our correlated random

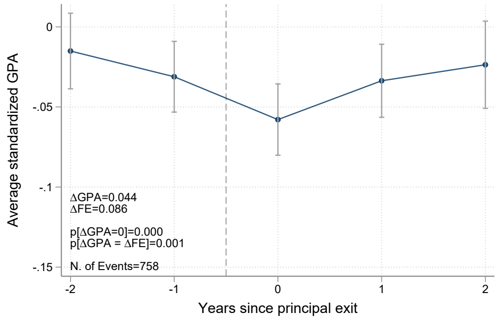
[40]This is because OLS estimates of worker and firm effects will be biased unless worker mobility is uncorrelated with the time-varying residual components of wages, a strong assumption on workers' mobility if one considered some specific models of wage determination (e.g., Gibbons et al., 2005).

effect approach. Figure B.3, Panel B, shows a binned scatterplot of the predicted course grades against our measure of principal effectiveness. We find that, after including our school-level controls, principal effectiveness is not significantly related to predicted course grades.

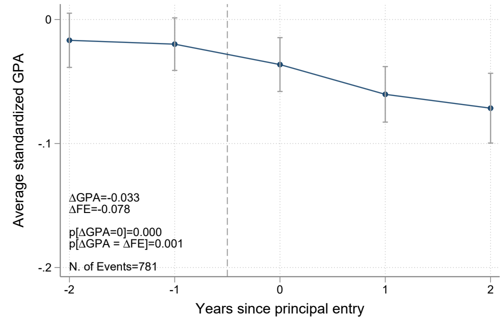
Finally, to complement the previous exercises, we also implement a falsification test similar to that in Rothstein (2010). We focus on a subset of students who switched schools at the end of primary and who were consequently exposed to more than one principal. The intuition of the test is simple: if the effectiveness of the principal in the school of destination s_{t+1} impacts GPA growth in the school of origin s_t, that would be evidence of model misspecification. We consider two sets of students. First, students who were "forced" to switch to another school because their school of origin *did not* offer secondary education. Second, students who switched from schools that *did* offer secondary education. For this exercise, we use "jackknife" estimates of principal effectiveness as the dependent variable, i.e., estimates of principal effectiveness in a sample that leaves out all observations of the students who switched schools. As shown by Table B.1, we fail to find evidence of a positive correlation between growth in course grades (the *pre-assignment* variable) and the effectiveness of their future principal (the *treatment* variable), in both cases.[41]

[41] It is worth noticing that failing to reject the null hypothesis that future principals have an impact on current achievement does *not* guarantee that there is no sorting. Consequently, we take this evidence only as suggestive.

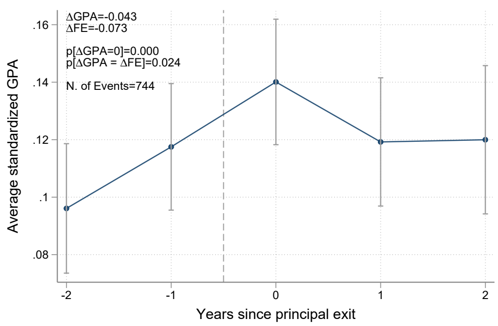
Figure B.1: Impacts of Principal Entry and Exit on Students' Performance



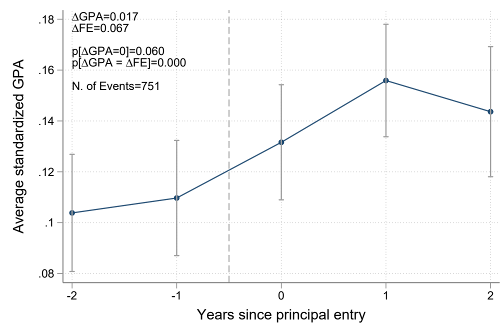
A. Low Value-Added Principal Exit



B. Low Value-Added Principal Entry



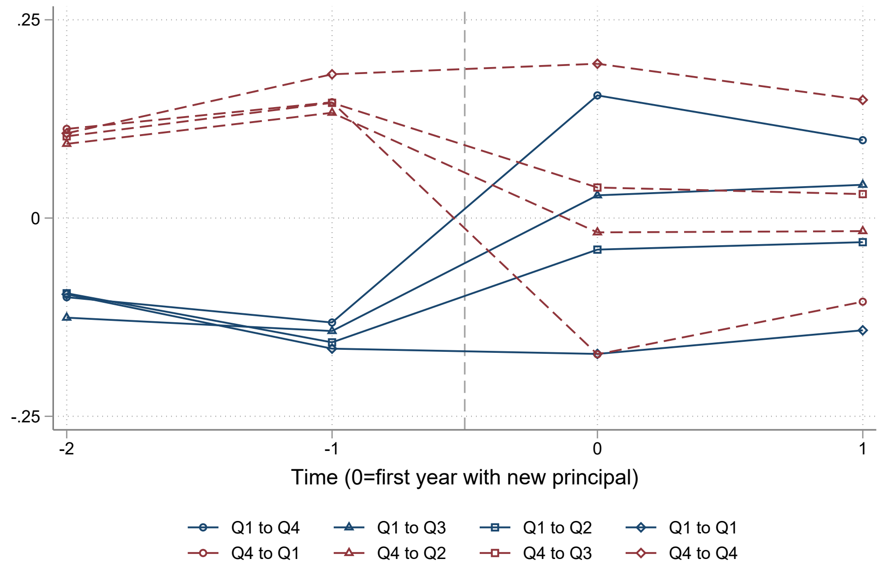
C. High Value-Added Principal Exit



D. High Value-Added Principal Entry

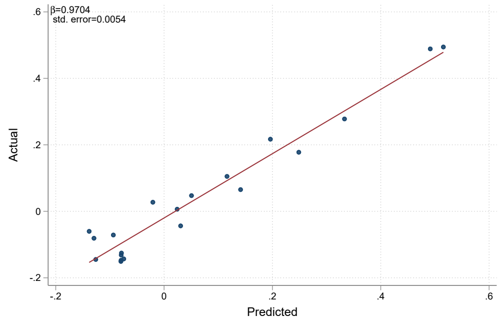
Notes: These figures plot event studies of standardized course grades as principals arrive at or leave a school at year t=0. Panels A and B plot the impact of the exit and entry of a low value-added principal (principals with VA in the bottom 25% of the distribution) on mean course grades. Likewise, Panels C and D plot the impact of the exit and entry of a high value-added principal (principals with VA in the top 25% of the distribution) on mean course grades. To construct each panel, we first identify the set of principals who entered or exited a school between 2012 and 2015 and define event time as the school year relative to the year of entry or exit. We only include observations where we observe students before and after the change in principal and plot the average course grade and their confidence interval at the 95% (based on the standard error of the mean) for each relative year across principal's turnover. Each panel reports the change in mean grades' gains (current minus lag grades) from t=-1 to t=1 and the change in mean estimated VA. We report p-values from a test of the hypotheses that the change in achievement gains from t=-1 to t=1 equals the change in VA and that the change in achievement gains equals 0.

Figure B.2: Mean Residualized GPA of Teachers who change Principal, classified by Quartile of Principals' Mean Residualized GPA at Origin and Destination

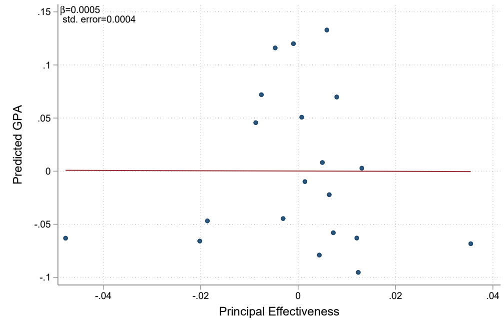


Notes: This figure plots the mean residualized course grades of teachers who changed principal between 2011 and 2016. We consider the first time a teacher switches to work under a new principal, but we do not condition on holding the old or new job relationship for a minimum number of years. Each principal is classified into quartiles based on mean residualized course grades of the students at her school. Course grades are residualized with respect to the same set of controls considered in our main specification (1), except teacher and principal fixed effects.

Figure B.3: Predicted Course Grades and Principal Effectiveness



A. Actual GPA and Predicted GPA



B. Principal Effectiveness and Predicted GPA conditional on School CRE

Notes: Panel A shows a binned scatterplot of the actual course grades against the predicted course grades. Panel B shows a binned scatterplot of the predicted course grades (after removing school controls) against our measure of principal effectiveness (after removing school controls). Each figure reports the coefficient from a regression of the y-axis variable on the x-axis variable. In panel B we leverage the Frisch–Waugh–Lovell theorem to obtain the β coefficient. We predict course grades based on a linear regression of course grades on student attributes, which are available for students who took the SIMCE national exams at some point.

Table B.1: Falsification Test

	"Forced" Switches		"Non-Forced" Switches	
	$\theta_{p(-i)}$ at school s_{t+1}	$\theta_{p(-i)}$ at school s_t	$\theta_{p(-i)}$ at school s_{t+1}	$\theta_{p(-i)}$ at school s_t
	(1)	(2)	(3)	(4)
Growth GPA	0.001	0.013	0.002	0.015
	(0.001)	(0.002)	(0.002)	(0.003)
Observations	96,448	96,448	39,257	39,257

Notes: This table shows the results from the validation exercise discussed above. We consider a sample of students who switched schools at the end of primary. In columns 1 and 2, we consider students "forced" to switch because their school *did not* offer secondary education. In columns 3 and 4, we consider students who switched from schools that *did* offer secondary education. For this exercise, we use "jackknife" estimates of principal effectiveness as the dependent variable, i.e., estimates of principal effectiveness in a sample that leaves out all observations of the students who switched schools. Robust standard errors clustered at the county level are presented in parentheses.

C Principals' Labor Market

C.1 Descriptive Analysis of Wages at Public and Private Schools

Public sector compensation usually does not include pay for performance (Finan et al., 2017), and although there is a good rationale for this,[42] it has been argued that fixed compensation schemes make it difficult to attract and keep the best personnel in public schools. This discussion, which has motivated several studies on the effects of pay for performance (Rothstein, 2015; Cullen et al., 2016; Biasi, 2021) and teachers' firing policies (Staiger and Rockoff, 2010; Boyd et al., 2011; Cowen and Winters, 2013), is also relevant to the Chilean case. To study this, we use administrative data on wages from public and subsidized private schools from 2015 to 2017. Figure C.1 presents some features of our data. Perhaps not surprisingly, we find that hourly wages (residualized with respect to year and municipality fixed effects) at public schools are significantly less spread and 0.09 log points lower than those at the voucher-private schooling sector. Like in the US, wages in Chilean public schools also rely less on pay-for-performance. On average, the bonus component of wages represents 22% of the principal's salary in voucher-private schools but only 9% in public schools.

To study whether workers' characteristics command the same price in public and voucher schools, we estimate the following Mincer type regression model:

$$\ln(\text{wage}_{pt}) = \alpha + \beta_0 \text{Voucher}_{pt} + \beta_1 [X_{pt} - \bar{X}] + \beta_2 \text{Voucher}_{pt} \times [X_{pt} - \bar{X}] + \rho_{m(p,t)} + \gamma_t + \epsilon_{pt}, \quad (4)$$

where $\ln(\text{wage}_{pt})$ represents the logarithm of the average hourly wage paid to principal p at time t, Voucher_{pt} is an indicator that equals one if the principal works at a voucher-private school (and zero otherwise), γ_t are year fixed effects, and $\rho_{m(p,t)}$ is a fixed effect at the level of the municipality in which principal p works at time t. The parameter of interest is β_2, and it represents the factor price differential between sectors. Importantly, the vector X_{pt} includes principal characteristics such as our measure of her effectiveness $\hat{\theta}_p$, tenure, tenure squared, an indicator for whether the principal is female, and for whether she has a permanent contract. This specification also allows us to study how the different components of wages relate to principal effectiveness. For this, we decompose the dependent variable $\ln(\text{wage}_{pt})$ into two components: $\ln(\text{base}_{pt})$ and $\ln(\text{wage}_{pt}/\text{base}_{pt})$, where "base" corresponds to the sum of the minimum legal wage and the statutory payments described in section 2,

[42] Performance pay for bureaucrats can create severe multi-tasking problems, where bureaucrats focus on the incentivized dimension of their job at the expense of the non-incentivized dimension (Holmstrom and Milgrom, 1987).

and base_{pt} corresponds to the total wage minus the bonuses.

Table C.1 presents the point estimates and bootstrap standard errors (100 replications) obtained from these regressions. Columns 1 and 2 show the association between the log wage of school principals and their characteristics, while columns 3 to 6, replicate this analysis but decompose log wages into its base and a bonus component. Our estimates reveal a sizable and statistically significant wage premium in voucher-private schools. On average, voucher schools pay 14% more than public schools, and most of this premium is driven by the bonus components of wages. Regarding the association between wages and principals' effectiveness, we fail to reject the null of no association between the variables in public schools; however, we find a modest, although statistically significant, association at voucher-private schools where increasing principal effectiveness by one standard deviation is associated with a 7.5% increase in wages, a correlation that is also driven by the bonus components of wages. The results in this table reveal other interesting patterns. For instance, we find that the tenure profile is salient at public schools, but not at voucher-private schools, a result consistent with the prevalence of fixed-wage schemes in the public sector. More interestingly, we find that the size of the gender wage gap is large—almost 11%—at voucher-private schools, but close to zero at public schools, a finding in line with recent evidence by Biasi and Sarsons (2022) showing that flexible pay reforms can increase the gender wage gap.

The relationship between wages and self-selection is a core topic in labor economics. Indeed, the seminal observation by Roy (1951) that insofar as higher quality workers demand higher compensation, employers paying higher wages can attract those workers has become pervasive in the economics literature. However, this view underestimates the role of labor demand. Higher wages might not suffice nor be the only relevant variable because workers' matching in the labor market also depends on: i) their idiosyncratic taste, i.e., workers might have specific preferences for the public or private sector (Dal Bó et al., 2013; Deserranno, 2019; Ashraf et al., 2020), and ii) the labor demand that they face, i.e., the personnel selection process of the employers constraints workers' choice *de facto*. Indeed, the intuition derived from models with two-sided selection (Abowd and Farber, 1982; Logan, 1996) is that schools could offset the "labor supply effect" by making informed choices; in other words, selection can accentuate or counteract the self-sorting of workers à la Roy. For the interested reader, in the next subsection we present a thorough exposition of a two-sided matching model for the labor market. We build on Logan (1996)'s model, which is itself a variant of the deterministic two-sided matching models studied in game theory, and simulate the allocation of talent under different selection schemes.

C.2 Two-sided Selection Model

This section builds on Logan (1996) to simultaneously investigate schools' preferences to offer a job and workers' choice given the job offers. The model is based on an underlying random matching model of the labor market, which itself is a stochastic variant of deterministic two-sided matching models studied in game theory (e.g., Roth and Sotomayor, 1990).[43] The timing of the model is the following:

- Workers apply to all available schools.
- Schools evaluate applicants and make offers according to a decision rule.
- Workers evaluate the received offers and choose the highest-utility alternative.

The School's Decision

Similar to Abowd and Farber (1982), an underlying random utility model is defined to describe the decision of schools regarding whether or not to make jobs available to particular workers. For school j, the utility of hiring worker i of ability θ_i is defined as:

$$U_j(i) = m_j + \beta_j \theta_i + \epsilon_{1ij}, \tag{5}$$

while j's utility of not hiring worker i is:

$$U_j(\neg i) = s_j + \epsilon_{0ij}, \tag{6}$$

where m_j represents market effects on the utility of hires in general (e.g., reflecting the need for filling the position), β_j is the increase in utility that the school would experience from hiring a worker of marginally higher quality, and s_j is simple a baseline utility that school j derives from its present state of staffing. Finally, ϵ_{1ij} and ϵ_{0ij} represent factors that are not known to the observer but that influence the utility of school j of hiring or not hiring worker i.

[43] This game is a random variant of the "college admissions" game of the formal game theory literature, and because the deterministic results are transferable to the random matching game, it is known that at least one stable matching of employers and workers exists such that no worker-employer pair who are not matched to each other can improve their utilities by abandoning any current pair and establishing a new match together.

When expression (5) is greater in value than expression (6), employer j makes a job available: $o_{ij} = 1$, zero otherwise. Thus, the exact probability that school j will make an offer depends on the distribution of the differences between the two error terms, as well as on the non-stochastic parts of (5) and (6). If ϵ_{1ij} and ϵ_{0ij} are *iid* type I extreme value, then the difference will follow a logistic distribution, and the probability that j will make an offer is given by:

$$Pr(o_{ij}) = \frac{\exp(\beta_{0j} + \beta_j \theta_i)}{1 + \exp(\beta_{0j} + \beta_j \theta_i)}, \tag{7}$$

where $\beta_{0j} = m_j - sj$, and the offer of unemployment is always available to the workers, i.e., $Pr(o_{i0}) = 1$.

The Worker's Decision

Assuming that employers act independently of one another, conditional on workers' quality θ_i, then each applicant would be presented some set O_k of offers from the employers as a whole. There will be $R = 2^J$ distinct possible offering sets when J employers make separate decisions. Given this, the probability that worker i obtain a given offering set O_k is given by:

$$Pr(S_{ik}) = \prod_{m \in O_k} Pr(O_{im} = 1) \prod_{n \in \bar{O}_k} Pr(O_{in} = 0), \tag{8}$$

where m is an element (offer) of set K and n is an element of the complement set of O_k. A worker will choose her most preferred offer from the offering set that she faces. This is specified as a second random utility model. The indirect utility that worker i obtains from the job offered by employer j is defined as:

$$V_{i(j)} = h_j + w_j \theta_i + v_{ij}, \tag{9}$$

where h_j represents a baseline level of payments and amenities, w_j is a pay-for-performance component offered by the employer, and v_{ij} represents idiosyncratic preferences of the worker for a given job. Workers evaluate simultaneously every job offer that they find available to choose the one that delivers the highest utility. If v_{ij} follows a type I extreme value distribution, then the probability that worker i selects job j given the set of offers O_k is given by this polytomous conditional logit:

$$Pr(A_{ij} \mid O_k) = \begin{cases} \frac{\exp(h_j + w_j\theta_i)}{\sum_{h\in O_k}\exp(h_h+w_h\theta_i)} & , \quad j \in O_k \\ 0 & , \quad j \notin O_k. \end{cases} \tag{10}$$

Given our assumptions about the distribution of the random components in (5), (6), and (9), and further assuming that these random components are mutually independent, the probability that worker i ends-up in job j is given by:

$$\begin{aligned} Pr(A_{ij}) &= \sum_{k=1}^{R} Pr(A_{ij} \mid S_{ik}) \times Pr(S_{ik}) \\ &= \sum_{k=1}^{R} Pr(A_{ij} \mid S_{ik}) \times \prod_{m\in O_k} Pr(O_{im}=1) \times \prod_{n\in \bar{O}_k} Pr(O_{in}=0) \\ &= \sum_{k:j\in O_k} \frac{\exp(h_j + w_j\theta_i)}{\sum_{h\in O_k}\exp(h_h+w_h\theta_i)} \times \prod_{m\in O_k} \frac{\exp(\beta_{0m}+\beta_m\theta_i)}{1+\exp(\beta_{0m}+\beta_m\theta_i)} \\ &\quad \times \prod_{n\in \bar{O}_k} \frac{1}{1+\exp(\beta_{0n}+\beta_n\theta_i)}. \end{aligned}$$

Importantly, from this model we can obtain the expected quality of the workforce in a given school, which depends on the choices of both sides of the labor market. The expected quality of the workforce in school j is given by:

$$\mathrm{E}[\theta_i \mid \text{school}=j] = \int_\theta \theta_i f_{\theta\mid\text{school}=j}(\theta_i \mid \text{school}=j)d\theta.$$

Numerical Simulation

We are interested in the allocation of worker quality in the public and private sectors. More specifically, we seek to understand how the allocation of principal effectiveness in a given sector depends on the *selection* parameter β_j and the *pay-for-performance* parameter w_j of the model. For this purpose, we will consider a particular case of the model with only two schools, one private and one public. In this setting, there are only four possible offering configurations from public and private schools $(p, v) \in \{(0,0),(0,1),(1,0),(1,1)\}$. Thus, the probability that worker i is at a public school given her quality is given by:

$$
\begin{aligned}
Pr(A_{ip} \mid \theta_i) = & \left(\frac{\exp(h_p + w_p\theta_i)}{\exp(h_p + w_p\theta_i) + \exp(h_v + w_v\theta_i)} \right. \\
& \left. \times \frac{\exp(\beta_{0p} + \beta_p\theta_i)}{1 + \exp(\beta_{0p} + \beta_p\theta_i)} \times \frac{\exp(\beta_{0v} + \beta_v\theta_i)}{1 + \exp(\beta_{0v} + \beta_v\theta_i)} \right) \\
& + \left(1 \times \frac{\exp(\beta_{0p} + \beta_p\theta_i)}{1 + \exp(\beta_{0p} + \beta_p\theta_i)} \times \frac{1}{1 + \exp(\beta_{0v} + \beta_v\theta_i)} \right). \qquad (11)
\end{aligned}
$$

In this case, the expected principal effectiveness in the public school is given by:

$$
\mathrm{E}[\theta_i \mid \text{Public}] = \int_\theta \theta_i f_{\theta|\text{Public}}(\theta_i \mid \text{Public}) d\theta \qquad (12)
$$

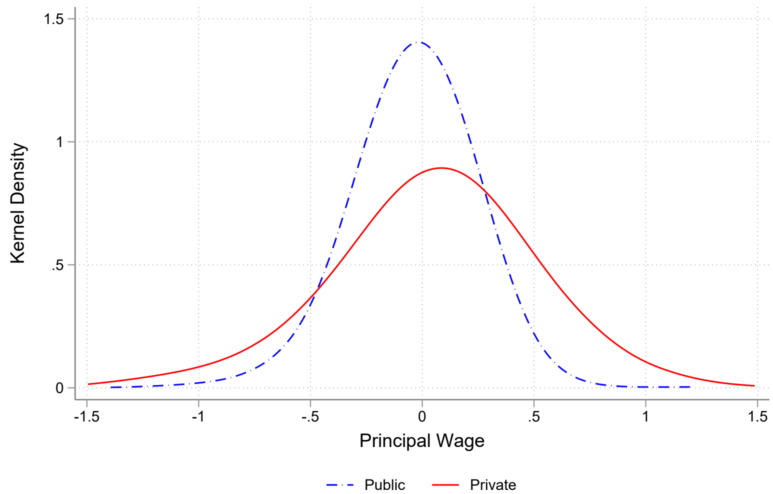
From Bayes' rule, we know that:

$$
f_{\theta|p}(\theta_i \mid \text{Public}) = \frac{Pr(A_{ip} \mid \theta_i) \times f_\theta(\theta_i)}{Pr(\text{Public})},
$$

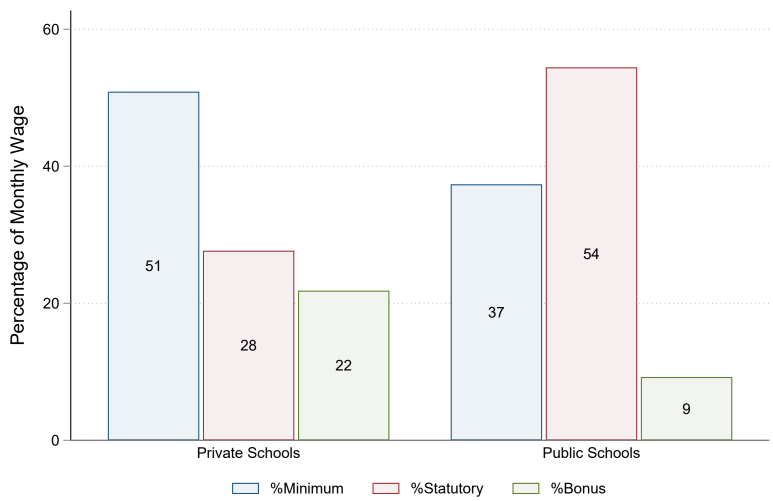
where $Pr(A_{ip} \mid \theta_i)$ is given by (11) and $Pr(\text{Public})$ is a scale factor equal to the fraction of public schools (0.5 in this case). Assuming that $f_\theta(\theta_i)$ is a standard normal, we can compute $\mathrm{E}[\theta_i \mid \text{Public}]$ using numerical integration. More importantly, we can study how this object depends on β_p and w_p, the two relevant parameters related to selection and payment policies in public schools, respectively.

Our simulation is presented in Figure C.2. Panel A, B, and C consider different personnel selection rules. Panel A shows a case where personnel selection is independent of worker quality. Panel B shows a case where a worker is selected if and only if her quality is above some threshold. Panel C shows the case where the likelihood of selecting a worker is increasing in proportion to her quality. Finally, Panel D shows the allocation of principal effectiveness given by equation (12). To construct this figure, we created a grid for β_p and w_p from 1 to 10, and compute $\mathrm{E}[\theta_p \mid \text{school type: Public}]$ for each cell of this grid.

Figure C.1: Principals' Wages



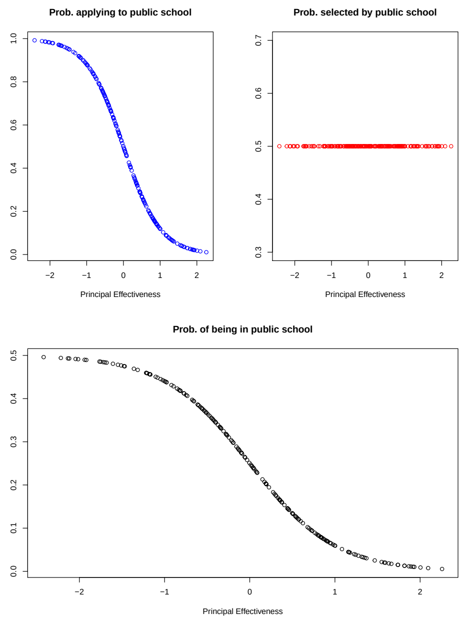
A. Residualized Log Wage



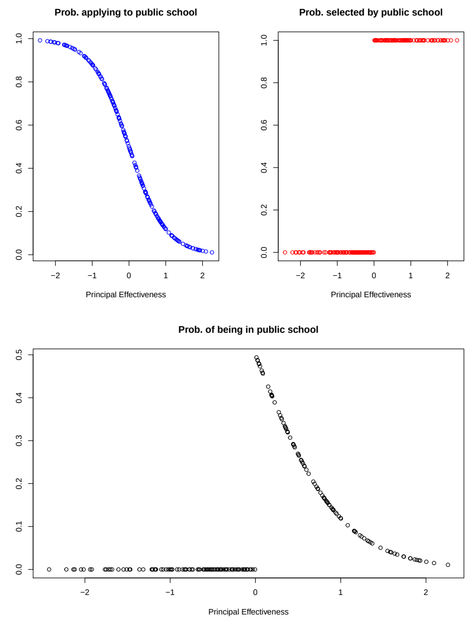
B. Wage Components

Notes: Panel A presents the distributions of log principals' wages in both public and subsidized private schools. Log principals' wages are residualized with respect to year and municipality fixed effects. Panel B decomposes the average monthly wage of school principals into the three components discussed in the data section: minimum legal wage, statutory payments, and bonuses. We present the share that each of these components represents of the principal' monthly wage, separately for subsidized private and public schools.

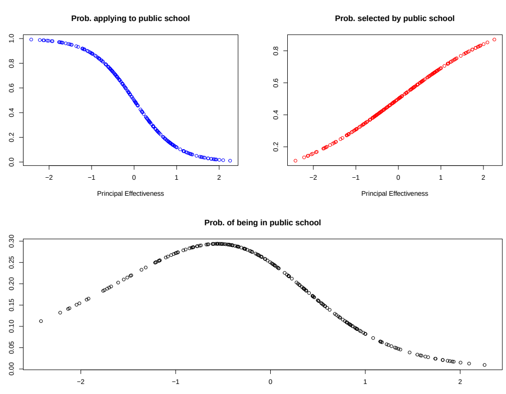
Figure C.2: Simulation of a two-sided matching model



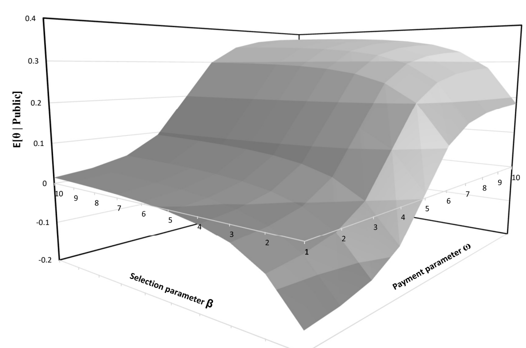
A. Selection at Random



B. Selection Above Cutoff



C. Selection on Quality



D. Allocation of θ_p if Selection on Quality

Notes: Panel A, B, and C show simple simulations that exemplify how personnel selection rules can counteract the self-selection component of labor markets. For this, we assume that the idiosyncratic preferences of principals follow a type I extreme value distribution, that principals do not anticipate the schools' selection rule, and that private schools have a larger pay-for-performance component ω than public schools. Panel D shows the allocation of principal effectiveness as a function of the selection and payment parameters. To construct this figure, we created a grid for β_p and w_p from 1 to 10, and computed E[θ_p | school type: Public] for each cell of this grid.

Table C.1: Principal Compensation and Principal Effectiveness

	ln(Wage)		ln(Base)		$\ln\left(\frac{\text{Wage}}{\text{Base}}\right)$	
	(1)	(2)	(3)	(4)	(5)	(6)
Private	0.105	0.139	-0.171	-0.129	0.276	0.268
	(0.008)	(0.009)	(0.010)	(0.012)	(0.008)	(0.013)
Principal Effectiveness	-0.016	-0.020	0.004	-0.003	-0.020	-0.018
	(0.015)	(0.015)	(0.021)	(0.019)	(0.015)	(0.015)
Principal Effectiveness × Private	0.087	0.075	-0.003	-0.008	0.091	0.083
	(0.033)	(0.030)	(0.039)	(0.036)	(0.030)	(0.029)
Female		-0.003		0.010		-0.014
		(0.007)		(0.011)		(0.007)
Female × Private		-0.114		-0.078		-0.036
		(0.016)		(0.020)		(0.015)
Age		0.032		0.043		-0.011
		(0.005)		(0.007)		(0.004)
Age × Private		0.005		-0.020		0.025
		(0.007)		(0.010)		(0.006)
Age^2		-0.000		-0.000		0.000
		(0.000)		(0.000)		(0.000)
Age^2 × Private		-0.000		0.000		-0.000
		(0.000)		(0.000)		(0.000)
Perm. Contract		0.067		0.078		-0.010
		(0.015)		(0.016)		(0.011)
Perm. Contract × Private		0.136		0.063		0.074
		(0.033)		(0.033)		(0.021)
Hours Contract		0.005		0.012		-0.008
		(0.005)		(0.005)		(0.008)
Hours Contract × Private		0.026		0.017		0.009
		(0.006)		(0.006)		(0.009)
College Degree		0.032		0.040		-0.008
		(0.016)		(0.022)		(0.014)
College Degree × Private		-0.019		-0.003		-0.016
		(0.026)		(0.034)		(0.024)
Observations	9,898	9,898	9,898	9,898	9,898	9,898
R-squared	0.181	0.303	0.163	0.241	0.227	0.234

Notes: This table presents the estimates from specification (4). We focus on a sample of principals for whom we have a standardized measure of effectiveness and detailed wage data from 2015 to 2017. Wage data is only available for public and subsidized private (voucher) schools. All specifications include year and municipality fixed effects. Bootstrapped standard errors (100 replications) clustered at the principal level are in parentheses.

D Data Files

This project combines students' performance and employer-employee records, provided by the Ministry of Education, with labor market outcomes coming from the Education Superintendency and the Civil Service. The authors did not have access to personal identifiers because the data files were anonymized by the Ministry of Education using a unique number. This appendix describes each data file used in the analysis.

Student performance: The Ministry of Education provided access to the performance records of all students between 2011 and 2016. For each student, we observe classroom and subject identifiers, as well as an identifier of the teacher by subject and classroom. For all students, we observe course grades by subject. For cohorts of students that take standardized exams, it is also possible to link our data to their test scores in the SIMCE exam. The SIMCE examination is only taken by students in some specific grades, usually 4th, 8th, and 10th grade, and it has not been systematically run every year in the country. Our main specification considers leads and lags of course grades. Thus, we only use four years of data (2012-2015). We exclude students for whom the teacher does not change in a given subject from one year to another, and we also exclude classes that had more than one teacher per year as well as the bottom and top one percent of classroom size outliers. We complement these data with records from the centralized admission system. Specifically, we add the average (at the school level) of the students' composite score used for college admission and the average (at the school level) of the students' scores in the college entrance exams of Math and Spanish.

Panel of school workers: The Ministry of Education provided access to a panel of school workers between 2008 and 2017. These records include 13,693 unique schools and 331,167 unique workers. For each worker, we observe the following characteristics: gender, age, tenure in the system, certification, type of contract, hours of contract, and her occupation within the school. Based on the latter, we identify the principal in each school by year. In cases with more than one principal in a given year, we choose the one with more hours of contract in the school (if there is a tie, then we choose the most senior worker).

We complement this data with records from the teachers' evaluation system. The Chilean evaluation system operates on the basis of four sources of evidence: a portfolio, an interview by a peer teacher with at least five years of experience, a written report of two school authorities on the basis of a set framework, and a self-evaluation report by the teacher following a given structure. Among the instruments the portfolio has the highest weighting in the process of establishing the competence level of the teacher being evaluated (60%), followed by the peer interview with 20% and the other two sources of evidence with 10% each. Based on this information, teachers are classified in four performance categories: "outstanding", "competent", "basic", or "unsatisfactory". For more details, see Avalos-Bevan (2018).

School characteristics: The Ministry of Education provided access to a panel of 13,693 schools between 2008 and 2017. These records include the following information for each school: type of administration (e.g., public, subsidized-private or private), an indicator if the school is in a rural area, its total enrollment, concentration of disadvantaged students, and the municipality where the school is located. Using the national representative survey CASEN, we add characteristics of the municipality where the school is located. Specifically, we add the following characteristics: average years of education, income per-capita, and the 2011 rates of crime, unemployment, and poverty. Moreover, from SIMCE surveys, we were able to recover the shares of low-income and high-income parents and the share of parents with a college degree.

For the analysis, we remove private schools that do not receive vouchers because we do not observe wages for those. Preschools, adults' schools, and special education schools are also excluded. All and all, we end-up with 11,320 schools.

Wages: The Superintendency of Education provided access to a monthly panel of workers from 2015 to 2017. These records correspond to reports that every school receiving vouchers must provide to the Superintendency in order to report the use of public resources. For each worker, we observe the school where she is working and detailed data on wages. Specifically, we observe worker's compensation by item. We classify the raw wage as the sum of these items, and we also classify these items into three categories:

- Minimum wage: corresponds to a per-hour legal-minimum payment for teachers, defined by the Ministry of Education.
- Statutory payments: include compensations regulated by law but unrelated to performance, such as payments for experience and for teacher certification. We include all payments defined by the Union Law of 1996 as well as other payments defined by subsequent Laws, such as: Mejoramiento, Condiciones Dificiles, Profesor Encargado, Excelencia Pedagogica, UMP, Titulo y Mencion, Planilla Complementaria), and other compensations assigned to those who work extra hours, in rural schools, or in schools where it is "difficult" to teach according to the Ministry of Education.

- Bonuses: encompasses compensations related to workers' performance, such as individual and collective performance bonuses (e.g., AVDI), payments from the national system of performance assessment (e.g., AEP, SNED), bonuses paid directly by the school owner in the case of private schools, and other discretionary payments and gratifications related to transportation, food, and holidays.

School Finance: The Superintendency of Education provided access to school finance for the year 2016. These records correspond to reports that every school receiving vouchers must provide to the Superintendency in order to report the use of public resources.

Teacher surveys: The Ministry of Education provided access to the survey responses of teachers. Every time students take the nationwide standardized exam SIMCE, teachers must fill out a survey created by the Ministry. For our analysis, we only consider questions about the school principal (e.g., The principal does a good job, the principal promotes a good work climate). According to the availability of the questions in each year, we took the surveys from 2009 to 2015 for teachers from 4th, 8th, and 10th grade.

In the SIMCE survey, every teacher must provide an answer within a range from 1 to 4 (or from 1 to 5 in some years), where 1 represents a high disagreement with the statement and 4 (or 5) represents a high level of agreement with it. We use their responses to create a dummy variable at the survey respondent level that equals one if the teacher "highly agrees" with the statement about the principal, i.e., her response is at the top of the specific scale for that question. Then, we take the average across respondents at the school-year level and assign this to the corresponding school principal.

Civil service: The Civil Service provided access to records of the contest implemented to elect principals in public schools from 2011 to 2016. While these contests are direct responsibility of the municipalities, the Civil Service oversees them and records data on them. For every school, we observe a panel of contests. Specifically, we observe when a contest was called and what was the outcome of the contest (whether the position was filled or not). Based on this, we create an identifier at the school-year level indicating if the school chose a principal through the new system each year.

Complaints against the schools: The Superintendency of Education provided access to all complaints filed against the school between 2014 and September 2018. These records have the number of complaints by category. The categories include: i) bullying and discrimination (also includes behaviors of sexual connotation against students or teachers), ii) denied enrollment (for instance, because of disciplinary measures), iii) poor infrastructure (includes lack of furniture), iv) teacher absenteeism (or lack of teachers), v) school accidents, vi) charge of extra fees (or ask for extra materials), vii) resource accountability (irregularities in the use of vouchers or misreporting of attendance).

Complaints are often filed by parents. While teachers could also file complaints through the Superintendency, most of the time, their complaints go directly to the Labor Directorate or justice system directly.

References

Abowd, J. M. and H. S. Farber (1982): "Job queues and the union status of workers," *ILR Review*, 35, 354–367.

Angrist, J. D., P. D. Hull, P. A. Pathak, and C. R. Walters (2017): "Leveraging lotteries for school value-added: Testing and estimation," *The Quarterly Journal of Economics*, 132, 871–919.

Ashraf, N., O. Bandiera, E. Davenport, and S. S. Lee (2020): "Losing prosociality in the quest for talent? Sorting, selection, and productivity in the delivery of public services," *American Economic Review*, 110, 1355–94.

Avalos-Bevan, B. (2018): "Teacher evaluation in Chile: highlights and complexities in 13 years of experience," *Teachers and Teaching*, 24, 297–311.

Biasi, B. (2021): "The labor market for teachers under different pay schemes," *American Economic Journal: Economic Policy*, 13, 63–102.

Biasi, B. and H. Sarsons (2022): "Flexible wages, bargaining, and the gender gap," *The Quarterly Journal of Economics*, 137, 215–266.

Boyd, D., H. Lankford, S. Loeb, M. Ronfeldt, and J. Wyckoff (2011): "The effect of school neighborhoods on teachers' career decisions," *Whither opportunity*, 377–396.

Callaway, B. and P. H. Sant'Anna (2020): "Difference-in-differences with multiple time periods," *Journal of Econometrics*.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018): "Firms and labor market inequality: Evidence and some theory," *Journal of Labor Economics*, 36, S13–S70.

Card, D., J. Heining, and P. Kline (2013): "Workplace heterogeneity and the rise of West German wage inequality," *The Quarterly Journal of Economics*, 128, 967–1015.

Chetty, R., J. Friedman, and J. Rockoff (2014): "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates," *American Economic Review*, 104, 2593–2632.

Cowen, J. M. and M. A. Winters (2013): "Do charters retain teachers differently? Evidence from elementary schools in Florida," *Education Finance and Policy*, 8, 14–42.

Cuesta, J. I., F. González, and C. L. Philippi (2020): "Distorted quality signals in school markets," *Journal of Development Economics*, 147, 102532.

Cullen, J. B., E. A. Hanushek, G. Phelan, and S. G. Rivkin (2016): "Performance information and personnel decisions in the public sector: The case of school principals," Tech. rep., National Bureau of Economic Research.

Dal Bó, E., F. Finan, and M. A. Rossi (2013): "Strengthening state capabilities: The role of financial incentives in the call to public service," *The Quarterly Journal of Economics*, 128, 1169–1218.

De Chaisemartin, C. and X. d'Haultfoeuille (2020): "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review*, 110, 2964–96.

Deserranno, E. (2019): "Financial incentives as signals: Experimental evidence from the recruitment of village promoters in Uganda," *American Economic Journal: Applied Economics*, 11, 277–317.

Finan, F., B. A. Olken, and R. Pande (2017): "The personnel economics of the developing state," *Handbook of Economic Field Experiments*, 2, 467–514.

Gibbons, R., L. F. Katz, T. Lemieux, and D. Parent (2005): "Comparative advantage, learning, and sectoral wage determination," *Journal of Labor Economics*, 23, 681–724.

Goodman-Bacon, A. (2021): "Difference-in-differences with variation in treatment timing," *Journal of Econometrics*.

Holland, P. W. (1986): "Statistics and causal inference," *Journal of the American Statistical Association*, 81, 945–960.

Holmstrom, B. and P. Milgrom (1987): "Aggregation and linearity in the provision of intertemporal incentives," *Econometrica*, 303–328.

Jackson, C. K., S. C. Porter, J. Q. Easton, and S. Kiguel (2022): "Who Benefits From Attending Effective High Schools?" .

Kremer, M. and E. Maskin (1996): "Wage inequality and segregation by skill," Tech. rep., NBER.

Logan, J. A. (1996): "Opportunity and choice in socially structured labor markets," *American Journal of Sociology*, 102, 114–160.

Neilson, C. (2019): "Targeted vouchers, competition among schools and the academic achievement of poor students," *Working Paper*.

Rambachan, A. and J. Roth (2021): "An honest approach to parallel trends," .

Romano, J. P. and M. Wolf (2005): "Stepwise multiple testing as formalized data snooping," *Econometrica*, 73, 1237–1282.

Roth, A. E. and M. Sotomayor (1990): "A study in game-theoretic modeling and analysis," *Econometric Society Monographs*, 18.

Roth, J. (2022): "Pretest with caution: Event-study estimates after testing for parallel trends," *American Economic Review: Insights*, 4, 305–22.

Rothstein, J. (2010): "Teacher quality in educational production: Tracking, decay, and student achievement," *The Quarterly Journal of Economics*, 125, 175–214.

——— (2015): "Teacher quality policy when supply matters," *American Economic Review*, 105, 100–130.

Roy, A. D. (1951): "Some thoughts on the distribution of earnings," *Oxford Economic Papers*, 3, 135–146.

Staiger, D. O. and J. E. Rockoff (2010): "Searching for effective teachers with imperfect information," *Journal of Economic Perspectives*, 24, 97–118.