# Rules Rather Than Discretion: Teacher Hiring and Rent Extraction

Ricardo Estrada

Paris School of Economics

- Preliminary version -

May 2014 *

**Abstract**

There is mounting research evidence on both the importance of teacher quality in the production of learning and on the difficulty of identifying who actually is (can be) a good teacher. In this study, I use a recent policy reform in Mexico to evaluate the effect on student outcomes of receiving a brand-new teacher hired through a standardized test versus one hired in a discretionary process with strong involvement from the teachers' union. My difference-in-differences results indicate that the allocation of test-hired teachers increases student achievement in junior-secondary schools. I also find that joint committees of state officials and union representatives allocate the discretionary-hired teachers to schools in more "desirable" localities - but with similiar pre-treatment trends on outcomes. Taken together, these results suggest the existence of an agency problem with potential rent-extraction that the use of a hiring rule can mitigate.

JEL: D82, I10, J23.

Keywords: teacher quality, teacher hiring, teachers' unions, student achievement, rent extraction.

*48 boulevard Jourdan, 75014 Paris, restrada@pse.ens.fr. I thank Marc Gurgand and Jérémie Gignoux for supervision and encouragement. I am also grateful for insightful comments to François Bourguignon, Taryn Dekelman, Alejandro del Valle, Pascaline Dupas, Francisco Ferreira, Paul Glewwe, Kenneth Houngbedji, Fabian Lange, Amine Ouazad, Jorge Santibanez and participants at seminars at the Paris School of Economics, the conferences NEUDC Harvard, LACEA Colegio de Mexico and MIEDC Minnesota, and workshops at Lisbon ISEG, LACEA Impact Evaluation Network GRADE, ZEW Mannheim University and Oxford Blavatnik School of Government. I thank the staff at Mexicanos Primero for kindly providing me with access to data.

# 1 Introduction

There is mounting research evidence on both the importance of teacher quality in the educational production function and on the difficulty of identifying at the time of hire who actually is (can be) a good teacher. The problem for school administrators (and parents) is that the characteristics they can typically observe at the time of hire, like experience and training certificates, are poor predictors of teacher quality – or irrelevant for recent graduates (see, for example, *Hanushek and Rivkin* (2006)).

The participation of current teachers in the selection of new teachers may ease this informational problem if the former have either superior information or higher ability to identify teacher quality using broader criteria.[1] However, asymmetric information between administrators and current teachers may lead to an agency problem if current teachers' optimize an objective function with different arguments from teacher quality. In this case, administrators would face a trade-off between hiring using rules that are second-best predictors of quality, but hard to manipulate, or giving discretionary powers to an agent endowed with better information, but with a distinct objective function.[2]

In this paper, I use a recent policy reform in Mexico to evaluate the effect on student outcomes of receiving a brand new teacher hired through a standardized test versus one hired in a discretionary process with strong participation from the teachers' union. The recruitment of teachers for public primary and junior-secondary schools in Mexico is centralized at the state level and is not directly associated to the filling-in of specific vacancies in schools. Prior to the reform, state officials would select almost all brand new teachers through a discretionary process in which the teachers' union would perform the role of a hiring agent. With the reform, each state now selectes a share of brand new teachers on the basis of a standardized test given for this purpose. Hence, every year states hire teachers using both the test and the discretionary methods. Importantly, both selection mechanisms are not tied, in principle, to specific school vacancies.

The reform did not change the mechanism to allocate teachers to schools. State-wide committees jointly chaired by state officials and representatives from the teachers' union are in charge at the beginning of every academic year to fill in from the stock of old and newly hired teachers the vacancies opened in schools by retirement, between-school transfers and the expansion of teaching positions. Joint committees allocate current teachers to available schools based on teachers' applications for specific school positions and a set of pre-defined criteria. In other words, the allocation of teachers to schools depends hence on teachers' preferences and committees' assessments of applicants' relative merit. Anecdotical evidence

[1] For example, Rockoff, Jacob, Kane and Staiger (2011) find that combining a broad set of measures of teacher characteristics is informative about the effectiveness of new math teachers in New York City, while the same measures have little predictive power as independent factors.

[2] *Hoxby* (1996)discusses a theoretical model in which teachers demand unions to influence the educational production function either because: 1) though they have the same objective function than administrators and parents (maximization of student achievement), teachers have better information about input efficiency or internalize externalities than the others neglect; alternatively 2) teachers may have a different objective function than administrators and parents, and hence a desire to set the school inputs that maximize their own objectives.

suggests that teachers' preferences over schools depend on the characteristics of the localities where schools are seated. Brand new teachers are allocated to available schools after current teachers' choices. A priori, whether test-hired and discretionary-hired teachers are allocated to schools with different characteristics is an empirical question that I investigate.

In this paper, I focus my analysis in Telesecundaria schools, a system of public junior secondary education (Grades 7 to 9). Telesecundarias are small schools catered to small communities and account for 20.6% of student enrollment in junior secondary education in Mexico.

I find in a cross-sectional OLS regression that brand new discretionary teachers are more likely to be assigned to schools located in "more desirable" localities – less poor and with larger penetration of public services –; and that past outcomes of schools do not predict assignment into treatment – conditional on locality characteristics. This result is consistent with a model in which committees allocate test and discretionary teachers based on teachers' preferences for locality characteristics and not on past school performance and in which committees give a higher weight to the preferences of the discretionary teachers.

I investigate the allocation of test and discretionary teachers to junior secondary schools using panel data with five yearly observations per school before treatment.[3] Visual inspection of the raw data – see Figure 1 – is consistent with the claim that allocation of teachers to schools is ortogonal to the evolution of past school outcomes. Schools that receive new test teachers have on average lower performing students than schools that receive discretionary teachers – as locality characteristics are correlated with student achievement. However, schools where test and discretionary teachers are assigned seem to follow similar pre-treatment trends in outcomes – and converge after treatment. I confirm the former pattern by running a school fixed-effects model in which I regress (school-level) student outcomes in a standardized exam on a vector of year dummies interacted with (eventual) treatment status and a set of time-variant school inputs and time-variant state effects. I believe this evidence firmly supports the plausibility of the "parallel trend" assumption necessary for the identification of a causal treatment effect in a difference-in-differences model.

My difference-in-differences estimates indicate that students benefit from having test teachers. I find that the allocation of a test teacher has a positive and sizable effect on student achievement. Moving from no test teachers in a school to only test teachers is associated with an increase in .45 standard deviations in the school's Mathematics test score and .37 standard deviations for the Spanish score, a result that is statistically significant at the ten-percent level for Mathematics and at the five-percent level for Spanish. This large effect indicates an important gap in the quality of teachers hired through both methods. Results are robust to different specifications and checks.

Summing up, my findings indicate that education officials hire on average teachers of less quality when they follow a discretionary process with participation from the teachers' union. Then, joint committees

[3] I restrict my analysis to a type of rural schools called Telesecundarias for data limitations.

of state officials and union representatives allocate these teachers to schools in more "desirable" localities. Taken together, these results suggest the existence of an agency problem with potential rent-extraction. State officials are able to reduce the extent of this problem by using a hard-to-manipulate rule to hire new teachers – a ranking based on a teacher test. These findings are particularly relevant for environments with weak institutions, where the lack of accountability (on actual teacher performance, for example) may exacerbate the incentives for rent-seeking behavior arising because of imperfect information about worker productivity.

# 2 A Reform to Teacher Hiring in Mexico

State governments (31) operate the public primary (grades 1 to 6) and junior-secondary (grades 7 to 9) schools in Mexico, while the Federal government sets the national curricula and provides states with the bulk of funding.[4] Teacher hiring, allocation to schools and promotions are centralized at the state level. State Ministries of Education are in charge of teacher hiring, which is not related to fill specific school positions. Joint committees of state officials and teachers' union representatives are responsible for the allocation of teachers to schools.

Starting in 2008, the Federal government championed a reform to teacher recruitment for public schools in the country, introducing standardized testing as a mechanism to hire teachers.

## 2.1 Discretionary Hiring

State Ministries of Education are responsible for hiring the brand new teachers required to fill the vacant positions in the school system. For brand new teachers, I mean those entering into a teaching position in the public education system. This definition excludes incumbent teachers that are transferred from one public school to another. For brevity, I will use as exchangeable the terms brand new teacher and new teacher from here and on.

Hiring is done at the teacher-type level (e.g. Primary school teacher, Mathematics junior high school teacher, etc.), but in principle is not related to specific vacancies at schools. The allocation of teachers to schools is defined later by a joint committee of state officials and teachers' union representatives in a process that I describe below.

Prior to 2008, there was no Federal regulation dictating how the process to hire new teachers should be conducted. Hirees must have university-level studies, though – even if not necessarily a degree at the moment of hiring. Teaching in junior high school education is not restricted to graduates from teacher schools and there is no teaching certificate required to join the profession at this educational level.

Beyond the schooling requirement, state governments enjoy wide discretion to set the selection method

[4] The Federal government is also in charge of managing public schools in the Federal District, where the capital of the country is located.

to hire – most of – those who join the public education system. From an aggregated point of view, vacancies equal the expansion of the stock of state teachers (new payroll positions) plus outflows from this stock – to retirement or other occupations, for example – (existing payroll positions). The bulk of state vacancies is generated by outflows from the stock of teachers. The State Ministries of Education rely heavily on the teachers' union to fill these positions. Basically, Ministries of Education hire candidates proposed by the union, which functions as an agent in charge of selecting the new hirees (*Iaies et al.*, 2006).

The teachers' union (SNTE by its Spanish acronym) is a national organization with 52 regional sections. Both affiliation and payment of fees to the union is mandatory – and automatic – for all teachers in public elementary schools. In other words, every teacher in a public elementary school is also a member of the teachers' union.

The role of the union as a hiring agent is informal and there is limited formal information about how the union runs the selection process. It is common though that retiring teachers propose to the union one would-be teacher with priority for hiring. This is probably the main mechanism used in practice to select new teachers.

Under a broader law that regulates the labor relations of public employees, the teachers' union would be formally entitled to select for hiring a number of positions equivalent to the 50% of the expansion in the stock of teachers in a year. State Ministries of Education would be entitled to select the other 50% of the new payroll positions, which would be typically allocated to new graduates from the public teacher schools in the state. Both quotas disappeared though with the 2008 reform.

State officials and the teachers' union have been subject to criticism for neglecting teacher quality when hiring. Strong, but not isolated, denounces include the selling of teaching positions and the practice of teachers going into retirement to bequeath their position to a relative. A national survey among elementary teachers found that one-third of interviewees thought that selling of teaching positions was a frequent practice (*Santibanez*, 2008). When the 2008 reform was announced, union leaders in at least two states publicly declared their opposition to the examination because it would take out the union members' right to bequest their position (*Elizondo*, 2011).

## 2.2 Test-based Hiring

In 2008, the Federal government introduced (under the umbrella of a broader agreement with the teachers' union) a plan to open to competitive examination all vacant teaching positions in public primary and junior secondary education in the country. The mechanics of the new examination is the following.

Competition is open to candidates willing to enter the teaching profession in public schools and current teachers with temporary contracts. There are hiring quotas for each group. In this paper, I focus only in the recruitment of brand-new teachers. Hiring is based on a national-standardized test held before the

beginning of the academic year. There is one exam for each type of teaching position (e.g. Primary school teacher, Mathematics junior high school teacher, etc.). The standardized exam is designed to measure cognitive skills, knowledge of the teaching subject and mastery of teaching methods.

Candidates are ranked by state and teacher type according to their exam results or, if states opt for it, a weighted average of the test score and other criteria (often undergraduate GPA). The number and type of available teaching positions by state and the exam results are widely publicized by media outlets and are available on a dedicated web page (concursonacionalalianza.sep.gob.mx). Civil-society organizations participate as monitors in different stages of the process, more visibly in the exam application. The teaching positions open to competition are not associated to specific schools. Some type of teaching positions are restricted to graduates of teacher schools or from specific college majors.

The reform met with strong opposition from state officials and local union leaders. In a compromising result, only new payroll positions funded by the Federal Government were filled through the test-based recruitment initially, though it was expected that progressively more vacancies were opened to test-based hiring.[5] Almost all states use the test-based recruitment to fill some of their vacancies since 2008 (30 of 32, including the Federal District). According to figures from the Federal Ministry of Education, from the 22,546 full-time vacancies opened to test hiring in the 2010 school year, 34% corresponded to new positions and the rest to existing payroll positions. There is not public information about the total number of new teachers hired through discretionary recruitment. They could amount to around 82% of all teachers hired according to my estimates.[6]

Hence, every year states select new teachers through both test-based and discretionary recruitment. The reform did not change the mechanism to allocate teachers to schools.

## 2.3 Telesecundaria Schools

As I explain below, I only observe the link between teachers and students at the school level. Hence, I focus my empirical analysis in Telesecundaria schools, a system of public junior secondary education (Grades 7 to 9). Telesecundarias are small schools catered to small communities. The small school size should increase the likelihood that I find a statistical significant teacher effect at the school level.

The typical Telesecundaria school in my sample has 72 students divided in 4 classrooms and is located in a locality with 890 inhabitants – all are median values. This educational system is widespread, though. According to figures from the Federal Ministry of Education, around 1.26 million students attended 18,000 Telesecundaria schools in the school year 2010, which amounts to 20.6% of student enrollment in junior secondary education in the country.

Telesecundaria students tend to be more rural, poorer and face in general more disadvantaged conditions than the average junior high school student. For example, in 2010 the average poverty rate in the

[5] The Teachers' Union agreed to cede its selection entitlement over the 50% of the federally-funded new payroll positions.
[6] In the Telesecundaria System.

localities where the Telesecundarias in my sample are located was 65%, while the national poverty rate was 46% – according to the National Council for the Evaluation of Social Policy (CONEVAL).

Contrary to general junior high schools, Telesecundarias have one teacher per classroom - as opposed to one teacher per topic. Instead of specialist teachers, Telesecundarias rely heavily on IT teaching support. The television programs that the Federal Ministry of Education produces specifically for this school system fill approximately 2 of the 6 hours of the school day. Hence, it seems reasonable to assume that the effect of teacher quality in Telesecundarias is lower than in educational systems in which teachers play a larger role in the learning process.

# 3 Data

## 3.1 Enlace Exam

I use the results in a national standardized test (Enlace) that students take at the end of the academic year to construct a panel dataset of school scores from 2005, the first year that the exam was given, to 2010 – five years before and one year after the treatment of interest.[7]

Test scores for grades 7 and 8 are only available since the school year 2008, as only 9 graders would take the exam in the 2005-2007 period. The grade 9 exam assessed materials of grades 7 to 9 before 2007, while after this year is focus in grade 9 materials. The test measures learning in Mathematics, Spanish and a rotating subject every year. I use the first two subjects for the analysis. I only use the results from grade 9 in my main estimations because the larger panel dimension.

The Mexican Evaluation of Scholastic Achievement of Educational Institutions (Enlace) is designed to assess the overall educational system and, hence, there is no bearing for students on GPA or graduation. However, Enlace results are widely reported by media outlets and non-governmental organizations.

Also, since 2009, the Federal Ministry of Education delivers monetary bonuses to teachers of classrooms and in schools in the top 15% in the – respective – score distribution; and to teachers in schools in the top 15% in the score gains distribution (gains with respect to the previous two years). Schools are classified by state into categories defined by locality (urban/rural and with high/low marginalization) and school (general/technical/telesecundaria/etc.) characteristics. A teacher can receive a bonus ranging from $2,000 up to $20,000 pesos (around USD PPP 260 and 2,600, respectively).

The publicity and the bonuses provide school agents with incentives to perform better and makes Enlace a medium-stake test.

[7] For simplicity, I will refer to the school year 2005-2006 as 2005 and so on, though the Enlace results from the 2005 school year correspond to the test given in the second quarter of the 2006 calendar year.

### 3.1.1 Detection of Cheating

Students take the Enlace exam in two school days towards the end of the academic year. Each State Ministry of Education allocates one exam coordinator per school to overview along the school principal the implementation of the test. The school principal selects one teacher per classroom to monitor the students during the application of the exam. It is forbidden that teachers monitor the classrooms they teach.[8] At the end of each day, the monitoring teachers must turn in the response sheets to the exam coordinator and the school principal, who pack the answer sheets into sealed boxes at the end of the second and final day of the exam. Information sheets distributed to principals and teachers state the subsequent use of a computer software to detect copying among students and provision of exam responses by a third-party.

The Federal Ministry of Education runs a software to detect test cheating using two statistical tools commonly used for this purpose, the K-Index and the Error Similarity Analysis (ESA) Index.[9] Both methods measure unusual agreement between the incorrect answers of two examinees in a multiple-choice test and, as both are based on a binomial distribution, have a fairly similar general structure. The focus in common incorrect answers comes from the idea that the number of similar correct answers increases with students' true achievement level, while the identical selection of responses given as distractors is informative of copying.

The two indexes are designed to give lower-bound estimates for a specific form of cheating: direct copying, e.g. one student copying from another or a larger group of students (and potentially teachers) exchanging responses during the application of the exam. Even in this case, copying will go undetected if it is restricted to a few answers – relatively to the total number of wrong responses– or if the source of copying do not have incorrect responses. E.g. if some one gives one string of correct answers to the whole students in a classroom. Moreover, both methods are unlikely to be informative about other forms of cheating that may involve students (like use of cheat sheets and impersonation) or teachers (e.g. giving students extra-time or teaching to the test).

There are not sanctions to either principals, teachers or students associated to suspected cheating. The individual exams that are flagged as suspicious of cheating are not taken into account though for the estimation of the school score that is reported in the official results. The Federal Ministry of Education delivers to the State Ministries a report with the list of the schools in which a high prevalence of cheating is detected.

[8] School principals should guarantee that at least two parents per classroom attend the exam as external observers. I do not have information about how extensively this policy is implemented.

[9] The Educational Teaching Service (ETS) routinely uses the K-Index to detect cheating in the several examinations they perform (SAT, GRE, GMAT, etc.), while the ESA Index is the basis for the, commercially available, *Scrutiny!* software. A detailed description of both methods can be found, respectively, at *Holland* (1996) and Belleza and Belleza (REF).

### 3.2 School and Locality Characteristics

I use the census of schools carried out by the Ministry of Education (Formato 911) to obtain yearly information about school inputs (school and class size, student characteristics and teachers' credentials). Using the census locality code, I retrieve information from the 2010 population census about the characteristics of the localities where the schools are located and from the National Commission for the Evaluation of Social Policy about the localities' poverty rate. I obtain from Google Maps the estimated travel distance by car from the schools' localities to the State capital.[10]

### 3.3 Census of Teachers

I benefit from extensive data of school personnel compiled due to a recent mandate of the Mexican Federal Congress. The data comprises the quarterly payrolls of public elementary schools from the 2nd quarter of 2010 – the last of the academic year 2009-2010 – to the 2nd quarter of 2011 – and so, it covers the full academic year 2010-2011. The Federal Ministry of Education (SEP) assembled the dataset using information supplied by the State Education Ministries. I track teachers through schools and quarters using their taxpayer number and construct a quarterly panel of school personnel inclusive of name, tax payer and population identification numbers, birth date, assigned school(s) and occupation information. The dataset does not include though information about hiring, education profile or assigned classrooms.

I do not observe directly in the data who are the teachers hired since 2008, when the test-based examination was implemented, and how these teachers were recruited. However, I can use the 2009 and 2010 censuses to identify the 2010 cohort of new teachers and match these to the available results of the test-based hiring. Hence, I focus my analysis here and after in the (24) states that opened vacancies for the Telesecundaria system in the 2010 test-based hiring.[11]

I identify the 2010 cohort of new teachers by first comparing the 2010-2nd-quarter census of Telesecundaria's personnel (the last of the school year 2009-2010) to the census of all personnel registered in any of the four quarterly censuses of the 2010-2011 academic year. I assume that all the 2010-2011 observations that I do not find in the 2nd quarter of 2010 correspond to Telesecundaria System's brand-new personnel in the 2010 school year. I drop observations from 7 states that report relatively few personnel in the 2nd quarter of 2010 and hence have a high, an likely unreliable, ratio of new/total personnel (larger than 20%) in the 2010 school year – the mean ratio in the remaining states is 6.7%.[12]

The SEP dataset includes a module with the list of the 2009 and 2010 test-selected applicants (626 and 550 teachers, respectively, in the 24 states). I merge this module to the main dataset using the national

[10] Using the Stata command traveltime.

[11] 2 states (Michoacan and Oaxaca) do not participate at all in the test-based examination and 6 states did not open to test-based competition any vacancy at the Telesecundaria system in 2010 (Baja California Sur, Colima, Nayarit, Queretaro, Sonora and Zacatecas).

[12] These states are Baja California, Guanajuato, Guerrero, Tabasco, Tlaxcala, Veracruz and Yucatan.

population number.[13] I am able to merge 68% (427 observations) of the 2009 and 72% (395 observations) of the 2010 test teachers to specific schools. I also find 12 (.8%) 2010 test teachers in supervision offices. The dataset includes a State Education Ministries' report on the candidates' hiring status. 39% of the the non-matched 2010 test teachers are declared to be in the waiting list for allocation to a school, 18% did not fulfill all the administrative requirements to be hired, 10% did not accept the assigned school and there is no status information for the remaining 33%.[14] I do not have good information to evaluate if the probability that I observe the allocation to a school of a test-selected teacher depends on teacher quality. I will come back to its potential implications though when I discuss the allocation of teachers to schools.

I find that 10% of the matched individuals hired in the 2010 test examination as new teachers were already in a Telesecundaria's payroll in the 2009 school year. In the extreme, 15 of the 16 test teachers hired in the state of Nuevo Leon fall in this case. This evidence suggests that some incumbent teachers – maybe hired under temporary contracts – where allowed to participate in the examination for brand-new teachers. I drop out the observations from incumbent teachers hired as new teachers in the test examination as well as all the observations from the state of Nuevo Leon.

The database is inclusive of teachers, administrative staff and principals. I identify as teachers all observations where at least in one quarterly database I observe a synonymous of the word "Teacher" or "Hours Telesecundaria" in the two variables with information about the post description .[15]

Overall, I have a database with 1,869 new teachers (19% test-selected) distributed in 1,661 schools in 15 states in the 2010 school year. In addition, I have information on the 415 test teachers hired in 2009 for whom I record the schools where they were allocated in 2009 and 2010 . I collapse then the dataset at the school-year level and merge it to the panel with school results and characteristics.

I obtained from the Federal Ministry of Education a list with the schools where the 2008 test teachers were initially assigned – the file does not have the teachers' population or tax identification number. I add this information to the panel of schools.

After merging, I have a panel with 1,638 schools in 15 states. The size of the database reduces to 892 schools in 13 states when I restrict to schools which have never received a 2009 or 2008 test teacher, which did not receive both a new test teacher and a new regular teacher in 2010 and for which there is at least 4 years of Enlace results. I finally trim schools with at least one year-to-year change in their school score larger than 2 standard deviations – which roughly corresponds to the top and bottom one percentiles of score changes. 9.8% of schools in the sample received – at least – a test teacher in the school year 2010.[16]

[13] The taxpayer identification number is not available in the test hiring module. Around 5% of the observations in the main personnel module have missing information for the national population number.

[14] I drop out all observations from the Federal District as there is no match among 2010 test-hired teachers.

[15] The actual keywords that I use are: Maestro, Mtro, Profesor, Docente, Horas Telesecundaria and H.S.M. I also use three payroll codes which I know from the data that are associated to teaching positions.

[16] I exclude all observations from the states of Morelos and Campeche because there is no left schools with only test teachers after these restrictions.

# 4 Allocation of Teachers to Schools

In each state, a joint committee of state officials and union representatives is in charge of allocating teachers to schools. Joint committees operate under state-level regulations heavily based on a 1973 agreement between the Federal Ministry of Education (SEP) and the Teachers' Union (SNTE). The 1973 SEP-SNTE agreement stipulates that all vacant positions at schools should be subject to competition among teachers currently employed in the public system. Committees must evaluate candidates according to their certifications, tenure, ability and discipline. Joint committees' decisions are mandatory for ministry and school administrators. The allocation of teachers to schools depends hence on teachers' preferences and committees' assessments of applicants' relative merit.

Broadly, the process works as follows: First, the joint committee announces to current teachers the list of schools in the state with available positions due to retirements, expansion of the school staff, etc. Second, interested teachers apply to specific schools. Third, the committee awards these positions. Afterwards, a process known as the *corrimiento* takes place. The school positions opened due to the between-school transfers done in the first-stage are now posted for applications. The process is repeated until no incumbent teacher is interested in the available school positions. Then, new teachers are assigned to these schools.

Joint committees do not have to follow the same criteria to allocate newly-hired teachers to schools with teaching vacancies and enjoy more discretionary power in this process. Hence, a priori, whether brand-new test and discretionary teachers are allocated to schools with different characteristics is an empirical question.

Table 1 shows descriptive statistics for the set of Telesecundaria schools that received either test (treatment) or discretionary (control) teachers hired in the 2010 school year – again, I focus only in schools receiving brand new teachers.[17] Schools that receive test (column 1) and discretionary teachers (column 2) are different, starting for – but not restricted to – the characteristics of the localities where they are located. A means by treatment status reported in table 1 are statistically significant at least at the 10-percent level, with the exception of the share of exams in the school detected for suspected cheating and the enrollment rate at the end of the academic year.

Test teachers tend to be allocated to schools in localities which are smaller, poorer, further away from the state capital and have less penetration of public services. For example, with respect to discretionary teachers, test teachers are allocated to schools located on average almost one hour further away from the state capital and with a poverty rate around ten-percentage points higher.

Regarding school characteristics, schools with test teachers have on average less students – 86.5 versus 104 in the control group –, an almost twice as much larger share of indigenous students (and indicator

[17] As said before, I exclude here and after schools that received both test and discretionary teachers and schools that received a test teacher hired in 2008 and 2009.

for poverty) – 23% versus 12.5% – and are less likely to have a principal with graduate school training – 18.4% versus 27.4%. As said before, Telesecundarias are small schools. Treatment schools have on average 86.6 students (18 less than those with discretionary teachers) distributed in around 4 classrooms (almost one less than the control schools); and so, the average class size in schools receiving new test teachers is 1.8 students larger than in schools with new discretionary teachers.

In the same line, test teachers are allocated to schools with – pre-treatment – lower performing students – from here and after I will refer to grade 9 scores as school scores unless I specify otherwise. Schools scores in treatment schools are lower on average by .13 standard deviation in Mathematics and Spanish. In both type of schools, the share of exams suspected of cheating is around 4.6%-6.7% , and the number of students that take this end-of-the-year exam is around 93-94 percent of those enrolled at the beginning of the school year.

As schools' location, inputs and outcomes are correlated, a regression analysis can be more informative about the process generating the allocation of teachers to schools than mere binary comparisons. Hence, I estimate a linear probability model in which the dependent variable is an indicator that turns 1 if the school received a test teacher in the school year 2010 (treatment) and 0 if received a discretionary teacher (control) in the same year. I regress this indicator on a vector of (past) school outcomes, inputs and locality characteristics, plus state fixed-effects. Table 2 reports the results.

First, it is noteworthy that, holding constant school inputs and locality characteristics, no single measure of student performance in the last two years predicts assignment into treatment. In the same line, the p-value associated to the test of the joint significance of all the (past) student-performance variables included in the model is very high (.662); and I cannot reject the null hypothesis that they are jointly insignificantly different from zero at conventional levels of statistical significance. So, the data does not seems to support an assignment model in which joint committees allocate test teachers based on past school performance.

Second, two variables of school inputs have a statistically significant relationship with the probability of receiving a test teacher. Class size is positively correlated with treatment status, while school size (total number of students) have a negative relationship with the probability of receiving a test teacher. However, the magnitude of the coefficients for class and school size is small (.00485 and -.00042).

Finally, it stands out the strong statistical relationship between treatment status and locality characteristics. Both the poverty rate and the share of households with electricity service in the locality have coefficients that indicate a statistically and economically significant relationship with the probability of treatment assignment. Summing up, *ceteris paribus*, test teachers are more likely to be allocated to schools in poorer with a lower penetration of public services like electricity.

Overall, the regression analysis indicates that treatment status is strongly correlated with locality characteristics and, in a lesser degree, with school inputs. Also, there is no observed relationship between

past school performance and the probability of assignment into treatment, once that locality characteristics are taking into account. These results are consistent with a model in which committees allocate test and discretionary teachers based on teachers' preferences for locality characteristics and not on past school performance and in which committees give a higher weight to the preferences of the discretionary teachers. This is encouraging evidence for a difference-in-difference analysis, in which is possible to control for both the effect of time-invariant locality characteristics and time-variant (observable) school inputs.

# 5 The Effect of Test-Hired Teachers

## 5.1 Identification and Estimation Methods

I am interested in estimating the average effect on student outcomes of assigning to a school a brand new teacher selected through a test-based examination (treatment) versus assigning a brand new teacher selected through a discretionary process with involvement from the teachers' union (control). My identification strategy takes advantage that the allocation of teachers to schools is not driven by past school performance, but by preferences of teachers over locality characteristics. With this purpose in mind, I estimate the following difference-in-differences model with school fixed-effects:

$$y_{st} = \beta_0 + \beta_1 T_{st} + \Gamma X_{st} + \tau_t + \alpha_s + v_{st} \tag{1}$$

Where $y_{st}$ is an outcome of school $s$ at time $t$, $T$ is the share of 2010 test teachers among total teachers in school $s$ at time $t$, $\beta$ is the parameter of interest, $X_{st}$ is a vector of time-variant school inputs - which include the number of 2010 brand new teachers in the school - and $\Gamma$ is the associated vector of parameters, $\tau_t$ is a vector of year effects fully interacted with state dummies, $\alpha_s$ is a school time-invariant (at least for the period of interest) component and $v_{st}$ is a disturbance term.

Under the parallel trend assumption, $\beta_1$ is an unbiased estimate of the average treatment effect on student achievement at the school level of receiving a new teacher hired using the test examination versus one hired in the discretionary process.

I estimate the model using panel data of Telesecundaria schools that receive either new test or discretionary-hired teachers in the 2010 school year. I have data from five years before treatment and one year after treatment. I focus in Telesecundarias because their small size should increase the likelihood that I find a statistically significant teacher effect at the school level. I approximate school outcomes with grade 9th outcomes.

There are reasons for which one could be interested in estimating the average treatment effect at the classroom level. For example, teachers may not teach all classrooms in a school. In Telesecundarias, teachers actually teach only one classroom per school. So, the classroom might be a more natural unit

to conceptualize the influence of a teacher. I cannot directly link teachers to classrooms in the data though. Beyond this limitation, the matching between teachers and students and potential within-school externalities can make the identification of a causal treatment effect at the classroom level more restrictive.

I describe and empirically investigate the allocation of teachers to schools and find strong support for the parallel trends assumption necessary for the identification of a causal treatment effect in a difference-in-differences model. I cannot do the same for the process generating the within-school allocation of teachers to students. Even when more detailed data is available, *Rothstein* (2010) gives a critical assessment of the typical assumptions about the assignment of students to teachers in which observational studies rely to identify teacher causal effects.[18]

The focus in the classroom level makes also easier to neglect within-school externalities associated to teacher quality. A higher quality teacher could for example free up other school resources - like principal's time - for the benefit of students in other classrooms. Also, higher quality teachers might have a direct effect on students in other classrooms through direct interactions.

My specification of treatment intensity provides a scaled-up treatment effect though. Assuming conditional independence of $T_{st}$, the difference-in-differences parameter $\beta_1$ captures the total (average) policy effect at the school level of increasing the share of test-hired teachers in a school from 0 to 1. Note that as I control for the number of brand new teachers in the school, I am comparing schools - of the same size - which receive, for example, one test teacher to schools which receive one discretionary teacher.

The causal interpretation of $\beta_1$ requires that the control schools give an accurate counter factual of the outcomes that the treatment schools would have had in the absence of treatment. Although it is impossible to test directly this assumption, I can take advantage of observing school outcomes for five years before treatment and test whether the secular trends in the treatment and control schools were the same in the pre-treatment period. Figure 1 gives a first approximation to the raw data. Here, I present the evolution of mean schools scores by (eventual) treatment status. The visual evidence is encouraging. As I describe in Table 1, schools that receive new test teachers in 2010 have on average lower school scores – pre-treatment – than schools that receive new discretionary teachers. But crucially for my identification strategy, the outcomes of both set of schools seem to follow the same time trend in the pre-treatment period and converge after treatment (year 2010).

More formally, following *Galiani et al.* (2005), I estimate a modified version of equation 1 in which I use a fixed-effects model to regress the outcomes under study - in separate regressions - on a vector of year dummies interacted with (eventual) treatment status - plus the set of time-variant school inputs and state-specific time trends. I only use observations from the five years in the pre-treatment period. Table 3 reports the full-estimated model and results.

In the same line that results in Table 2 and in Figure **??**, I do not find a statistically significant

[18] Although the Rothstein's critic is focused in the estimation of individual teacher effects which require more restrictive assumptions than the estimation of an aggregated effect.

relationship between treatment status and the the pre-treatment path of the four outcomes that I study: enrollment at the end of the academic year,[19] the share of students suspected of cheating in the school, and the school scores for Mathematics and Spanish in the Enlace exam. The coefficients for the interactions between (eventual) treatment status and year dummies in the regressions for final enrollment and share of flagged exams (columns 1 and 2) have, in general, small magnitudes and are not statistically significant. Two exceptions are the interactions between the year 2008 and eventual treatment status in the regression for final enrollment, and between year 2007 and eventual treatment status in the regressions for flagged exams. However, there is no evidence of a systematic pattern, the coefficients for the other years are smaller and given the large number of coefficients estimated - 16 - this result could arise because pure chance. In the regressions for the Mathematics and Spanish scores (columns 3 and 4), the corresponding coefficients are sightlier larger, but in no case a coefficient is statistically significant at the 10 percent-level, and I cannot reject the null hypothesis that the pre-intervention year dummies are the same for both control and (eventual) treatment schools at conventional levels of statistical significance. I interpret these results as strong evidence in favor of the parallel trend assumption necessary for the identification of a causal effect in a difference-in-differences model.

The total policy effect of using a recruiting method over other might comprise both: 1) the relative capacity of each method to identify and select teacher quality; and the propensity of (potential) candidates to apply through each of these methods. In other words, different sets of applicants might self-select into different recruitment methods. To what degree this happens or not, it is a question about the mechanisms through which the policy under study can relate to teacher quality.

Though $\beta_1$ is a relevant parameter from the policy point of view, it does not have the ceteris paribus interpretation of a parameter in the educational production as discussed by (*Todd and Wolpin*, 2003). Notably, there is no control for parental inputs that might react to changes in teacher quality induced by the policy. In principle, parents might increase or decrease the inputs they provide to students if they observe a change in teacher quality and teacher quality is a complement or a substitute of parental inputs. The estimated $\beta_1$ will underestimate the true difference in quality between the two groups of teachers if teacher quality is a substitute for parental time - as a recent paper by *Pop-Eleches and Urquiola* (2013) with data on Romanian high school suggests.

The incomplete take-up rate among test-hired teachers - described in the data section - could bias my estimates if changes the quality distribution of these teachers. I do not have information to directly investigate if the quality of the test-selected teachers that I observe in schools differ from those that I cannot match to any school. Given that test teachers are allocated on average to schools in localities with less desirable characteristics – poorer, with less coverage of public services, etc. – it is possible that the test selected candidates with better outside options - likely those with higher productivity - are those

[19] Measured as the number of students that take the Enlace exam over the number of students registered at the beginning of the academic year.

who are less likely to take up these teaching positions. If this is the case, the estimated $\beta_1$ will have a downward bias with respect to what would have happened with a full take-up rate.

## 5.2 Main Results

I present the main results of my difference-in-differences estimation in Table 4. The model in the all regressions controls for class size, school size, the share of indigenous students in the school, principal's attendance of graduate school and a vector of interactions between year and state dummies to capture state-specific time trends. Standard errors are clustered at the school level.

First, I do not observe an effect on the enrollment rate at the end of the academic year (see column 1). The coefficient of interest has a negative sign, but the magnitude is small (.03) and the point estimate is not different from zero at conventional levels of statistical significance. The same goes for the effect of test teachers on exam cheating (column 2). Going from not having test teachers in a school to having only test teachers is associated with a reduction of 2.2 percentage points in the share of students with flagged exams – which is not trivial compared to the mean in the control group (3.2%) –, but the result is not significant at conventional levels of statistical significance.[20]

As it is possible to see in columns 3 and 4, the allocation of new test-hired teachers has a positive effect on student achievement. The treatment coefficients in the student achievement regressions have a large magnitude (.45 standard deviation for Mathematics and .37 standard deviation for Spanish), and are statistically significant at the ten-percent level in the first case and the five-percent level in the second. In other words, moving from no test teachers in a school to only test teachers is associated with an increase in .45 standard deviations in the school's Mathematics test score and .37 standard deviations for the Spanish score. This large effect indicates an important gap in the quality of teachers hired through both methods.

## 5.3 Robustness checks

Telesecundarias cater small and isolated villages where local youth face high transportation costs to attend a regular junior high school located in larger locality. Given the limited school choice, I do not expect to observe that the allocation of test teachers change the composition of students in a school. A large correlation between the allocation of test teachers and the composition of students in the school would indicate the presence of underlying differential trends between treatment and control schools that my test on past outcomes dies not detect. I run hence my main model using the share of female and indigenous students in the school as outcomes. Table 5 reports the results.

I do not observe any effect of test teachers on both the share of indigenous and female students in the

[20] As explained in section 3.1.1, the Federal Ministry of Education measures exam cheating using a detection algorithm designed to give lower-bound estimates for direct copying, e.g. one student copying from another or a larger group of students (and potentially teachers) exchanging responses during the application of the exam.

school – see columns 1 and 2, respectively. The point estimates of interest have a small magnitude (.016 and .005 respectively) and in no case are statistically significant at conventional levels.

In the same line, I estimate my main model using two different measures of the school principal's education, indicators for whether the principal: graduated from college (column 3) and attended graduate school (column 4) - the reference is graduation from a teacher school. Again, finding a correlation between treatment status and principal's schooling would indicate that something else than the allocation of test teachers is going on at the treated schools. Results are reassuring. The coefficients of interest are small and not statistically significant.

For further robustness checks, I run a modified version of my main equation that allows each school to have an specific – linear – time trend. Hence, I first run separate regressions of every outcome on a full set of school dummies interacted with a linear time trend – using the five pre-treatment years – and use the residuals to predict – for the whole time period – a de-trended outcome. Then, I take the de-trended outcomes to my main specification. Table 6 reports the results. The same story emerges from this estimation. There is no effect of test teachers on enrollment at the end of year and neither on exam cheating – columns 1 and 2. However, there is a large effect on student achievement – see columns 3 and 4. Point estimates have both a magnitude and statistically significance similar to those shown in table 4.

# 6 Conclusions

Summing up, I find that education officials hire on average teachers of less quality when they follow a discretionary process with participation from the teachers' union than when they rely on the results from a standardized test. Then, joint committees of state officials and union representatives allocate these teachers to schools in more "desirable" localities. Taken together, these results suggest the existence of an agency problem with potential rent-extraction.

In the same line, Duflo, Dupas and Kremer (2012) observe that weaker institutional settings in Kenya increase the probability that hiring committees hire a teachers' relative.

State officials are able to reduce the extent of the agency problem by using a hard-to-manipulate rule to hire new teachers – a ranking based on a teacher test – instead of relying on a discretionary process with strong participation from the teachers' union. Test teachers significantly increase student achievement.

These results are in contrast with the previous literature – mainly U.S. based – on the relationship between teacher scores and achievement tests – summarized in *Hanushek and Rivkin* (2006). The evidence emerging from this literature indicates that though teacher scores might be more informative about teacher quality than other teacher characteristics (like experience and education), they are still a modest predictor of teacher quality. Again, the interaction between the institutional context and the alternative

hiring methods is key to determine the attractiveness of teacher tests as a policy to identify teacher quality.

The findings described in this study are particularly relevant for environments with weak institutions, where the lack of accountability (on actual teacher performance, for example) may exacerbate the incentives for rent-seeking behavior arising because of imperfect information about worker productivity.

## References


Duflo, E., P. Dupas, and M. Kremer, School governance, teacher incentives, and pupil-teacher ratios: Experimental evidence from kenyan primary schools, *NBER Working Papers 17939*, National Bureau of Economic Research, Inc, 2012.

Elizondo, C., *Por eso estamos como estamos*, Random House Mondadori, 2011.

Galiani, S., P. Gertler, and E. Schargrodsky, Water for life: The impact of the privatization of water services on child mortality, *Journal of Political Economy*, *113*(1), 83–120, 2005.

Hanushek, E. A., and S. G. Rivkin, *Teacher Quality*, *Handbook of the Economics of Education*, vol. 2, chap. 18, pp. 1051–1078, Elsevier, 2006.

Holland, P., Assesing unusual agreement between the incorrect answers of two examinees using the k-index, *Program Statistics Research*, 1996.

Hoxby, C. M., How teachers' unions affect education production, *The Quarterly Journal of Economics*, *111*(3), 671–718, 1996.

Iaies, G., M. Legarralde, and R. Inclan-Garza, Evaluación del sistema educativo de mexico, *Tech. rep.*, Fundacion Centro de Estudios en Politicas Publicas (CEPP), Consorcio Latinoamericano de Evaluacion de los Sistemas Educativos (CLESE) y Fundacion Konrad Adenauer (FKA), 2006.

Pop-Eleches, C., and M. Urquiola, Going to a better school: Effects and behavioral responses, *American Economic Review*, *103*(4), 1289–1324, doi:10.1257/aer.103.4.1289, 2013.

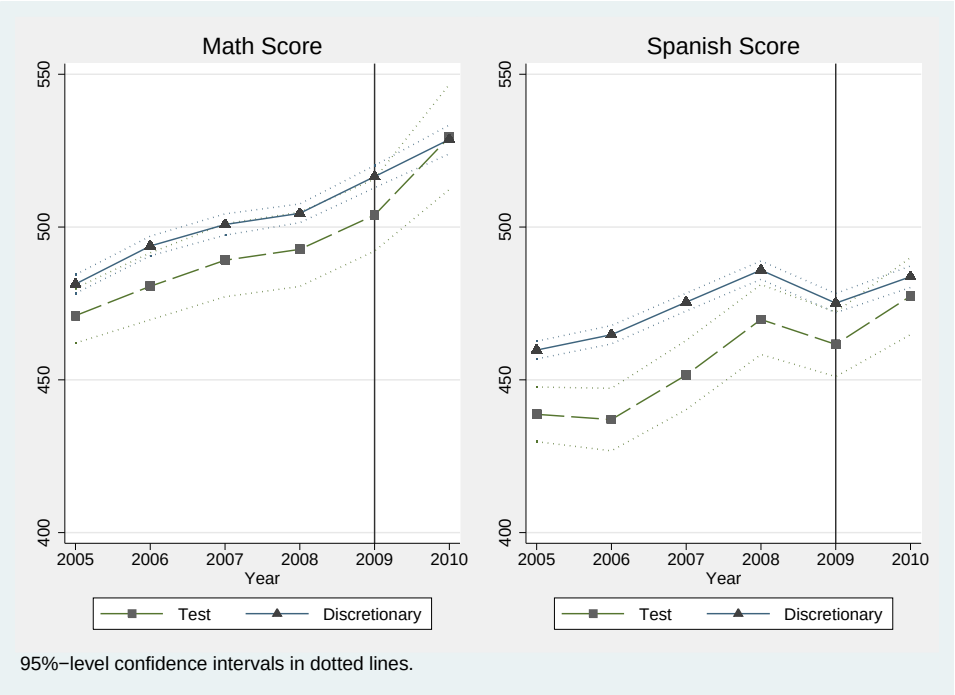
Rockoff, J. E., B. A. Jacob, T. J. Kane, and D. O. Staiger, Can you recognize an effective teacher when you recruit one?, *Education Finance and Policy*, *6*(1), 43–74, 2011.

Rothstein, J., Teacher quality in educational production: Tracking, decay, and student achievement, *The Quarterly Journal of Economics*, *125*(1), 175–214, 2010.

Santibanez, L., Reforma educativa: El papel del snte, *Revista Mexicana de Investigacion Educativa*, *13*(037), 419–443, 2008.

Todd, P. E., and K. I. Wolpin, On the specification and estimation of the production function for cognitive achievement, *Economic Journal*, *113*(485), F3–F33, 2003.

# Figures




Notes: Sample is composed of Telesecundaria schools which received a brand-new teacher in the year 2010. The two lines show the evolution of yearly means of school scores in the Enlace exam by eventual treatment status. Enlace scores are standardised at the national level with mean 500. Source: Enlace 2006-2011. Enlace results for 2005 corresponds to the academic year 2005-2006 and so on.

# Tables

Table 1: Descriptive Statistics Across Schools with Brand-New Teachers in 2010: Means (Standard Deviations) in 2009

| VARIABLES | (1) Received Test Teacher in 2010 | (2) Received Discretionary Teacher in 2010 |
|---|---|---|
| Math Score Grade 9 | 503.9 | 516.5 |
| | (55.79) | (52.67) |
| Spanish Score Grade 9 | 461.6 | 475.1 |
| | (49.89) | (45.34) |
| Flagged Exams Grade 9 | 0.0672 | 0.0463 |
| | (0.193) | (0.137) |
| Enrollment Grade 9 | 0.932 | 0.940 |
| | (0.0800) | (0.0774) |
| School size | 86.55 | 104.1 |
| | (65.53) | (91.54) |
| Class size | 19.67 | 17.90 |
| | (8.826) | (6.756) |
| Share indigenous students | 0.230 | 0.125 |
| | (0.423) | (0.321) |
| Principal has grad school | 0.184 | 0.274 |
| | (0.390) | (0.441) |
| Locality Population | 17,391 | 47,626 |
| | (84,667) | (194,363) |
| Hours to state capital | 2.844 | 1.997 |
| | (1.916) | (1.305) |
| Locality Poverty Rate | 0.745 | 0.638 |
| | (0.161) | (0.165) |
| Share hhs electricity | 0.883 | 0.952 |
| | (0.218) | (0.0738) |
| Observations | 87 | 805 |

Notes: All school statistics are for the school year 2009 (one year before treatment). Enlace scores are standardised at the national level with mean 500 and standard deviation 100. Sample is composed of Telesecundaria schools which received a brand-new teacher in the year 2010. Source: Enlace 2009, school census 2009 and population census data 2010. Enlace and school census results for 2009 corresponds to the academic year 2009-2010.

Table 2: Probability of Receiving a Test Teacher (OLS)

| VARIABLES | (1)<br>Received<br>Test Teacher in 2010 |
|---|---|
| Math 9 Lag 1 | -0.0302 |
| | (0.0224) |
| Spanish 9 Lag 1 | 0.00490 |
| | (0.0218) |
| Share Exams Flagged Lag 1 | -0.000703 |
| | (0.0910) |
| Final Enrollment Lag 1 | -0.111 |
| | (0.139) |
| Math 9 Lag 2 | 0.0114 |
| | (0.0285) |
| Spanish 9 Lag 2 | -0.00656 |
| | (0.0247) |
| Share Exams Flagged Lag 2 | 0.00230 |
| | (0.103) |
| Final Enrollment Lag 2 | 0.0934 |
| | (0.0671) |
| Class size Lag 1 | 0.00485** |
| | (0.00232) |
| Students Lag 1 | -0.000426*** |
| | (0.000133) |
| Share indigenous students Lag 1 | -0.0179 |
| | (0.0389) |
| Principal has grad school | -0.00310 |
| | (0.0188) |
| Locality Poverty Rate | 0.177** |
| | (0.0817) |
| Kms to state capital | 0.000121 |
| | (0.000148) |
| Locality Population | 4.10e-08 |
| | (2.92e-08) |
| Share hhs electricity | -0.602*** |
| | (0.166) |
| Share hhs sewage | 0.0364 |
| | (0.0436) |
| Constant | 0.457** |
| | (0.203) |
| Observations | 892 |
| R-squared | 0.175 |
| State Fixed Effects | Yes |
| F statistic Ho Var 1-8=0 | 0.733 |
| Prob>F | 0.662 |

Notes: Sample is composed of schools which received a brand-new teacher in the year 2010. Source: Enlace, school census and population census data. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Difference-in-Differences: Pre-Treatment Trends in Outcomes

| VARIABLES | (1) Final Enrollment (%) | (2) Share Flagged Exams (%) | (3) Math Score Grade 9 | (4) Spanish Score Grade 9 |
|---|---|---|---|---|
| 2006 X Share Test Teachers | -0.0291 | 0.0251 | 4.067 | -6.476 |
| | (0.0394) | (0.0309) | (13.01) | (12.30) |
| 2007 X Share Test Teachers | 0.0128 | -0.0741** | 9.893 | 16.75 |
| | (0.0459) | (0.0342) | (16.22) | (15.66) |
| 2008 X Share Test Teachers | 0.0822* | -0.00270 | 7.428 | 16.53 |
| | (0.0424) | (0.0464) | (18.39) | (13.86) |
| 2009 X Share Test Teachers | 0.00285 | -0.0180 | 2.043 | 11.22 |
| | (0.0411) | (0.0484) | (19.08) | (13.23) |
| Observations | 4,403 | 4,404 | 4,404 | 4,404 |
| R-squared | 0.076 | 0.048 | 0.138 | 0.137 |
| Number of id | 881 | 881 | 881 | 881 |
| Time-varying covariates | Yes | Yes | Yes | Yes |
| State X Year Dummies | Yes | Yes | Yes | Yes |
| F statistic Ho Var 1-4=0 | 1.910 | 1.848 | 0.140 | 1.226 |
| Prob>F | 0.107 | 0.118 | 0.967 | 0.298 |

Notes: All results are for grade 9th outcomes. Enlace scores (columns 3 and 4) are standardised at the national level with mean 500 and standard deviation 100. Source: Enlace 2006-2011, school census data 2006-2011 and Registro Maestros 2010-2011. Enlace and school census results for 2005 corresponds to the academic year 2005-2006 and so on. Standard errors in parentheses are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Difference-in-Differences: Results

| VARIABLES | (1) Final Enrollment (%) | (2) Flagged Exams (%) | (3) Math Score Grade 9 | (4) Spanish Score Grade 9 |
|---|---|---|---|---|
| Share Test Teachers | -0.0338 | -0.0224 | 44.56* | 37.07** |
| | (0.0411) | (0.0280) | (26.54) | (17.91) |
| New Teachers | -0.00151 | -0.00502 | -7.406** | -4.451 |
| | (0.00878) | (0.00548) | (3.745) | (2.921) |
| Constant | 0.941*** | 0.0253* | 500.9*** | 471.5*** |
| | (0.0129) | (0.0141) | (6.810) | (5.356) |
| Observations | 5,285 | 5,286 | 5,286 | 5,286 |
| R-squared | 0.076 | 0.043 | 0.181 | 0.142 |
| Number of id | 882 | 882 | 882 | 882 |
| Time-varying covariates | Yes | Yes | Yes | Yes |
| State*Year Dummies | Yes | Yes | Yes | Yes |
| Mean Control in 2010 | 0.938 | 0.0323 | 528.7 | 483.7 |

Notes: All results are for grade 9th outcomes. Enlace scores (columns 3 and 4) are standardised at the national level with mean 500 and standard deviation 100. Source: Enlace 2006-2011, school census data 2006-2011 and Registro Maestros 2010-2011. Enlace and school census results for 2005 corresponds to the academic year 2005-2006 and so on. Standard errors in parentheses are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Difference-in-Differences: Other Outcomes

| VARIABLES | (1) Share Indigenous Students | (2) Share Repeaters | (3) Principal is college grad | (4) Principal has grad school |
|---|---|---|---|---|
| Share Test Teachers | 0.0167 | -0.00582 | 0.132 | -0.0468 |
| | (0.0512) | (0.0103) | (0.130) | (0.0449) |
| New Teachers | 0.00752 | -0.000959 | -0.0208 | 0.0186 |
| | (0.0110) | (0.00128) | (0.0224) | (0.0202) |
| Class size | -2.54e-05 | -0.000328** | 0.000795 | -0.00463* |
| | (0.00127) | (0.000133) | (0.00277) | (0.00265) |
| School size | -0.000175 | 4.98e-06 | -0.000473 | 0.00118** |
| | (0.000220) | (1.49e-05) | (0.000478) | (0.000591) |
| Constant | 0.132*** | 0.0149*** | 0.239*** | 0.198*** |
| | (0.0220) | (0.00243) | (0.0430) | (0.0448) |
| Observations | 5,286 | 5,286 | 5,286 | 5,286 |
| R-squared | 0.040 | 0.045 | 0.018 | 0.029 |
| Number of id | 882 | 882 | 882 | 882 |
| Time-varying covariates | Yes | Yes | Yes | Yes |
| State*Year Dummies | Yes | Yes | Yes | Yes |

Notes: Source: School census data 2006-2011 and Registro Maestros 2010-2011. School census results for 2005 corresponds to the academic year 2005-2006 and so on. Standard errors in parentheses are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Difference-in-Differences with School-specific Linear Trends

| VARIABLES | (1) Final Enrollment (%) | (2) Flagged Exams (%) | (3) Math Score | (4) Spanish Score |
|---|---|---|---|---|
| Share Test Teachers | -0.0432 | -0.0283 | 50.24* | 33.08* |
| | (0.0606) | (0.0577) | (25.65) | (17.60) |
| New Teachers | 0.00169 | -0.0207*** | 2.315 | -3.904*** |
| | (0.00378) | (0.00457) | (1.820) | (1.449) |
| Constant | 0.0661 | 0.813 | -24.66 | 130.7 |
| | (1.395) | (1.799) | (671.6) | (531.5) |
| Observations | 5,285 | 5,286 | 5,286 | 5,286 |
| R-squared | 0.207 | 0.200 | 0.380 | 0.354 |
| Number of id | 882 | 882 | 882 | 882 |
| Time-varying covariates | Yes | Yes | Yes | Yes |
| State*Year Dummies | Yes | Yes | Yes | Yes |

Notes: All outcomes are de-trended using school-specific linear time trends. All results are for grade 9th outcomes. Enlace scores (columns 3 and 4) are standardised at the national level with mean 500 and standard deviation 100. Source: Enlace 2006-2011, school census data 2006-2011 and Registro Maestros 2010-2011. Enlace and school census results for 2005 corresponds to the academic year 2005-2006 and so on. Standard errors in parentheses are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1