

Wp

WORKING PAPERS

RULES VS. DISCRETION IN PUBLIC SERVICE:  
TEACHER HIRING IN MEXICO  
08/2017  
N° 2017/11

# RULES VS. DISCRETION IN PUBLIC SERVICE: TEACHER HIRING IN MEXICO

Estrada, R.

## RULES VS. DISCRETION IN PUBLIC SERVICE: TEACHER HIRING IN MEXICO

Estrada, R.

CAF – Working paper N° 2017/11  
08/2017

### ABSTRACT

In this paper, I use a unique empirical setting that allows me to compare the performance of teachers hired in a discretionary process led by the teachers' union in Mexico with the performance of those hired on the basis of a screening rule (test scores on a standardized exam). My results show that the discretionary hires perform considerably worse than the rule-based hires (as measured by value added to student achievement). The evidence presented here shows the impact of personnel selection mechanisms on the quality of public service delivery.

Small sections of text, that are less than two paragraphs, may be quoted without explicit permission as long as this document is stated. Findings, interpretations and conclusions expressed in this publication are the sole responsibility of its author(s), and it cannot be, in any way, attributed to CAF, its Executive Directors or the countries they represent. CAF does not guarantee the accuracy of the data included in this publication and is not, in any way, responsible for any consequences resulting from its use.

© 2017 Corporación Andina de Fomento

# REGLAS VS DISCRECIÓN EN EL SERVICIO PÚBLICO: SELECCIÓN DE DOCENTES EN MÉXICO

Estrada, R.

CAF - Documento de trabajo N° 2017/11  
08/2017

## RESUMEN

En este estudio, utilizo un contexto empírico único que me permite comparar el desempeño de docentes seleccionados en un proceso discrecional liderado por el sindicato de maestros en México con el de docentes seleccionados en un concurso basado en un examen estandarizado. Mis resultados muestran que los docentes seleccionados mediante el método discrecional tienen un desempeño considerablemente menor que el de los docentes seleccionados mediante el concurso (midiendo desempeño como el valor agregado al logro académico de los estudiantes). La evidencia presentada en este estudio revela la importancia de los métodos de selección de personal para la calidad de los servicios públicos.

Small sections of text, that are less than two paragraphs, may be quoted without explicit permission as long as this document is stated. Findings, interpretations and conclusions expressed in this publication are the sole responsibility of its author(s), and it cannot be, in any way, attributed to CAF, its Executive Directors or the countries they represent. CAF does not guarantee the accuracy of the data included in this publication and is not, in any way, responsible for any consequences resulting from its use.

© 2017 Corporación Andina de Fomento

# Rules vs. Discretion in Public Service: Teacher Hiring in Mexico

Ricardo Estrada

CAF–Development Bank of Latin America\*

August 2017

## Abstract

In this paper, I use a unique empirical setting that allows me to compare the performance of teachers hired in a discretionary process led by the teachers' union in Mexico with the performance of those hired on the basis of a screening rule (test scores on a standardized exam). My results show that the discretionary hires perform considerably worse than the rule-based hires (as measured by value added to student achievement). The evidence presented here shows the impact of personnel selection mechanisms on the quality of public service delivery.

JEL: I21, J51, M51.

Keywords: hiring methods, public service delivery, teachers' unions, teacher performance, teacher hiring.

---

\*Avenida Madero 900, Piso 15, Torre Catalinas Plaza, Ciudad de Buenos Aires, Argentina, email: restrada@caf.com. I thank the editor Oriana Bandeira, two anonymous referees and François Bourguignon, Marc Gurgand, Jérémie Gignoux and Andrea Ichino for detailed comments and suggestions. I am grateful for important insights to Yann Algan, Philippe Aghion, Joshua Angrist, Pablo Brasiolo, Alejandro del Valle, Pascaline Dupas, Gustavo Fajardo, Francisco Ferreira, Benjamin Feigenberg, Paul Glewwe, Miguel Jaramillo, Fabian Lange, Dario Maldonado, David Margolis, Thomas Zwick and participants at multiple seminars and conferences, and the Max Weber Economics Group at EUI. I also thank Jorge Santibanez for his insights about the Mexican education system, Kenneth Hounghbedji for sharing Stata routines, Rafael de Hoyos for sharing both expertise and data on the Programa Escuelas de Calidad and the staff at Mexicanos Primero for kindly providing me with access to data. Financial support from CONACYT and the Max Weber Programme at EUI is gratefully acknowledged. A previous version of this manuscript was circulated under the title "Rules Rather than Discretion: Teacher Hiring and Rent Extraction".

# 1 Introduction

Recruiting skilled and motivated individuals to deliver high-quality public services is a key challenge for governments everywhere, but is of particular importance for governments in developing countries (where weak institutions are more prevalent). The public sector's standard contractual employment arrangement, which provides low incentives, paired with a low accountability environment can hinder employee effort and prevent effective service delivery. These characteristics might also make employment in the public sector more attractive to lazy and corrupt individuals (Finan, Olken, and Pande, 2017).

The main challenge in differentiating good applicants from bad applicants is that employers usually have imperfect information about applicants' quality. In the absence of an objective measure that captures quality, organizations often rely on discretionary hiring—the subjective judgment of staff members to evaluate applicants. However, giving discretion to individuals may open the door to bias and, in an environment with low accountability, to favoritism or even rent extraction. Moreover, weak institutions may hinder the implementation of optimal hiring policies. In this setting, two related questions gain relevance: How costly is it to use discretion-based hiring within weak institutions? And how does discretion-based hiring fare in comparison to feasible alternatives? We have limited answers to these questions, though, because conducting a proper analysis requires both exogenous variation in recruitment methods—in otherwise similar organizations—and the possibility of measuring worker performance. In this paper, I address these challenges by using a unique setting that allows me to compare the performance (value added to student achievement) of teachers hired in a discretionary process led by the national teachers' union in Mexico to those hired on the basis of a screening rule.

Teachers provide an interesting case to learn about hiring with imperfect information. Ample research shows both that variance of teacher effectiveness is high and that identifying who actually is (or could be) a good teacher at the moment of hiring is difficult (Hanushek and Rivkin, 2006). The problem for policy makers (and parents) is that the characteristics that can easily be observed when recruiting are, at best, modest predictors of teacher performance (Rivkin, Hanushek, and Kain, 2005; Rockoff et al., 2011).<sup>1</sup>

In Mexico, public school teachers enjoy high (hourly) wages and high job security (Section 2), which makes teaching positions highly coveted, but does not necessarily translate into good teacher performance. In a recent survey, 70 percent of principals from junior secondary schools (grades 7 to 9) reported that the late arrival of teachers, absenteeism, or lack of pedagogical preparation hindered the quality of education provided at the school (OECD, 2013). Along the same lines, classroom observations in a representative sample of public schools revealed that 50 percent of instructional time is lost in Mexico City's primary schools (Bruns and Luque, 2014). Moreover, university students majoring in education (the main pathway to a teaching job) score

---

<sup>1</sup>Current teachers might have superior information or a higher ability to identify good teachers using broader criteria (Jacob and Lefgren, 2008; Rockoff and Speroni, 2011).

.2 standard deviations (SD) lower in mathematics (and .15 SD in Spanish) than students with health-related majors on a national standardized exam held at the end of secondary education (De Hoyos, Estrada, and Vargas, 2016).

State governments operate the public primary (grades 1 to 6) and junior-secondary schools, while the federal government sets the national curricula and provides states with the bulk of their education budget. The hiring of teachers and their assignment to schools is centralized at the state level, not the school level. Until 2014, the national teachers' union (the National Educational Workers Union, or SNTE) had the leading voice in a highly discretionary process used to hire new teachers. This process was controversial (criticisms included widespread nepotism and the selling of jobs) and led to a reform in 2008 that introduced a national examination based on a standardized exam to fill some teaching positions. In September 2013, a further step was taken to end discretionary hiring. The Mexican Senate passed a constitutional amendment that required all vacant teaching positions in public schools to be filled through a national and competitive examination as of 2014. "The inheritance and sale of jobs has ended," Education Secretary Emilio Chuayffet said. "Merit is the ideal means of access to, and progress in, a teaching career." (Weissenstein, 2013).

The 2013 reform scaled up the test-based (rule-based) recruitment process adopted nationwide after the 2008 reform, which means that both hiring systems co-existed from 2008 to 2013.<sup>2</sup> In this paper, I compare the performance of teachers hired in the discretionary process with those hired on the basis of a screening rule (test scores in a standardized exam). Specifically, I estimate the causal effect on student achievement of allocating to a school a discretionary hire vs a rule-based hire. For the empirical analysis, I exploit: the variation in hiring introduced by the 2008 reform; a new personnel data set that allows me to identify new teachers by hiring status and link them to a panel of student scores on a national standardized exam; and, for causal inference, a centralized allocation process of teachers to schools that does not depend on past trends in school outcomes.

I focus on a set of junior secondary schools that received new teachers in the 2010-2011 academic year (henceforth referred to as 2010 for simplicity's sake). Although the rule-based hiring was introduced in 2008, not all schools immediately received a rule-based hire. Both because not all new teachers were selected using the rule-based examination and because not all schools have vacant positions every year. So, I compare schools that in 2010 received either rule-based hires only or discretionary hires only and had previously not received a rule-based hire.

I am able to identify the 2010 cohort of teachers recruited through each of the hiring systems by using extensive data of school personnel compiled following a recent mandate of the Mexican Federal Congress. I match the teachers' data to panel data (which I compiled) on

---

<sup>2</sup>The name of the rule-based recruitment is (in Spanish) *Concurso Nacional para el Otorgamiento de Plazas Docentes*. The states of Michoacan and Oaxaca, strongholds of CNTE a smaller teachers' union, did not adopt the rule-based hiring. The 2013 reform also changed the evaluation and promotion of teachers, which were not part of the 2008 reform.

school characteristics and student scores in a national standardized test from 2005 to 2010 (five years before and one year after the allocation of teachers). I focus on Telesecundaria schools, a public system of junior secondary education that accounts for 20 percent of total enrollment at this schooling level. Telesecundarias are small schools, catering to small communities, with one teacher per classroom.<sup>3</sup>

To identify a causal effect, I rely on a centralized allocation process of teachers to schools that does not depend on past school trends in school outcomes. Figure 1 shows a first approximation to the raw variation in the data using a binary specification (the evolution of student achievement in schools that received new teachers in 2010). Before treatment, schools that received rule-based hires had lower performing students on average than schools that received discretionary hires. But, crucially for the identification of a causal effect in a difference-in-difference model (based on the parallel-trend assumption), both sets of schools follow similar trends in outcomes during the five-year period before treatment. After treatment, the outcomes converge.

My difference-in-differences estimates confirm that the allocation of rule-based hires has a positive and sizable effect on student achievement. In my preferred dose specification, moving from no rule-based hires in a school to only rule-based hires increases the school's mathematics test score by .53 SD and the Spanish score by .32 SD, a result that is statistically significant at the five-percent level in both cases. Given a one-to-one relation between teachers and classrooms in a school, this effect can be roughly interpreted as the average effect of a rule-based hire on student achievement in a classroom. Results are robust to different specifications and checks (Section 5.4).

The rule-based hiring might not have only selected (and attracted) better quality applicants, it might also have changed the (informal) incentives that the rule-based hires faced on the job, by diluting the role of union connections. I find that rule-based hires had a better academic background than discretionary hires as measured by average university GPA. I do not find though an statistical association between university GPA (or other observed hires' characteristics) and teacher performance—a common feature in the literature on teacher effectiveness (Rivkin, Hanushek, and Kain, 2005; Rockoff, Jacob, Kane, and Staiger, 2011). Although once on the job both types of hires faced the same formal incentives, the relational contracts of rule-based and discretionary hires were likely different. The union retained a large say in teacher promotion and transfers to other schools (Section 5). So, the discretionary hires might have needed less on-the-job effort to secure transfers(promotions) than the rule-based hires. The idea that incentives can improve teacher performance has support in the literature on education in developing countries (see Mbiti 2016; Evans and Popova 2016; Murnane and Ganimian 2014 for recent literature reviews). In India, for example, Muralidharan and Sundararaman (2011)

---

<sup>3</sup>Instead of specialist teachers, Telesecundarias rely heavily on IT teaching support. The television programs produced by the Secretariat of Public Education for this school system fill approximately two of the six hours of the school day.

find that giving well-designed financial incentives to teachers increases student test scores by .27 SD and .17 SD in math and language, respectively, while Duflo, Hanna, and Rya (2012) find that a combination of monitoring and financial incentives increases average test scores (math and reading) by .17 SD. In Kenya, Duflo, Dupas, and Kremer (2012) find that hiring a contract teacher increases student test scores only when parents (who participate in teacher hiring) are trained in school governance.

The main contribution of this paper is to the nascent literature on recruiting for public service delivery (see Finan, Olken, and Pande, 2017 for a review). In recent studies, Ashraf, Bandiera, and Lee (2016) in Zambia, Deserranno (2017) in Uganda and Dal Bó, Finan, and Rossi (2013) in Mexico use experimental variation to study how key attributes of public sector jobs (e.g. salary, career prospects and social value) affect the quality of the applicant pool (and the quality of service delivery in the case of the first two studies). The evidence presented here shows the importance of personnel selection mechanisms for the quality of public service delivery. Specifically, I document how recruiting on the basis of a rule rather than on discretion can improve efficiency in an environment with weak incentives and institutions.<sup>4</sup>

I also contribute to the literature on the effects of teachers' unions on school quality. There is a vivid controversy on whether and how teachers' unions push to set school inputs and policies that maximize school quality or deviate resources from this objective. This literature is relatively thin though, as there are important empirical challenges related to 1) the lack of appropriate variation to identify the causal effects of unions on school inputs and school quality; and 2) the absence of detailed data on specific mechanisms through which unions shape the education production function. In related studies, Hoxby (1996), Lovenheim (2009) and Kingdon and Teal (2010) use observational data to study the effect of unionization on school inputs and student outcomes. In this paper, I evaluate how the teachers' union affects school quality through a specific mechanism: teacher hiring.

## 2 Setting

### 2.1 Teachers' Union: SNTE

The SNTE is a national organization, formed in 1943, with more than one million members. Although there are other small teachers' unions in some states, SNTE affiliates almost all teachers, principals and workers in primary and junior secondary public schools, public teachers' schools, and the state and federal ministries of education. SNTE comprises a national executive committee and 55 regional sections (around two sections per state).

Both affiliation and payment of fees (about 1 percent of base salary) to the union is manda-

---

<sup>4</sup>More generally, Oyer and Schaefer (2011) assess the literature on personnel economics and argue for research directed at improving our understanding of how firms conduct hiring and the merits of different recruitment strategies.



tory and automatic for all teachers in public elementary schools, inclusive of new teachers. However, there is little accountability on the way that the SNTE leadership disburses the collected fees (Santibanez and Jarillo, 2007). The national leader of the union (Elba Esther Gordillo) was arrested in February 2013 on embezzlement charges (of around USD 200 million).

The teachers' union has an important say in the operation of the public education system, with direct participation in several bodies with decision-making authority. For example, the union is formally represented (jointly with education officials) on the state committees that define the hiring, allocation and promotion of teachers in schools. Through these joint committees, the union is also involved in the appointment of school principals and regional supervisors.

More broadly, the union is actively involved in the political system. SNTE has traditionally supported the Institutional Revolutionary Party (PRI), but formed alliances with the National Action Party (PAN) when this right-wing party was voted into the presidency of Mexico in 2000. SNTE created its own party, the New Alliance Party (PANAL), in 2005. As a result, union leaders are regularly elected as representatives and senators and hold leadership positions on the education committees of both the chamber of representatives and the senate.

The participation of union leaders is not limited to the legislative power. Santibanez (2008) argues that the union leadership trades electoral and political support for the appointment of union members to mid- and high-ranking positions in the federal and state ministries of education. More visibly, the head of the federal under-secretary for basic education in the 2006-2012 administration was an in-law of the union's national leader of the time.

The political clout of SNTE is reflected in the government's budget for education and the labor conditions of teachers. Public expenditures in education amount to 5.4 percent of GDP and 20 percent of total public expenditures—though still short of the minimum 8 percent of GDP mandated by an SNTE-backed amendment to the Federal Education Law approved in 2002. Tellingly, 86 percent of the education budget for primary schools is spent on teachers' salaries (the highest proportion among OECD countries) (OECD, 2014). Such a level of expenditure translates into high relative wages for public school teachers.

## **2.2 Teacher Compensation**

Table 1 compares the (mean) wages of primary and junior secondary teachers in public and private schools to those of the other college-educated workers in the labor market (using information from the National Labor Force Survey, ENOE), revealing a couple of notable issues.

First, public school teachers have a higher wage on average (both monthly and hourly) than private school teachers. The raw difference in the mean (hourly) wage is around 25 percent (and 21 percent after controlling for age and gender).

Second, though teachers from primary and junior secondary public schools tend to have a lower monthly wage than the rest of college-educated workers, the scenario changes once the number of hours worked is taken into account. The mean hourly wage of the elementary

school teachers is around 36 percent greater than the mean wage of the other college-educated workers (31 percent after controlling for age and gender), and it is similar to that of the teachers from other schooling levels and workers occupied as managers. The better pay of public school teachers in Mexico contrasts to the norm in the rest of Latin America (Mizala and Ñopo, 2014).

In addition, public school teachers have better non-wage benefits than private sector workers, including entitlement to more paid vacation time, a higher end-of-the-year bonus and practically guaranteed labor security (dismissal cases are rare, as public school teachers facing conflict with school administrators or parents are typically transferred to another school).

The statistics in Table 1 do not consider selection issues, but indicate that teachers are relatively well remunerated and a career as a teacher in the public service offers, at least for some sectors of the population, better working conditions than those of alternative jobs.

## 3 Teacher Hiring

### 3.1 Discretionary Hiring

State ministries of education are responsible for hiring teachers entering the profession in the public education system. Hiring is done according to the type and level of the teacher (primary school teacher, mathematics secondary school teacher, etc.) and in principle is not related to specific vacancies in schools.

Before 2014 (the year when the 2013 reform kicked in), the teachers' union would play a prominent role in the hiring process by directly selecting a share of the total number of hires in every state (Guevara and Gonzalez, 2004). Acting as a *de facto* hiring agent is actually a common function among unions in the public sector in Mexico. The federal law that regulates the labor relations for public employees establishes that unions have the right to directly select 50 percent of the hires for both new (permanent) positions and vacant positions that cannot be filled by a current employee.<sup>5</sup>

Although there is limited formal information about how exactly the teachers' union would run this selection process, it is clear that the union had ample discretion over who selected for hiring. That said, new hires had to comply with schooling requirements in terms of college education and type of degree.<sup>6</sup>

A controversial, but well established practice was entitlement, which allowed teachers who

---

<sup>5</sup>Prior to the 2008 reform, it was also considered that the teachers' union would informally control in a majority of states, through its influence in the appointment of state officials, the hiring done directly by the state ministries of education (see, for example, Santibanez (2008)). In support of these criticisms, a 2003 report carried out by the Federal Ministry of Education found that 19 of 32 states would not have in place any formal mechanism to select new hires and 5 would only conduct formal evaluations of applicants in a restricted manner (Guevara and Gonzalez, 2004).

<sup>6</sup>New hires must have university-level studies, though not necessarily a degree at the moment of hiring. Hires must be either graduates from teachers schools or from fields of studies related to the subject of the teaching position. No teaching certificate is required for hiring.

were retiring to pass on their job to a direct relative (offspring). A central argument used by local union leaders in their public opposition to the 2008 reform was that the use of a competitive examination would remove the union members' right to pass on their position (Elizondo, 2011).

Furthermore, strong and widespread media coverage and policy reports denounced the union's selling of teaching positions. A national survey among elementary teachers found that one-third of interviewees thought that selling of teaching positions was a frequent practice, while another third said it was done with limited frequency (Este Pais, 2005). An internal document of the teachers' union in the state of Oaxaca leaked to a national newspaper estimated the selling price of a teaching position in a Telesecundaria school—like those analyzed in this paper—to be up to 300,000 pesos (around USD 17,000) (Del Valle, 2015).

### **3.2 Rule-based Hiring**

In 2008, the federal government, under the umbrella of a broader agreement with the teachers' union, introduced a plan to fill all vacant teaching positions in public primary and junior secondary education through a competitive examination process. The exam (named the *Concurso Nacional para el Otorgamiento de Plazas Docentes* in Spanish) was open to candidates willing to enter the teaching profession in public schools and current teachers with temporary or part-time contracts, with hiring quotas for each group. In this paper, I focus on the recruitment of new teachers only.

The exam was based on a national, standardized test designed to measure cognitive skills, knowledge of the teaching subject, mastery of teaching methods and ethics. There was one exam for each type of teaching position (e.g. primary school teacher, mathematics junior secondary school teacher, etc.). Some types of positions were restricted to graduates of teacher training schools or with specific college majors. The number and type of available teaching positions by state and the exam results were widely publicized by media outlets and were available on a dedicated web page. The advertised positions were not associated with specific schools.

A national committee (*Comision Nacional Rectora* in Spanish) composed of high-ranking education officials and union representatives was put into place to oversee the design and implementation of the exam (with the support of a technical committee), while similarly integrated state committees (*Comites Estatales de Seguimiento* in Spanish) were in charge of implementation at the local level. Civil society organizations participated as monitors in different stages of the process, more visibly when the candidates took the exam, which was held before the beginning of the academic year.

The federal ministry of education processed the exams and ranked applicants by state and teacher type according to their exam results or, if states opted for it, a weighted average of the test score and other criteria (often university GPA). The state committees distributed hiring offers according to the applicants' rankings until all available positions were filled. The committees had discretion in matching applicants to specific schools (Section 5). In 16 states (half

of the total), selected applicants had to pass an additional examination (typically a health exam). Hires with an exam score below a state threshold had to undergo remedial training, as defined at the state level. Also, applicants had to correctly answer at least 30 percent of the questions to be considered fit for a teaching position. Only one percent of applicants scored below this threshold though.

The reform's emphasis on rule-based hiring met with strong opposition from state officials and local union leaders, so as a compromise, only new payroll positions funded by the federal government were filled through the rule-based recruitment initially, with other vacancies progressively opened to the rule-based hiring.<sup>7</sup> That said, in 2008, almost all states began to use rule-based hiring to fill some vacancies. According to figures from the Secretariat of Public Education, from the 22,546 full-time vacancies opened to rule-based hiring in 2010, the cohort that I study, 34 percent corresponded to new positions and the rest to existing payroll positions. There is no public information about the total number of new teachers hired through discretionary recruitment, though they could amount to around 85 percent of hires in 2010, according to my estimates.<sup>8</sup> At that time, no regulation prohibited test applicants to be recruited through the discretionary process (which, as I show later, some were). Then in 2013, the constitutional reform made testing the mandatory mechanism to fill all vacancies as of 2014.

## 4 Data

### 4.1 Enlace Exam

The data from my main outcomes on student achievement come from the Mexican Evaluation of Scholastic Achievement of Educational Institutions (Enlace), a national standardized test that students in primary and secondary school would take at the end of the academic year and which was given from 2005 to 2014. I use the individual test results to construct a panel dataset of school scores from 2005 to 2010 (five years before and one year after the treatment of interest).<sup>9</sup>

I observe ninth-grade scores for the whole period and seventh- and eight-grade scores from the school year 2008, when students from these grades started to sit the exam.<sup>10</sup> The test measured learning in mathematics, Spanish and a rotating subject, which I exclude from the analysis. Enlace scores were standardized at the national level with a mean of 500 and a SD of 100. I use a panel of ninth-grade scores in my main estimations because of the longer time dimension (2005-2010). However, I also present results using the data as a panel of classrooms for the 2008-2010 period.

---

<sup>7</sup>The teachers' union agreed to cede its selection entitlement over the 50 percent of the federally-funded new payroll positions.

<sup>8</sup>In the Telesecundaria System.

<sup>9</sup>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.

<sup>10</sup>The ninth-grade exam assessed materials of grades seventh to ninth before 2007, while after this year focuses in ninth-grade materials.

Enlace was designed to assess the overall educational system and, hence, had no bearing on students' GPA or graduation. However, Enlace results were widely reported by media outlets and NGOs. Also, since 2009, the Secretariat of Public Education delivered monetary bonuses to teachers of high-performing classrooms and schools within specific categories. Bonuses were distributed to teachers (and school principals) of classrooms or schools in the top 15 percent of the score distribution, and classrooms in the top 15 percent of the score gains distribution. Schools are classified by state into categories defined by locality characteristics (urban/rural and with high/low marginalization) and school type (general/technical/Telesecundaria/etc.). A teacher could receive a bonus ranging from MXN \$2,000 up to MXN \$20,000. This is around USD purchasing power parity 260 and 2,600, respectively; or 16 percent to 160 percent of the mean monthly wage of the new teachers in my sample.

## **4.2 School and Locality Characteristics**

I use the census of schools carried out by the Secretariat of Public Education (Formato 911) to obtain annual 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.<sup>11</sup>

## **4.3 Census of Teachers**

I benefit from extensive data of school personnel compiled as the result of a recent mandate of the Mexican Federal Congress. The data comprises the quarterly payrolls of public elementary schools from the second quarter of 2010 (the last quarter of the 2009-2010 academic year ) to the second quarter of 2011. 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 complete information about hiring, education profile or assigned classrooms.

I do not directly observe in the data who was hired since 2008—since rule-based examination was implemented. However, I use the 2009 and 2010 censuses to identify the 2010 cohort of new teachers. Then, I match these observations to the list of rule-based hires, which is available in the dataset. In this fashion, I create a dataset with the list of all schools that received (at least) one new teacher in 2010, selected through either the rule-based or the discretionary method. Using the school identification code, I link this dataset to the panel of Enlace school scores and the information on school and locality characteristics.

---

<sup>11</sup>Using the Stata command `traveltime`.

I restrict the dataset to schools that have never received a 2009 or 2008 rule-based hire, have at least 4 years of Enlace results, and did not receive both a rule-based hire and a discretionary hire in 2010. Finally, I cut schools with at least one year-to-year change in their school score larger than 2 SD, which roughly corresponds to the top and the bottom one percentile of score changes. This gives me a dataset with 1,427 schools in 15 states, with 6.7 percent of schools in the sample having received at least one rule-based hire in the school year 2010.

I give a detailed account of the procedure followed to generate this dataset in online Appendix A.

#### **4.4 Telesecundaria Schools**

I model the link between teachers and students at the school level for identification and data restrictions, so I focus my empirical analysis on the junior secondary Telesecundaria schools. The median school in my sample has 83 students, 4 classrooms and is located in a locality with 1,000 inhabitants. The small school size should increase the likelihood that I find a statistically significant teacher effect at the school level.

The Telesecundaria system was created to serve localities with under 1,500 inhabitants, but over time it expanded to larger rural communities and suburban areas. According to figures from the Secretariat of Public Education, around 1.26 million students attended 18,000 Telesecundaria schools in 2010, which amounts to 20.6 percent of total enrollment in junior secondary education.

Being mostly rural, Telesecundaria students tend to face more disadvantaged conditions than the average junior secondary school student. For example, in 2010 the average poverty rate in Telesecundarias communities in my sample was 62 percent, while the national poverty rate was 46 percent, according to the National Council for the Evaluation of Social Policy (CONEVAL).

Telesecundarias have one teacher per classroom, in contrast to general secondary schools that have one teacher per topic. Instead of specialist teachers, Telesecundarias rely heavily on IT teaching support. The television programs produced by the Secretariat of Public Education for this school system fill approximately two of the six hours of the school day. Hence, the effect of teacher quality in Telesecundarias is likely lower than in educational systems in which teachers play a larger role in the classroom.

### **5 Allocation of Teachers to Schools**

The 2008 reform introduced (within-state) variation in teacher hiring to the education system. Using this variation to meaningfully compare student outcomes in schools that receive rule-based versus discretionary hires requires that the matching between schools and teachers of the two types be independent of potential outcomes. In a difference-in-differences framework, such a requirement would be violated if state authorities allocated rule-based(discretionary)

hires to schools that would have improved with respect to the comparison schools even in the absence of rule-based hires; either because rule-based hires were allocated to schools 1) where student outcomes were already in a better trajectory before 2010, or 2) where improvements in student outcomes were expected in 2010 (perhaps for contemporaneous increases in other school inputs). Figure 1, presented in the introduction, shows a first piece of evidence against the first concern. I proceed now to study the institutional setting that governs the allocation of teachers to schools and generates such patterns.

The allocation of teachers to schools is centralized at the state level and is run by joint committees of state officials and union representatives. Joint committee decisions are mandatory for education officials and school principals, so the decisions of who to hire and what hiring procedure to follow are not a choice made by schools. The allocation is done in two steps. First, school vacancies are opened for applications among current teachers, which leads to sequential inter-school transfers, as the position left by a teacher who fills a vacancy is open for application to teachers from other schools. This process, known as *el corrimiento*, stops when no current teacher is interested in the available school positions. Then, new hires are assigned to these schools.

The allocation of current teachers (inter-school transfers) is highly regulated and depends on teachers' preferences and a joint committee's evaluation of their merits.<sup>12</sup> Typically, teachers progress in their careers by moving from schools in more isolated and poorer localities to schools in more urban and wealthier localities. The allocation of new hires (both rule-based and discretionary hires) does not have to follow the same criteria, and joint committees enjoy more discretionary power in this process. In principle, one could expect no difference in the characteristics of schools where rule-based and discretionary hires are allocated. However, such differences could arise if teachers prefer localities with better amenities, and union representatives are able to influence committee decisions and give a higher weight to the preference of the discretionary hires. The data are consistent with this hypothesis.

For simplicity, I start the analysis making a binary comparison between the schools that received either type of hire. Table 2 shows descriptive statistics for the set of Telesecundaria schools that received either rule-based (treatment) or discretionary (control) hires in the 2010 school year.<sup>13</sup> Rule-based hires tend to be allocated to schools that have lower performing students and a less-educated principals, are based in poorer localities further from the state capital and have lower penetration of public services.

Specifically, students in schools where rule-based hires are allocated have, pretreatment, lower scores in mathematics (by .22 SD) and Spanish (.19 SD)—on the Enlace exam—than those in schools with new discretionary hires. Consistent with these differences, treated schools are

---

<sup>12</sup>Joint committees operate under state-level regulations heavily based on a 1973 agreement between the Secretariat of Public Education and the SNTE. Joint committees must evaluate candidates according to their schooling, tenure, ability and discipline.

<sup>13</sup>I focus only in schools receiving new hires and exclude here and after schools that receive both rule-based and discretionary hires and schools that received rule-based hires hired in 2008 and 2009.

based in localities that are on average around 45 minutes further from the state capital and have higher poverty rates (by around 12 percentage points). Differences in school inputs are less clear. Class size is similar (19.1-19.8) in the two groups of schools, but schools with rule-based hires are less likely to have a principal with graduate school training (20 percent versus 32 percent).

As schools' location, inputs and outcomes are all 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 1 if the school received a rule-based hire in the school year 2010 (treatment) and 0 if it received a discretionary hire (control) in the same year. Then, in a second regression, I use the share of rule-based hires among the teachers in the school as dependent variable. I regress these variables on a vector of changes in (past) school outcomes, inputs and locality characteristics, plus state fixed-effects. Table 3 reports the results, which are qualitatively similar across both specifications.

Holding constant school inputs and locality characteristics, no single measure of change in student test scores in the last two years predicts assignment into treatment. Moreover, the p-value associated with the test of the joint significance of the change in student-performance variables included in the model is high (.267 in the binary and .547 in the dose specification); and I cannot reject the null hypothesis that they are jointly insignificantly different from zero at conventional levels of statistical significance. I obtain similar results if I use levels instead of rates of change to measure past school performance. So, the data do not support an assignment model in which joint committees allocate rule-based hires based on past school performance.

In contrast, the strong statistical relationship between treatment status and locality characteristics shown in the bivariate analysis remains after conditioning on student performance and school inputs. Rule-based hires are more likely to be allocated to schools in localities that are poorer, have a lower penetration of electrical service and are further from the state capital. Again, the relationship between treatment status and school inputs is less clear.

To sum up, the regression analysis indicates that treatment status is strongly correlated with locality characteristics and, to a lesser degree, with school inputs. Also, there is no observed relationship between past school performance and the probability of assignment into treatment, once locality characteristics are taken into account. These results are consistent with the hypothesis that committees allocate rule-based and discretionary hires by considering teachers' preferences for locality characteristics and not past school performance. They are also consistent with the hypothesis that committees give a higher weight to the preferences of the discretionary hires. The participation of the teachers union in the process to allocate hires to schools also lessens the concerns about the capacity of state officials to assign rule-based hires to schools where test scores were expected to improve in 2010. In Section 6.3, I investigate further the hypothesis that rule-based hires were assigned to schools with contemporaneous improvements in school inputs. Overall, the evidence presented in this section is encouraging for a difference-in-



difference analysis, in which it is possible to control for both the effect of time-invariant locality characteristics and time-variant (observable) school inputs.

## 6 The Effect of Rule-Based Hires

### 6.1 Identification and Estimation Methods

To estimate the average effect on student outcomes of assigning to a school a new teacher selected in a rule-based examination (treatment) versus a new teacher selected in a discretionary process (control). I take advantage of the centralized allocation process of teachers to schools that is not based on past school trends. With this purpose in mind, I estimate the following difference-in-differences model with school fixed-effects:

$$y_{st} = \beta_0 + \beta_1 share_{st} + \Gamma X_{st} + \tau_t + \alpha_s + v_{st} \quad (1)$$

Where  $y_{st}$  is an outcome of school  $s$  at time  $t$ ,  $share_{st}$  is the share (proportion) of 2010 rule-based hires among total teachers in school  $s$  at time  $t$ ,  $\beta_1$  is the parameter of interest,  $X_{st}$  is a vector of time-variant school inputs,  $\Gamma$  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 component and  $v_{st}$  is a disturbance term.

I estimate the model using panel data of Telesecundaria schools that receive new hires, either rule-based or discretionary-hires, in the 2010 school year. I focus on Telesecundarias because their small size should increase the likelihood that I can statistically observe a teacher effect at the school level. I approximate school outcomes with ninth-grade outcomes in my main estimations. I do so because I observe ninth grade test scores for a longer period than eighth and seventh grade test scores, and hence I can study and control better for pretreatment trends in outcomes.<sup>14</sup>

The model in all regressions controls for the number of 2010 hires in the school, class size, school size, the share of indigenous students, 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.

Under the parallel trends assumption, the difference-in-differences parameter  $\beta_1$  captures the total (average) effect of increasing the share of rule-based hires in a school from 0 to 1 (or from 0 to 100 percentage points). Or in other words,  $\beta_1$  measures the difference in the (total) value-added (teacher quality) to student outcomes between rule-based and discretionary hires.<sup>15</sup>

I model the relationship between teachers and students at the school level for identification,

---

<sup>14</sup>I have ninth-grade test scores from 2005 to 2010 (five years before and one year after treatment), while I only have 7th and 8th-grade test scores from 2008 to 2010.

<sup>15</sup>The share of rule-based hires among total teachers in treated schools is 41 percent. The distribution of this share is available in online Appendix B.

though there are reasons to be interested in estimating the average treatment effect at the classroom level. For example, teachers may not teach all students in a school. In Telesecundarias, teachers are actually with the same group of students in one classroom all day, so the classroom might be a more natural unit to conceptualize their influence. I cannot directly link teachers to students in the data, though. Even if I could, the matching between teachers and students and the potential within-school externalities of teacher quality make the identification of a causal effect at the classroom level more restrictive.

In my empirical investigation of allocation of teachers to schools, I find strong support for the parallel trend assumption necessary for the identification of a causal effect at the school level 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 are available, Rothstein (2010) gives a critical assessment of the typical assumptions made about the assignment of students to teachers in which observational studies rely on to identify teacher causal effects.<sup>16</sup>

The focus at the school level allows me also to neglect the within-school externalities associated with teacher performance. More effective teachers could, for example, free up other school resources, like a principal's time, for the benefit of students in other classrooms. Also, better teachers might have a direct effect on students in other classrooms, for example, through personal interactions.

My specification of treatment intensity provides a scaled-up treatment effect. Given a one-to-one relation between teachers and classrooms in a school,  $\beta_1$  can also be interpreted as the average effect of a rule-based hire on student achievement in a classroom.

## 6.2 Past Trends in Outcomes

The causal interpretation of  $\beta_1$  requires that the control schools give an accurate counterfactual of the outcomes that the treated schools would have had in the absence of treatment. Although it is impossible to directly test this assumption, I take advantage of observing school outcomes for five years before treatment and test whether the trends in both sets of schools were the same during the pretreatment period.

Figure 1 shows the raw data, presenting the evolution of mean school scores by (eventual) treatment status using a binary specification. The visual evidence is encouraging. Schools that receive rule-based hires in 2010 have on average lower school scores, pretreatment, than schools that receive discretionary hires. But crucially for my identification strategy, the outcomes of both sets of schools seem to follow similar paths during the pretreatment period, and then converge after treatment (year 2010).

More formally, I estimate a modified version of equation 1 in which I regress the outcomes under study, in separate regressions, on a vector of year dummies interacted with eventual treat-

---

<sup>16</sup>Although the Rothstein's critic is focused in the estimation of individual teacher effects which require stronger assumptions than the estimation of an aggregated effect.

ment status (using the share of new hires among teachers in the school), plus a set of time-variant school inputs, state-specific time trends and school fixed-effects. I use only observations from the five years in the pretreatment period. Table 4 reports the results.

Along the same lines of results in Table 3 and Figure 1, I do not find a consistent statistical relationship between treatment status and the pretreatment path of the three outcomes that I study: enrollment at the end of the academic year,<sup>17</sup> and the scores for mathematics and Spanish on the Enlace exam. Only one of the 12 estimated coefficients for the interactions between eventual treatment status and year dummies is statistically significant.

The interaction between eventual treatment status and the year 2008 in the regression for final enrollment (column 1) is highly statistically significant and has a large magnitude (10 percentage points). However, the other three coefficients in the same regression have small magnitudes (less than or equal to 1.4 percentage points) and are not statistically significant. Similarly, no coefficient of interest in the regressions for mathematics and Spanish scores (columns 2 and 3) is statistically significantly different from zero at conventional levels. The magnitude of three of the eight coefficients for mathematics and Spanish are relatively large, of around .10 to .15 national SD. But, again, in no case is a coefficient statistically significant at the 10-percent-level. Furthermore, all coefficients have a negative sign. Hence, if anything rule-based hires would be allocated to schools with a negative trend in test scores, which if true would produce a downward bias in my estimates. Overall, I cannot reject the null hypothesis that the pre-intervention year dummies are the same for both control and treatment schools at conventional levels of statistical significance in two of the three regressions.

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.

### 6.3 Main Results

I present the main results of my difference-in-differences estimation in Table 5. I first evaluate whether the allocation of rule-based hires changes the enrollment rate at the end of the academic year (see column 1), and I do not observe any effect. The coefficient of interest has a negative sign, but the magnitude is very small (1.3 percentage points) and the point estimate is not different from zero at conventional levels of statistical significance.

Turning to the main outcome of interest, the allocation of rule-based hires has a positive impact on student achievement in mathematics and Spanish (columns 3 and 4), which is both statistically and economically meaningful. The treatment coefficients in the achievement regressions have a large magnitude (.53 national SD for mathematics and .32 national SD for Spanish), and are statistically significant at the one-percent level in the two cases. These coefficients imply that moving from having no rule-based hires in a school to only having teachers

---

<sup>17</sup>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 rule-based hires increases the school's mathematics(Spanish) test score by .53 SD (.32 SD). As mentioned before,  $\beta_1$  captures the total (average) effect of increasing the share of rule-based hires in a school from 0 to 1. However, this parameter can be roughly interpreted as the average effect of a rule-based hire on student achievement in a classroom.<sup>18</sup>

## 6.4 Characteristics of the Hires

I turn now to study differences in observable characteristics between the individuals hired through each method. Unfortunately, the teacher census only includes career and education information of rule-based hires and discretionary hires who applied to the rule-based examination.<sup>19</sup> To better understand the specifics of this subsample of discretionary hires, I estimate my main model using only them as a comparison group. Results (in Table 9, columns 1 and 2) show that the difference in teacher performance between the discretionary hires in this subsample and the rule-based hires is similar to the one found in the main results.<sup>20</sup>

Table 10 shows summary statistics of hire's characteristics by recruitment method. Compared to the discretionary hires, rule-based hires tend to be younger (by 2.1 years), to be less likely of having worked in the public sector (as a teacher or in other occupation) (by 30 percentage points) and to have a higher university GPA (by .5 SD). The difference in average university GPA is substantial. There are not statistically significant differences in gender composition and private sector experience between the two groups of hires.

To study the relationship between hire characteristics and teacher performance, I adapt my main model to run a regression between value-added to student test scores and hires' characteristics. To do so, I first collapse the pretreatment years data in one pretreatment period and then I estimate changes between the pretreatment and posttreatment periods. I regress this measure of changes in student test scores on a vector of (2010) hire's characteristics (the school average in case of more than one hire in the school) and a vector of changes in school inputs. Columns 3 and 4 in Table 9 report the results. As it is possible to observe, I do not find a statistical relationship between hires' characteristics and their teaching performance. None of the coefficients for hire's characteristics is statistically significant at conventional levels. This result must be taken with a grain of salt as measurement error could bias the coefficients toward zero and

---

<sup>18</sup>It is possible that behind the documented average effect there is a pattern in which the two group of teachers tend to focus on different types of students. For example, discretionary hires could lean more towards the welfare of relatively low-achieving students and focus the teaching on these, while rule-based hires focus the teaching on high-achieving students. If learning gains depend strongly on the level of past achievement, such specialization could produce the average effect observed, with important implications for the interpretation of this result. I investigate this possibility and find that the effect of rule-based hires in both mathematics and Spanish achievement is similar at the 25th and 75th percentile of the school score distribution – and mimics the main results – (see Table 11 in online Appendix C). In other words, the allocation of a test-hired teacher benefits both high and low-achieving students in a school.

<sup>19</sup>I do not observe though the specific teaching position they applied (Telesecundaria or other), year of application nor the test score obtained.

<sup>20</sup>In the same fashion, the pretreatment trends in outcomes in this sample follow parallel trends (see Appendix D).

the sample size be too small to detect small effects. However, the weak association between teacher characteristics and teacher performance is a common result in the literature on teacher effectiveness (Rivkin, Hanushek, and Kain, 2005; Rockoff et al., 2011).

## **7 Robustness Checks**

### **7.1 Spurious Outcomes**

Telesecundarias cater to small, isolated villages where local youth face high transportation costs to attend a regular junior secondary school located in a larger locality. Given the limited school choice, I do not expect to observe that the allocation of rule-based hires changes the composition of students in a school at the beginning of the academic year. A large correlation between the allocation of rule-based hires and the composition of students in the school would indicate the presence of underlying differential trends between treatment and control schools that the test on past outcomes fails to detect, but that could bias the reported results. With this purpose in mind, I run my main model using as outcomes the share of indigenous students and of students repeating grades in the school. Table 96 reports the results.

I do not observe any correlation between the allocation of rule-based hires and the share of indigenous students or repeaters in the school (see columns 1 and 2, respectively). The point estimates of interest have a small magnitude (4.5 percent and .02 percent respectively) and are not statistically significant at conventional levels.

### **7.2 Contemporary Shocks in Other Inputs**

I also estimate my main model using as outcomes two measures of school inputs: an indicator for whether the principal has a graduate education and class size. The idea is to investigate the presence of potential contemporaneous shocks that could introduce bias in the main results. The concern would be that for some policy (about which I do not have any information) or the interest of education officials, the schools where rule-based hires were allocated experienced at the same time an improvement in (other) school inputs. Results are, again, in Table 5.

Both the coefficients of interest for the principal's education (column 3) and class size (column 4) have a negative sign and small magnitude (4.6 percentage points and .04 students). The point estimate for the principal's education is marginally significant, while the one for class size is not statistically significant at the ten-percent level. So, I do not find evidence of contemporaneous positive shocks in other inputs in the schools that receive rule-based hires—at least in the school inputs that I can observe.

### **7.3 School Time Trends**

For further robustness checks, I modify the main equation to allow each school to have a 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 pretreatment years) and use the residuals to predict (for the whole time period) a de-trended outcome. Then, I apply the de-trended outcomes to my main specification. Panel A in Table 7 reports the results.

The same story as before emerges from this estimation. There is no effect of rule-based hires on enrollment at the end of year (column 1), but there is a large effect on student achievement (columns 2 and 3). Point estimates have both a magnitude and statistical significance similar to those shown in the main results in Table 5.

### **7.4 Binary Treatment**

Up to now, I have modeled treatment as the share of rule-based hires in the school with the purpose of providing a scaled up estimate of a teacher effect in a regression run at the school level. One could be worried though that a few schools with a high share of rule-based hires drive the results reported so far. Hence, I produce estimates using a binary specification for treatment: whether or not at least one rule-based hire is allocated to the school. The counterfactual is again the allocation of a newly-hired discretionary hire(s). I restrict the estimation to small schools, those with three classrooms or less, as it is highly unlikely that I can find a statistically significant effect in large schools. Panel B in Table 7 reports the results which, again, are reassuring.

No changes in the conclusions arise from this set of estimates: the allocation of rule-based hires does not have an effect over enrollment at the end of the academic year, but it increases student achievement, measured by the mathematics and Spanish scores on the Enlace exam.

### **7.5 Heterogeneous Response**

In support of the parallel trend assumption, I showed in Table 4 that treatment and control groups followed similar paths in the outcomes under study during the five years prior to treatment. However, another source of concern arises from the potential interaction between the outcome variables and exogenous characteristics with unbalanced distributions between schools with rule-based and discretionary hires (Meyer, 1995; Heckman, Ichimura, and Todd, 1997). For example, such as whether the impact of rule-based hires is larger in isolated localities.

Abadie (2005) proposes a semi-parametric difference-in-differences estimator to deal with non-parallel outcome dynamics for the treated and control groups due to differences in observed characteristics. The estimation uses a two-step strategy in which first a propensity score is estimated and then a matching estimator re-weights the control observations on the propensity score and imposes as common support a balanced sample in pretreatment characteristics between the

treated and the control groups.

Hence, I estimate first the propensity score from a linear model of the probability that a school receives a rule-based hire in the school year 2010 as a function of similar pretreatment characteristics to those listed in Table 3 plus two lags of final enrollment and exam cheating. The model includes third-order polynomial functions for all the (continuous) variables. I impose a common support by dropping both treatment observations of which the propensity score is higher than the maximum propensity score of the control observations, and control observations of which the propensity score is lower than the minimum propensity score of the treated observations. The Abadie estimator matches differences in pretreatment and posttreatment outcomes for the treated to weighted averages of differences in pretreatment and posttreatment outcomes for the untreated. Hence, I use only observations for one year before and one year after treatment. I use a binary definition of treatment and restrict the estimation to small schools (those with three classrooms or less). Results are reported in Panel C on Table 7.

The main results are consistent with those in Panel B on the same table. The allocation of a rule-based hire does not have an discernible effect on the share of students in the school that take the Enlace exam at the end of the academic year (column 1). Once again, I observe a positive effect on mathematics scores of .27 national SD, which is statistically significant at the one-percent level and a positive effect on the Spanish score of .15 national SD, statistically significant at the ten-percent level.

### **7.5.1 Classroom Cohorts**

I have run all the regressions on a panel of ninth-grade scores because to study and control for pretreatment trends in outcomes using the longest possible time series. The Enlace exam was administered to ninth graders beginning in 2005, whereas it was expanded to the seventh through the ninth grades in only 2008. As an alternative measure, then, I can use the outcomes of students in the eighth and the seventh grade in 2009 and 2008, respectively, the two years prior to treatment, to investigate the robustness of my results to the choice of pretreatment outcomes.

Figure 2 shows the raw data in a binary specification and Panel A on Table 8 the falsification test to investigate if differences in pretreatment trends are correlated to eventual treatment status (in a dose specification). I do not find any correlation between the allocation of rule-based hires in 2010 and changes in student outcomes between 2008 and 2009. The coefficient for the partial correlation between eventual treatment status and changes in enrollment at the end of the academic year (column 1) has a very small magnitude (.06 percentage points) and is not statistically significant at conventional levels. The coefficient of interest is relatively large in the regression for mathematics score (.17 SD, see column 2), but it is not statistically significant at conventional levels and the sign is negative. The magnitude of the corresponding coefficient in the regression for the Spanish score is smaller (.05 SD, see column 3) and is not statistically significant. This test gives additional support for the identification of a causal effect based on

the parallel trend assumption.

The results, available in Panel B on Table 8, are in line with those presented earlier. First, there is no observed effect on enrollment at the end of the academic year. The coefficient of interest has a small magnitude and no statistical significance, see column 1. Then, there is a significant effect on student achievement (see columns 2 and 3), measured by both mathematics and Spanish scores. The coefficients of interest confirm that rule-based hires have a large and economically significant effect on student outcomes, as with those reported in the main specification (Table 5).

## 8 Discussion

I presented several checks supporting the identification of the causal effect on student achievement of allocating a school with a rule-based hire versus a discretionary hire. Results show that children in schools that received rule-based hires learned considerably more math and Spanish than children in schools that received discretionary hires. I turn now to discuss differences in the incentives and selection patterns associated with each hiring method that could have potentially contributed to this large gap in teacher performance.

In theory, rule-based hiring could have increased average hire quality either by improving the applicant pool or by better screening of the candidates. With respect to pooling, the promise of a meritocratic recruitment and a teaching career less dependent on union connections might have attracted higher quality (and more performance-oriented) candidates to the rule-based process—particularly if applicant quality was not positively correlated with connections to the union. One lesson from the nascent literature on recruiting for public service delivery is that job attributes can significantly affect the applicant pool (Deserranno, 2017; Ashraf, Bandiera, and Lee, 2016; Dal Bó, Finan, and Rossi, 2013). In addition, testing could also have increased (average) applicant quality if quality was negatively correlated with the cost of test preparation or positively correlated with (expected) test scores.<sup>21</sup> Although, I have limited information about the hire’s characteristics, Section 6 documents that rule-based hiring led to the recruitment of higher quality applicants as measured by their average university GPA.

With respect to screening, there is a large literature that shows that teacher test scores frequently predict teacher performance (value added to student achievement) (see for reviews: Wayne and Youngs, 2003; Hanushek and Rivkin, 2006; and Glewwe, Hanushek, Humpage, and Ravina, 2011), particularly when the test measures knowledge of the teaching subject. Teacher test scores tend to have a stronger correlation with teacher effectiveness than characteristics like education, experience (beyond the two years) and salaries. However, even when significant, teacher test scores capture only a small portion of the overall variation in teacher performance.

---

<sup>21</sup>Note that testing imposes an application cost that if it is independent of applicant quality can have a stronger deterrent effect on higher quality applicants. See Angrist and Guryan (2008) for a model of the effect of teacher testing on selection.



In other words, teacher test scores are unlikely to be a silver bullet to detect high-quality applicants holding constant the applicant pool. However, basing hiring decisions solely on test scores might still be an improvement (in terms of pooling and screening) over using discretion in this context.

In principle, current teachers might have superior information or a higher ability to identify teacher quality using broader criteria (Jacob and Lefgren, 2008; Rockoff and Speroni, 2011). However, two elements could lead to the selection of lower quality applicants. The first is the prevalence of evaluation bias. The second, and more worrisome, is the abuse of discretion that motivated the introduction of the rule-based hiring. Note that if an agent decides to sell a teaching position (demands a bribe from applicants), the side payment is maximized by selecting the worse applicant, the one who has the lowest outside option in the labor market and hence the incentive to pay the highest bribe.<sup>22</sup>

Once on the job, both types of hires faced the same formal incentives. The 2008 reform did not change the regulations for teacher evaluation, promotion or firing. The practice of granting tenure automatically after three months on the job gave job security to all hires, and union affiliation and payment of dues was mandatory. However, the relational contracts of rule-based and discretionary hires were likely different. The union held a large say in teacher promotion and transfers to other schools, and anecdotal information and the results presented in Section 5 suggest that having close ties to the union helped teachers achieve better school (locality) assignments. Accordingly, discretionary hires might have needed less on-the-job effort to secure transfers (promotions), while rule-based hires—having been appointed through a meritocratic system—might have seen job performance as a more promising way to progress in their teaching career. The prevalence of high levels of absenteeism and instructional time lost suggest there was a large potential for improvement on this margin (Bruns and Luque, 2014). Although I do not have information on absenteeism or classroom activities to empirically investigate this hypothesis, the idea that incentives can improve teacher performance has support in the literature on education in developing countries (see Mbiti 2016; Evans and Popova 2016; Murnane and Ganimian 2014 for recent literature reviews).

## 9 Conclusions

One of the biggest challenges of the public sector is to recruit skilled and motivated individuals. This is especially true in contexts with low accountability. However, we still know little about how should organizations hire the workers they need to deliver high-quality services. The evidence presented here contributes to filling this gap by showing the importance of personnel selection mechanisms for the quality of public service delivery. As documented in this paper, Mexico's recruiting of new teachers through a discretionary process run by the teachers'

---

<sup>22</sup>Similarly, a retiring teacher who maximizes family income has the incentive to select the relative with the worst outside option.

union led to the hiring of individuals who had much worse performance (value added to student achievement) than those selected on the basis of a screening rule. One contribution of this paper is to evaluate an ambitious reform in a large education system (with around 1 million teachers).

The reform evaluated here was introduced amid considerable criticism for the abuse of discretionary hiring, including accusations of selling of positions and passing them on to relatives upon retirement. By reducing the opportunities for manipulation, rule-based hiring might not only have selected (and attracted) better-quality applicants, but it might also have changed the (informal) incentives that rule-based hires faced on the job, diluting the role of union connections. Unfortunately, I do not have the appropriate data to empirically investigate these mechanisms. I therefore leave it to future research to study the relationship between personnel selection mechanisms and both applicants' and hires' quality (selection and screening), and the effect of these selection mechanisms on the incentives that hires face once they join the bureaucracy.

## References

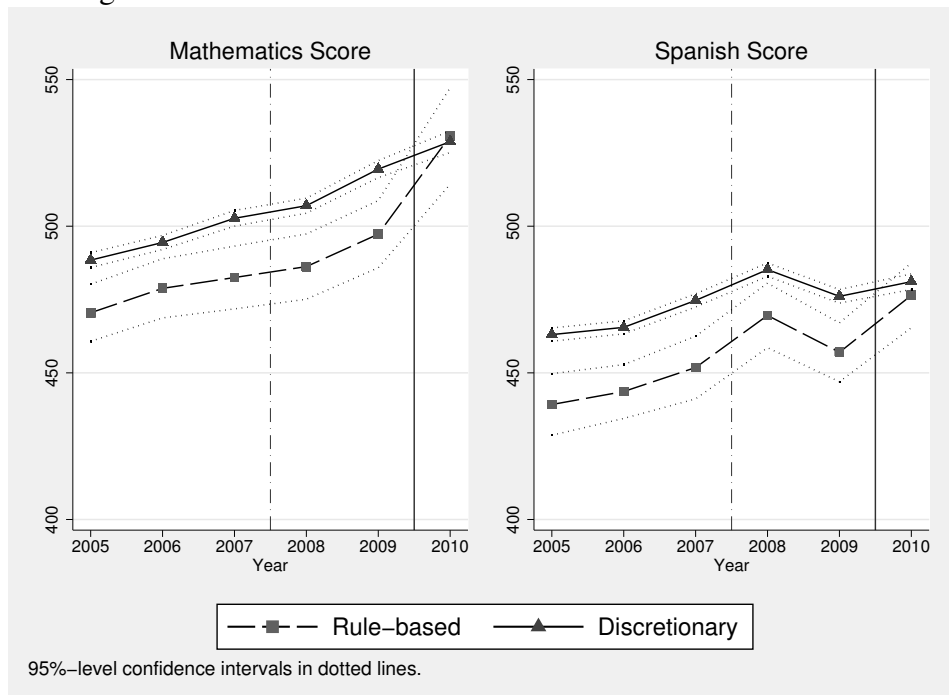
- Abadie, A. 2005. “Semiparametric Difference-in-Differences Estimators.” *Review of Economic Studies* 72:1–19.
- Angrist, J.D., and J. Guryan. 2008. “Does teacher testing raise teacher quality? Evidence from state certification requirements.” *Economics of Education Review* 27:483–503.
- Ashraf, N., O. Bandiera, and S.S. Lee. 2016. “Do-gooders and go-getters.” Working paper.
- Bruns, B., and J. Luque. 2014. *Great teachers: How to raise student learning in Latin America and the Caribbean*. World Bank Publications.
- 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 Hoyos, R., R. Estrada, and M.J. Vargas. 2016. “Do test scores predict labor market outcomes? Evidence from a national examination in Mexico.” mimeo.
- Del Valle, S. 2015. “Exige CNTE plazas... para venderlas.”
- Deserranno, E. 2017. “Financial Incentives as Signals: Experimental Evidence from the Recruitment of Village Promoters in Uganda.” Working paper, mimeo Northwestern University.
- Duflo, E., P. Dupas, and M. Kremer. 2012. “School Governance, Teacher Incentives, and Pupil-Teacher Ratios: Experimental Evidence from Kenyan Primary Schools.” NBER Working Papers No. 17939, National Bureau of Economic Research, Inc, Mar.
- Duflo, E., R. Hanna, and S.P. Rya. 2012. “Incentives work: Getting teachers to come to school.” *The American Economic Review* 102:1241–1278.
- Elizondo, C. 2011. *Por eso estamos como estamos*. Random House Mondadori.
- Este País, F. 2005. “Lo que piensan nuestros maestros: Encuesta National sobre Actitudes y Valores de Maestros y Padres de Familia.” *Revista Este País*, April, pp. .
- Evans, D.K., and A. Popova. 2016. “What really works to improve learning in developing countries? An analysis of divergent findings in systematic reviews.” *The World Bank Research Observer* 31:242–270.
- Finan, F., B.A. Olken, and R. Pande. 2017. “The personnel economics of the state.” 2:467–514.
- Glewwe, P.W., E.A. Hanushek, S.D. Humpage, and R. Ravina. 2011. “School Resources and Educational Outcomes in Developing Countries: A Review of the Literature from 1990 to 2010.” NBER Working Papers No. 17554, National Bureau of Economic Research, Inc, Oct.

- Guevara, M.d.R., and L.E. Gonzalez. 2004. "Atraer, formar y retener profesorado de calidad. Reporte sobre la situación de México." Working paper, OECD.
- Hanushek, E.A., and S.G. Rivkin. 2006. *Teacher Quality*, Elsevier, vol. 2 of *Handbook of the Economics of Education*, chap. 18. pp. 1051–1078.
- Heckman, J.J., H. Ichimura, and P.E. Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Review of Economic Studies* 64:605–54.
- Hoxby, C.M. 1996. "How Teachers' Unions Affect Education Production." *The Quarterly Journal of Economics* 111:671–718.
- Jacob, B.A., and L. Lefgren. 2008. "Can Principals Identify Effective Teachers? Evidence on Subjective Performance Evaluation in Education." *Journal of Labor Economics* 26:101–136.
- Kingdon, G., and F. Teal. 2010. "Teacher unions, teacher pay and student performance in India: A pupil fixed effects approach." *Journal of Development Economics* 91:278–288.
- Lovenheim, M.F. 2009. "The Effect of Teachers' Unions on Education Production: Evidence from Union Election Certifications in Three Midwestern States." *Journal of Labor Economics* 27:525–587.
- Mbiti, I.M. 2016. "The need for accountability in education in developing countries." *The Journal of Economic Perspectives* 30:109–132.
- Meyer, B.D. 1995. "Natural and Quasi-experiments in Economics." *Journal of Business & Economic Statistics* 13:151–61.
- Mizala, A., and H. Ñopo. 2014. "Measuring the Relative Pay of Latin American School Teachers at the turn of the 20th Century." Working Papers No. 2014-15, Peruvian Economic Association, Aug.
- Muralidharan, K., and V. Sundararaman. 2011. "Teacher performance pay: Experimental evidence from India." *Journal of political Economy* 119:39–77.
- Murnane, R.J., and A. Ganimian. 2014. "Improving educational outcomes in developing countries: Lessons from rigorous impact evaluations." *NBER Working Papers Series*, pp. .
- OECD. 2013. "Education Policy Outlook: Mexico." Working paper.
- . 2014. "Mexico Country Note - Education at a Glance 2014: OECD Indicators." Education at a glance country notes, Organization for Economic Cooperation and Development, October.
- Oyer, P., and S. Schaefer. 2011. *Personnel Economics: Hiring and Incentives*, Elsevier, vol. 4 of *Handbook of Labor Economics*, chap. 20. pp. 1769–1823.

- Rivkin, S.G., E.A. Hanushek, and J.F. Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica* 73:417–458.
- Rockoff, J.E., B.A. Jacob, T.J. Kane, and D.O. Staiger. 2011. "Can You Recognize an Effective Teacher When You Recruit One?" *Education Finance and Policy* 6:43–74.
- Rockoff, J.E., and C. Speroni. 2011. "Subjective and objective evaluations of teacher effectiveness: Evidence from New York City." *Labour Economics* 18:687–696.
- Rothstein, J. 2010. "Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement." *The Quarterly Journal of Economics* 125:175–214.
- Santibanez, L. 2008. "Reforma Educativa: El papel del SNTE." *Revista Mexicana de Investigacion Educativa* 13:419–443.
- Santibanez, L., and B. Jarillo. 2007. "Conflict and Power: the Teachers' Union and Education Quality in Mexico." *Well-being and Social Policy* 3:21–40.
- Wayne, A.J., and P. Youngs. 2003. "Teacher Characteristics and Student Achievement Gains: A Review." *Review of Educational Research* 73:89–122.
- Weissenstein, M. 2013. "Mexican Senate passes major education reform." *Associated Press*, September, pp. .

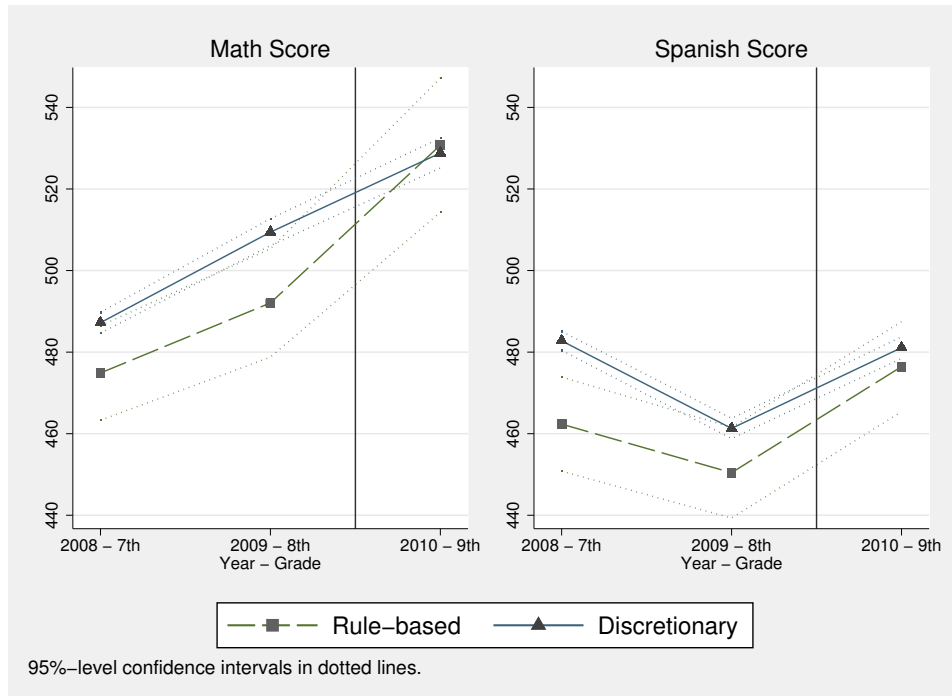
# Figures

Figure 1: Evolution of Exam Scores in Schools with 2010 Hires



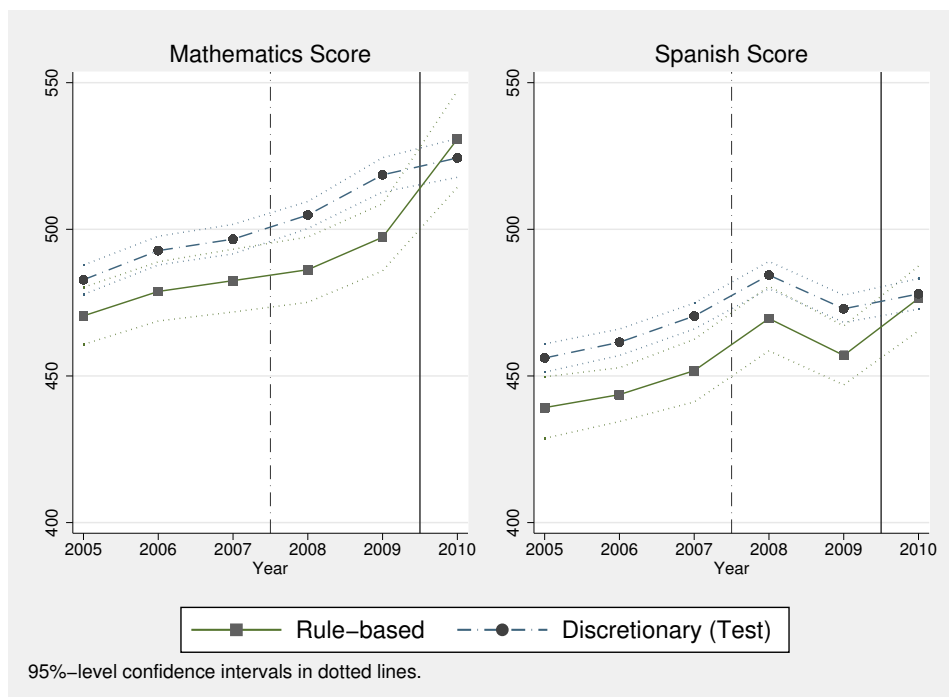
Notes: The graph plots the evolution of yearly means of school (ninth grade) scores in the Enlace exam by eventual treatment status. The vertical dotted line indicates the periods before and after test-based hiring was introduced in public schools. The vertical solid line indicates the periods before and after treatment. Sample is composed of Telesecundaria schools which received a new teacher in the year 2010. Enlace scores are standardized at the national level with mean 500 and standard deviation 100. Source: Enlace exam. Enlace results for 2005 corresponds to the academic year 2005-2006 and so on.

Figure 2: Evolution of Exam Scores in Schools with 2010 Hires: Panel of Classrooms



Notes: The graph plots the evolution of yearly means of grade scores in the Enlace exam by eventual treatment status. The vertical solid line indicates the periods before and after treatment. Sample is composed of Telesecundaria schools which received a brand-new teacher in the year 2010. Enlace scores are standardized at the national level with mean 500 and standard deviation 100. Source: Enlace exam. Enlace results for 2008 correspond to the academic year 2008-2009 and so on. Results for 2008 correspond to 7th grade classrooms and so on.

Figure 3: Evolution of Exam Scores in Schools with 2010 Hires Who Applied to the Rule-based Hiring



Notes: The graph plots the evolution of yearly means of school (ninth grade) scores in the Enlace exam by eventual treatment status. The label Discretionary(Test) refers to schools with discretionary hires who applied to the rule-based hiring (in any year and for any teaching position). The vertical dotted line indicates the periods before and after test-based hiring was introduced in public schools. The vertical solid line indicates the periods before and after treatment. Sample is composed of Telesecundaria schools which received a new teacher in the year 2010. Enlace scores are standardized at the national level with mean 500 and standard deviation 100. Source: Enlace exam. Enlace results for 2005 corresponds to the academic year 2005-2006 and so on.



## Tables

Table 1: Relative Wages of Primary and Junior Secondary Teachers

	Wage (pesos)		Wage gap (ln hourly wage)				N
	Monthly Mean	Hourly Mean	Raw		Adjusted		
			Beta	SE	Beta	SE	
Public Prim.-JS Teachers	8,280	70.7					24,599
Private Prim.-JS Teachers	7,110	56.5	-0.2470	0.0207	-0.2087	0.0195	2,487
Other Teachers	8,767	74.7	0.0083	0.0087	-0.0137	0.0084	20,215
Managers	13,285	73.3	-0.0226	0.0101	-0.0310	0.0100	13,405
Others	9,670	53.6	-0.3603	0.0062	-0.3111	0.0061	133,081
Observations	193,787						

Notes: Columns 1 and 2 report mean monthly and hourly wages of college-educated wage earners by occupation categories. Columns 3 and 4 the difference in means (and the standard error) between the base category (in the first row) and the others. Columns 5 and 6 report the coefficients (and standard errors) for the occupation categories in a Mincerian wage equation (the baseline category is in the first row). Source: National Labor Force Survey (ENOE) of the third quarter of 2006 to the second quarter of 2010. The top and bottom 2 percentiles of the wage distribution are trimmed. Sample: Individuals with college education who report positive wage earnings. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 2: Characteristics of Schools with 2010 Hires: Means in 2009 (one year before treatment)

	(1)		
	Difference	Rule-based hires	Discretionary hires
Math Score (std)	-22.19*** (5.76)	497.29	519.47
Spanish Score (std)	-19.08*** (4.62)	457.00	476.09
Flagged Exams (%)	-0.00 (0.01)	0.05	0.05
Final Enrollment (%)	-0.01 (0.01)	0.93	0.94
School size	-26.66** (10.18)	87.86	114.53
Class size	-0.72 (0.78)	19.07	19.78
Share indigenous students	0.06 (0.03)	0.16	0.10
Principal has grad school	-0.12* (0.05)	0.20	0.32
Locality Population	-7668.85 (4063.97)	6248.29	13917.14
Hours to state capital	0.73*** (0.11)	2.81	2.08
Locality Poverty Rate	0.12*** (0.02)	0.74	0.62
Share hhs electricity	-0.05*** (0.01)	0.90	0.95
Share hhs sewage	-0.10*** (0.03)	0.64	0.73
Observations	1408		

Notes: All school statistics are for the school year 2009 (one year before treatment). Column 1 reports standard errors, in parenthesis, for a t-test on the equality of means in columns 2 and 3. Enlace scores, suspected cheating and final enrollment correspond to ninth grade results. Enlace scores are standardized at the national level with mean 500 and standard deviation 100. Flagged exams report the share of exams flagged by a cheating detection algorithm run by the Secretariat of Public Education. Final enrollment is the number of Enlace takers over the number of students enrolled at the beginning of the academic year. Sample is composed of Telesecundaria schools which received a new teacher in the year 2010. Source: Enlace, school census and population census 2010. Enlace and school census results for 2009 correspond to the academic year 2009-2010. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 3: Probability of Receiving a rule-based hire (OLS)

VARIABLES	(1) Received rule-based hire(s)	(2) Share of rule-based hires
Change Math Score Lag 1	5.95e-05 (0.000182)	8.84e-05 (0.000123)
Change Spanish Score Lag 1	-6.22e-05 (0.000243)	-6.40e-05 (0.000164)
Change Math Score Lag 2	-4.37e-05 (0.000205)	1.27e-05 (0.000110)
Change Spanish Score Lag 2	-0.000146 (0.000271)	-0.000150 (0.000164)
Class size Lag 1	0.00265* (0.00141)	-0.000624 (0.000740)
Students Lag 1	-0.000171** (7.93e-05)	-1.89e-05 (3.33e-05)
Share indigenous students Lag 1	-0.0851** (0.0345)	-0.0405*** (0.0150)
Principal has grad school	-0.0100 (0.0116)	-0.00790** (0.00367)
Locality Poverty Rate	0.111* (0.0602)	0.0553* (0.0316)
Hours to state capital	0.0319*** (0.0120)	0.0226** (0.00937)
Locality Population	2.29e-07 (1.58e-07)	1.43e-07*** (5.13e-08)
Share hhs electricity	-0.285** (0.130)	-0.247** (0.121)
Share hhs sewage	-0.00608 (0.0341)	-0.000322 (0.0208)
Observations	1,395	1,395
R-squared	0.161	0.163
State Fixed Effects	Yes	Yes
F statistic Ho Var 1-4=0	0.267	0.547
Prob > F	0.899	0.701

Notes: Enlace scores correspond to ninth grade results and are standardized at the national level with mean 500 and standard deviation 100. Sample is composed of Telesecundaria schools which received a new teacher in the year 2010. Source: Enlace, school census and population census 2010. Robust standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 4: Difference-in-Differences: Pretreatment Trends in Outcomes

VARIABLES	(1) Final Enrollment (%)	(2) Math Score (std)	(3) Spanish Score (std)
2006 X Share rule-based hires	0.0133 (0.0310)	-3.726 (9.864)	-10.04 (10.09)
2007 X Share rule-based hires	-0.00979 (0.0394)	-8.248 (13.93)	-6.066 (15.15)
2008 X Share rule-based hires	0.102*** (0.0386)	-16.53 (13.79)	-9.823 (13.54)
2009 X Share rule-based hires	-0.00182 (0.0337)	-2.338 (15.08)	-2.356 (12.77)
Observations	6,867		
R-squared	0.072	0.127	0.107
Number of id	1,396		
F statistic Ho Var 1-4=0	3.111	0.683	0.382
Prob > F	0.0146	0.604	0.822

Notes: Results are for ninth grade outcomes. Final enrollment (column 1) is the number of Enlace takers over the number of students enrolled at the beginning of the academic year. Enlace scores (columns 2 and 3) are standardized at the national level with mean 500 and standard deviation 100. Regressions include class size, school size, the share of indigenous students, an indicator for principal's attendance of graduate school, a vector of interactions between year and state dummies, and school fixed effects. Source: Enlace 2006-2011, school census data 2006-2011 and Registro Maestros 2010-2011. Enlace and school census results for 2005 correspond 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: Results

VARIABLES	(1) Final Enrollment (%)	(2) Math Score (std)	(3) Spanish Score (std)
Share rule-based hires	-0.0130 (0.0304)	52.64*** (15.94)	31.63*** (11.20)
Observations	8,262		
R-squared	0.068	0.162	0.108
Number of id	1,397		
Mean Control	0.938	528.8	481.1

Notes: Results are for ninth grade outcomes. Final enrollment (column 1) is the number of Enlace takers over the number of students enrolled at the beginning of the academic year. Enlace scores (columns 2 and 3) are standardized at the national level with mean 500 and standard deviation 100. Regressions include the number of 2010 hires in the school, class size, school size, the share of indigenous students, an indicator for principal's attendance of graduate school, a vector of interactions between year and state dummies, and school fixed effects. Source: Enlace 2006-2011, school census data 2006-2011 and Registro Maestros 2010-2011. Enlace and school census results for 2005 correspond 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: Other Outcomes

VARIABLES	(1) Share Indigenous Students	(2) Share Repeaters	(3) Principal has grad school	(4) Class size
Share rule-based hires	0.0456 (0.0340)	-0.00211 (0.00692)	-0.0464* (0.0281)	-0.0463 (0.382)
Observations	8,270			
R-squared	0.048	0.031	0.050	0.358
Number of id	1,397			

Notes: Outcomes are measured at the beginning of the academic year. Regressions include the number of 2010 hires in the school, school size, a vector of interactions between year and state dummies, and school fixed effects. 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 7: Difference-in-Differences: Robustness Checks

VARIABLES	(1) Final Enrollment (%)	(2) Math Score (std)	(3) Spanish Score (std)
A. School-Specific Linear Time Trends			
Share rule-based hires	-0.0130 (0.0304)	52.68*** (15.94)	31.67*** (11.20)
Observations	8,262		
R-squared	0.064	0.044	0.062
Number of id	1,397		
B. Binary Treatment (Small Schools)			
Rule-based hire(s)	0.00177 (0.0160)	35.99*** (10.09)	19.87*** (7.004)
Observations	3,810		
R-squared	0.081	0.183	0.119
Number of id	646		
Mean Control	0.945	521.2	476.2
C. Abadie Semiparametric DID Estimator (Small schools)			
Rule-based hire(s)	0.0243 (0.0214)	27.10*** (10.15)	15.28* (7.860)
Observations	414		

Notes: Results are for ninth grade outcomes. Final enrollment (column 1) is the number of Enlace takers over the number of students enrolled at the beginning of the academic year. Enlace scores (columns 2 and 3) are standardized at the national level with mean 500 and standard deviation 100. Panel A: Outcomes are de-trended using school-specific linear time trends. Panels A and B: Regressions include the number of 2010 hires in the school, school size, a vector of interactions between year and state dummies, and school fixed effects. Panels B and C: Small schools are those with 3 or less classrooms in 2010. Panel C reports results of Abadie's (2005) semi-parametric difference-in-differences estimator. The propensity score is constructed using a regression in which treatment status is regressed on the covariates listed in Table 2 plus two lags of the change in final enrollment and exam cheating. Third-order polynomials are included for all continuous variables. Estimation of results is restricted to observations in the common support of the propensity score. 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 8: Difference-in-Differences: Panel of Classrooms

VARIABLES	(1) Final Enrollment (%)	(2) Math Score (std)	(3) Spanish Score (std)
A. Pre-treatment Outcomes			
Share rule-based hires	0.000598 (0.0341)	-16.68 (12.94)	5.160 (9.937)
Observations	2,765		
R-squared	0.057	0.155	0.252
Number of id	1,396		
Mean Control in 2009	0.929	509.4	461.3
B. Results			
Share rule-based hires	0.0186 (0.0344)	43.66*** (14.77)	21.22** (10.06)
Observations	4,160		
R-squared	0.044	0.212	0.157
Number of id	1,397		
Mean Control in 2010	0.938	528.8	481.1

Notes: Final enrollment (column 1) is the number of Enlace takers over the number of students enrolled at the beginning of the academic year. Enlace scores (columns 2 and 3) are standardized at the national level with mean 500 and standard deviation 100. Regressions include the number of 2010 hires in the school, class size, school size, the share of indigenous students, an indicator for principal's attendance of graduate school, a vector of interactions between year and state dummies, and school fixed effects. Source: Enlace 2009-2011, school census data 2009-2011 and Registro Maestros 2010-2011. Enlace and school census results for 2008 corresponds to the academic year 2008-2009 and so on. Standard errors in parentheses are clustered at the school level. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1



Table 9: Difference-in-Differences: Teachers Who Applied to the Rule-based Hiring

VARIABLES	(1)	(2)	(3)	(4)
	Math Score (std)	Spanish Score (std)	Math Score (std)	Spanish Score (std)
Share rule-based hires	55.50*** (16.57)	36.06*** (12.16)	51.64*** (16.46)	34.14*** (11.37)
Hire university GPA			-1.729 (3.376)	-3.172 (2.806)
Female Hires			6.009 (6.241)	1.314 (5.016)
Hire Age			0.230 (0.724)	0.200 (0.507)
Hire public sector experience			-6.677 (6.328)	-6.065 (5.102)
Hire private sector experience			-2.164 (9.935)	-10.15 (6.660)
Hire graduation year			-0.707 (1.097)	-1.427 (0.907)
Observations	2,426		408	408
R-squared	0.197	0.149	0.050	0.076
Number of id	415			
Mean Control in 2010	524.3	478		

Notes: Results are for ninth grade outcomes. Enlace scores (columns 1 to 4) are standardized at the national level with mean 500 and standard deviation 100. Regressions in columns 1 and 2 include the number of 2010 hires in the school, class size, school size, the share of indigenous students, an indicator for principal's attendance of graduate school, a vector of interactions between year and state dummies, and school fixed effects. Regressions in columns 3 and 4 substitute levels for changes between the pretreatment and posttreatment periods in the variables shown in columns 1 and 2. The university GPA is standardized with mean 0 and SD 1. 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 10: Teacher Characteristics – Test Applicants: Means

	(1) Difference	Test Teacher	Discretionary Teacher
Age	-2.14*** (0.60)	25.93	28.06
Female (%)	-0.03 (0.06)	0.55	0.58
Public sector experience (%)	-0.31*** (0.05)	0.20	0.50
Private sector experience (%)	0.03 (0.03)	0.11	0.09
College graduation year	1.31*** (0.39)	2007.66	2006.34
College GPA (std)	0.51*** (0.11)	0.37	-0.14
Observations		96	456

Notes: Column 1 reports differences in means and standard errors, in parenthesis, for a t-test on the equality of means in columns 2 and 3. The university GPA is standardized with mean 0 and SD 1. Sample: 2010 hires who applied to the rule-based recruitment. Source: National Register of Teachers 2010-2011. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

# Online Appendix

## A Census of Teachers

I benefit from extensive data of school personnel compiled as the result of a recent mandate of the Mexican Federal Congress. The data comprises the quarterly payrolls of public elementary schools from the second quarter of 2010 (the last of the 2009-2010 academic year) to the second quarter of 2011. The Secretariat of Public Education 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 complete information about hiring, education profile or assigned classrooms.

I do not directly observe in the data which are the new teachers hired since 2008—when the rule-based examination was implemented. However, I use the 2009 and 2010 censuses to identify the 2010 hires. Then, I match these observations to the list of rule-based hires, which is available in the dataset. Hence, I focus my analysis on the (24) states that opened vacancies for the Telesecundaria system in 2010 using rule-based hiring.<sup>23</sup>

I identify the 2010 cohort of hires comparing the census of Telesecundaria's personnel for the second quarter of 2010 (the last of the 2009-2010 school year) 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 new personnel in the 2010 school year. I drop observations from the state of Guerrero because there are large missing values for school assignment in the 2nd-quarter of 2010. I also drop four states that report relatively few personnel in the second quarter and hence have a high, and likely unreliable, ratio of new to total personnel in the 2010 school year.<sup>24</sup>

The secretariat's dataset includes a module with the list of the 2009 and 2010 rule-based applicants (573 and 492 teachers, respectively, in 19 states). I merge this module with the main dataset using the national population number.<sup>25</sup> I merge 76.8 percent of the 2009 and 89.4 percent of the 2010 rule-based hires to specific schools.<sup>26</sup> I obtain the test scores of all the rule-based hires in the sample by merging the data by full name with the official results available on the website of the secretariat.

---

<sup>23</sup>2 states (Michoacan and Oaxaca) do not participate at all in the test-based examination and 6 states did not open to test-based recruitment any vacancy at the Telesecundaria system in 2010 (Baja California Sur, Colima, Nayarit, Queretaro, Sonora and Zacatecas).

<sup>24</sup>These states are Baja California, Tabasco, Veracruz and Yucatan.

<sup>25</sup>The taxpayer identification number is not available in the rule-based hiring module. Around 5 percent of the observations in the main personnel module have missing information for the national population number.

<sup>26</sup>I add a smaller number of observations from applicants who were put in waiting list for hiring (16 in 2009 and 68 in 2009). Not surprisingly, the matching rate for this group is considerably lower (50 percent for 2009 and 15 percent for 2010)

I find that 15.5 percent of the matched individuals hired in the 2010 test examination as new teachers were already on a Telesecundaria's payroll in the 2009 school year. At the extreme, 15 of the 16 individuals hired in the state of Nuevo Leon fall into this case. This evidence suggests that some incumbent teachers, maybe hired under temporary contracts, were allowed to participate in the rule-based examination for new teachers. As I am interested in studying the performance of teachers who are effectively new, I drop 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 and the Federal District.

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

I collapse the teachers dataset at the school-year level and merge it into the panel with school results and characteristics. I merge to this panel a list with the schools where the 2008 rule-based hires were initially assigned (the file does not include the teachers' population or tax identification numbers).

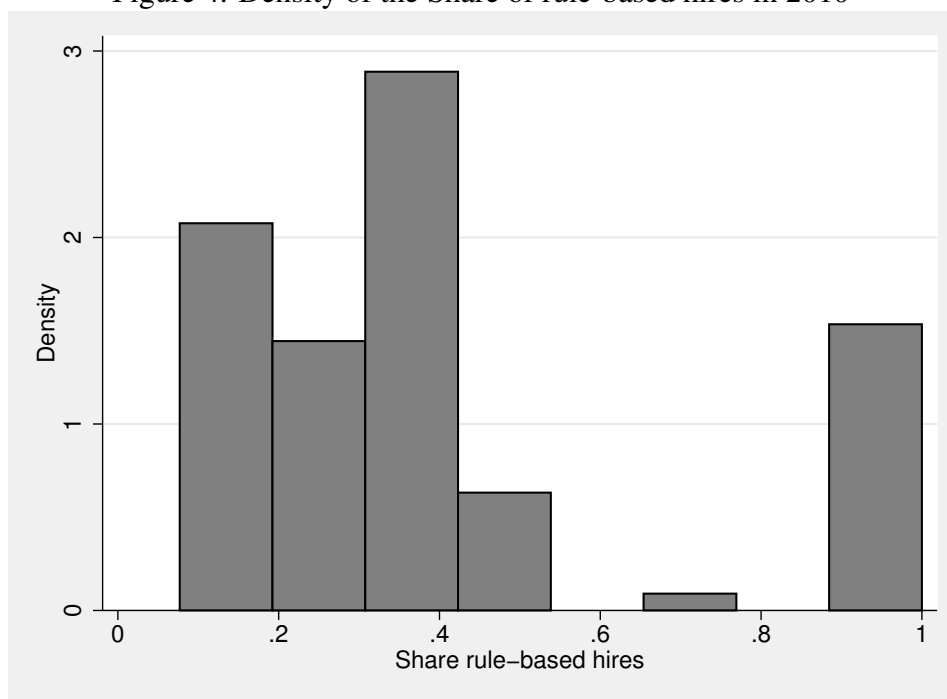
I restrict the dataset to schools that have never received a 2009 or 2008 rule-based hire, did not receive both a new rule-based hire and a new regular teacher in 2010 and have at least 4 years of Enlace results. I finally cut schools with at least one year-to-year change in their school score larger than 2 SD, which roughly corresponds to the top and bottom one percentile of score changes. I obtain a dataset with 1,427 schools in 15 states after these restrictions. 6.7 percent of schools in the sample received at least one rule-based hire in the 2010 school year.<sup>27</sup>

---

<sup>27</sup>I exclude all observations from the states of Morelos and Campeche because there is no left schools with only rule-based hires after these restrictions.

## B Distribution of the Share of Rule-Based Hires

Figure 4: Density of the Share of rule-based hires in 2010



Notes: The graph shows the distribution of the share of rule-based hires among the total number of teachers in Telesecundaria schools that received new rule-based hires in the year 2010. Source: National Register of Teachers and school census 2010-2011.

## C Distributional Effects of Rule-Based Hires

Table 11: Difference-in-Differences: Results – 25th and 75th Percentiles

VARIABLES	(1) Math Score pct 25	(2) Math Score pct 75	(3) Spanish Score pct 25	(4) Spanish Score pct 75
Share rule-based hires	55.75*** (16.90)	51.81*** (19.30)	29.86** (13.35)	32.46*** (11.46)
Observations	8,262			
R-squared	0.131	0.170	0.079	0.099
Number of id	1,397			
Mean Control	466.5	589.5	422.7	537.3

Notes: Outcomes in columns 1 and 2 (3 and 4) are the 25th and 75th percentiles of the school Enlace score distribution in mathematics (Spanish). Enlace scores are standardized at the national level with mean 500 and standard deviation 100. Regressions include the number of 2010 hires in the school, class size, school size, the share of indigenous students, an indicator for principal's attendance of graduate school, a vector of interactions between year and state dummies, and school fixed effects. Results are for ninth grade outcomes. Source: Enlace 2006-2011, school census data 2006-2011 and National Register of Teachers 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

## D Pre-Trends: Teachers Who Applied to the Rule-Based Hiring

Table 12: Difference-in-Differences: Teachers Who Applied to the Rule-based Hiring

VARIABLES	(1)	(2)
	Math Score (std)	Spanish Score (std)
2006 X Share rule-based hires	-8.690 (10.85)	-14.02 (10.54)
2007 X Share rule-based hires	-1.460 (15.29)	-3.456 (16.31)
2008 X Share rule-based hires	-21.62 (15.16)	-20.51 (14.67)
2009 X Share rule-based hires	-10.41 (16.75)	-5.930 (13.38)
Observations	2,011	
R-squared	0.148	0.139
Number of id	414	
F statistic Ho Var 1-4=0	0.842	0.991
Prob > F	0.499	0.412

Notes: Results are for ninth grade outcomes. Enlace scores (columns 1 and 2) are standardized at the national level with mean 500 and standard deviation 100. Regressions include class size, school size, the share of indigenous students, an indicator for principal's attendance of graduate school, a vector of interactions between year and state dummies, and school fixed effects. 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