Inducing Teacher Retention in Remote Locations: Evidence from Peru*

Max Perez Leon†

October 23, 2019

Abstract

In this paper, I estimate the extent to which cash bonuses retain teachers in remote locations of Peru. Many governments utilize financial incentives to fill education and health care positions in undesirable locations. The Peruvian public education system is an ideal ground to study the effectiveness of such policies. As part of a larger education reform in 2015, the Peruvian government implemented a salary bonus scheme favoring regions far from urban centers. Because the bonuses are added to base salaries that are uniform across regions, teachers face wage offers that change discontinuously across space. These discontinuities allow the use of well-developed program evaluation methods. Specifically, I apply a matching estimator to a boundary discontinuity design to recover the average treatment effect on the treated at the (unknown) boundary of the bonuses. I find that the bonuses offered by the Peruvian government have a heterogeneous effect: temporary teachers react to the bonuses with a sizeable increase in their retention rate, while permanent teachers with tenure respond very little. One explanation is that permanent teachers choosing these remote locations have a much higher retention rate to begin with.

*I am indebted to Sam Kortum, Lorenzo Caliendo, John Eric Humphries for their invaluable advice and encouragement and to Joseph Altonji, Barbara Biasi, Rafael Dix-Carneiro, Cecilia Fieler, Yuichi Kitamura, Francis Kramarz, Yusuke Narita, Michael Peters and Edward Vytlacil for their insightful comments. I also thank Antoine Arnaud, Julian Aramburu, Paula Calvo, Lucas Zavala and the participants of the Yale Trade and Labor Lunch. Alexander Quispe, Lesly Flores, Thyara Sanchez, Rodrigo Rivarola provided excellent fieldwork assistance. This research was funded by the Department of Economics at Yale University.

† Yale University, max.perezleonquinoso@yale.edu.

1 Introduction

How costly is it for governments to curtail the "brain drain" of public workers leaving undesirable locations? The education and health sector are the two sectors where this issue is most tangible and hits both developed and developing countries. In this paper, I study a policy implemented by the Peruvian government that aims to alleviate the unequal distribution of teachers in the education sector. As part of a larger education reform implemented in 2015, this policy established a salary bonus schedule targeted at schools in remote and potentially undesirable locations. In order for teachers to receive a location-specific bonus, the only condition is to work at the schools selected by the government as remote. The bonus schedule ultimately intends to retain teachers in remote locations and, consequently, reduce the unequal distribution of teachers across space. In this paper, I investigate to what extent this policy altered teacher retention and teacher composition in remote schools. I find a substantial increase in retention, but this effect is limited to teachers with a particular type of contract.

The Peruvian education system provides a unique setting in which to explore the impacts of location-specific bonus schedules. First, the bonuses are significant, varying from 5% to 25% of teachers' average salary. In the aggregate, the bonus schedule represents approximately 4.7% of the total payroll. Second, teachers face the same institutional setup: an eight-step schedule for base salaries, nationwide reallocation rules, and job tasks that are similar across the whole system. Third, the legal framework allows teachers to move freely across Peru without ever leaving the public education system. The ubiquity and magnitude of the public education system, in principle, allows teachers to move to any location in Peru. Finally, teachers can only be hired under two types of contracts: permanent and temporary. This simple legal structure turns out to be crucial to gain insight on the efficacy of the monetary incentives because of the differential effect between the two contracts.

In this paper, my goal is to quantify the effect of location-specific monetary incentives on teacher retention, but more importantly show how the effects are heterogeneous across different bonuses and how incentives differ considerably depending on the teacher's contract. These estimates are obtained using teacher level data. Then, I confirm the robustness of my estimates by estimating retention rates at the school level. I also show that bonuses had no effect on teacher composition within schools in term of the distribution of contracts and qualifications. Finally, I re-examine the estimates under a discrete choice model, I provide some insights on why we observe such distinct heterogeneity in teacher retention rates across contracts.

To investigate how monetary incentives affect teacher retention and school's teacher composition I exploit geographical boundaries induced by the location-specific bonus schedule established in 2015. Part of the bonus schedule is defined at the school level by cutoffs in terms of two variables: time to provincial capital and population size of the town surrounding the school. These cutoffs imply a discontinuity which has been previously studied by Alva, Bobba, Leon, Neilson, and Nieddu (2017) and Castro and Esposito (2018) under a regression discontinuity design (RDD). Nevertheless, the entire bonus schedule is not defined by clear cutoffs. To accommodate the entire bonus schedule, I propose a new approach under a **boundary discontinuity design (BDD)** (Bayer, Ferreira, and McMillan, 2007; Black, 1999; Dell, 2010; Keele and Titiunik, 2015).[1] The BDD is the extension of the regression discontinuity design to spatial settings. Methodologically, I depart from the previous approaches and apply matching estimators (Abadie and Imbens, 2011; Heckman, Ichimura, and Todd, 1997) to the BDD. Intuitively, I compare teachers or schools that receive a bonus to geographically close counterparts that do not get the bonus. I can then compare two observations that face the same local amenities and local markets but that are differently affected by the policy. An appealing feature of this approach is that it can easily incorporate more complex geographic information such as topographical data or travel times, which are not incorporated naturally in a regression framework.

The boundary discontinuity design suits the specific bonus schedule setup and it can be easily paired with the matching estimator approach. The combination of the BDD with matching estimators provides several benefits. First, the location-specific bonus schedule induces an "invisible" geographic boundary between schools. Even though the boundary is unknown, I can discern the treated from control group[2]. The fact that two adjacent schools are treated differently implies that a boundary runs between them. The matching approach allows me to "recover" the boundary by focusing on pairs of schools with different treatment status that are close to each other. Second, the assumption that unobserved variables change continuously at the discontinuity (e.g., unobserved amenities, labor markets) fits naturally in the geographical space. Finally, the continuity assumption has a consistent interpretation under a discrete choice model, which provides a framework to evaluate the results.

To study the heterogeneous effect of the bonus on teachers I group teachers along two dimensions: contract type and qualification. Teachers can be hired under as permanent or

[1] I use the "boundary discontinuity design" term coined by Bayer, Ferreira, and McMillan (2007), which is also called Geographic Regression Discontinuity by Keele and Titiunik (2015).

[2] This is not be confused with a fuzzy boundary discontinuity design. In that design, the identity of the treatment group is not observed.

temporary contract.[3] Qualification is defined by the possession of an educational college degree. While all permanent teachers are qualified, temporary teachers can be qualified or unqualified. **Permanent teachers** have tenure and have the first priority to move between schools. **Qualified temporary teachers** can then cover unfilled positions and if some are still available, they can be filled by **unqualified temporary teachers**. All three types of teachers are eligible to receive the location-specific bonuses.

In terms of teacher retention, my results show that the bonus schedule has a heterogeneous effect across the three types of teachers. The current bonus schedule does not modify permanent teachers' retention rate. This result is robust to different estimation procedures and across different type of bonuses. The results are interesting when compared to qualified temporary teachers. They show the strongest effect for the bonuses. For qualified temporary teachers I find a 4-8 percentage point increase in retention rates when faced with a bonus that represents 25% of the average salary.[4] Nevertheless, this effect is not linear. For this group of teachers, I find no effects on retention when faced with 17% salary bonus in the coca-growing region of Peru and at the same time I find an increase between 5-10 percentage points when faced with a 5% salary bonus. Finally, unqualified teachers behavior is unaltered by the presence of the bonuses.

National and state governments are required to supply a public service, most noticeably education and health, in isolated and potentially undesirable locations.[5] Educated workers, including teachers and health professionals, tend to move from non-urban towards urban locations. This would not be problematic if the whole population moved at the same pace, but public servants usually move to urban locations at a higher speed than public service demand (e.g. students). This creates a problem for governments that still need to provide services to the population that remains in remote locations. To tackle this problem, governments design personnel policies that try to counter the preferences that drive their workforce towards cities. The main instrument used by policymakers are monetary incentives and the problem has been studied extensively.

This paper lies in the intersection of several literatures. First, this paper contributes to the economics of education literature that studies the effect of location-specific bonuses on teacher retention and location decisions (Alva, Bobba, Leon, Neilson, and Nieddu, 2017; Castro and Esposito, 2018; Clotfelter, Glennie, Ladd, and Vigdor, 2008). I add to this liter-

[3] Permanent contracts are known as indefinite contracts and provide tenure. Teachers can only access permanent contracts after approving a national entrance exams. The most recent entrance exams have taken place in 2009, 2010, 2011 and 2015 (Majerowicz and Montero, 2018).

[4] The base retention rate is between 16 to 25%.

[5] In the US, teacher shortage is a recurring issue across states which disproportionately affects rural school districts.

ature by showing the heterogeneous effect of the monetary incentives across teacher contracts and across type of bonuses. The literature has focused on a specific bonus out of the whole bonus schedule. I provide new results showing the heterogeneous effects across different bonus types and contracts within the same institutional context. More generally, this paper is related to the study of teacher pay on teacher turnover rates that has been extensively examined by Cabrera and Webbink (2018); Dolton and Klaauw (1999); Dolton and von der Klaauw (1995); Gilpin (2011); Hanushek, Kain, and Rivkin (2004); Imazeki (2005); Murnane and Olsen (1990) and wage differential on school selection by Biasi (2018).

This paper is also related to recent literature that tries to uncover migration or location decision such as Bayer, Ferreira, and McMillan (2007); Bayer, McMillan, Murphy, and Timmins (2016); Bishop and Murphy (2011) and Diamond (2016). In the same direction, Kennan and Walker (2011) and Albouy (2009) provide explanations on how monetary incentives such as expected income and taxes shape the migration/distribution of workers across space.

Methodologically, this paper contributes to the boundary discontinuity design literature by providing another approach to a setup where the treatment status is known but the boundary is not. The BDD approach has been applied to study policies that change discontinuously across space such as the case of minimum wage increases in Card and Krueger (1994) and Dube, Lester, and Reich (2010) or right-to-work laws in Holmes (1998). Black (1999) and Bayer, Ferreira, and McMillan (2007) applied a similar design using school district borders to study neighbor preferences; and Dell (2010) used the border of mining *mita* to study long-run effects of colonial institutions in Peru. With the increasing availability of geographic information, this approach has been used in a variety of settings in economics and political science (Keele and Titiunik, 2015; Posner, 2004; Shapiro, 2018).

This paper also contributes to the matching estimator literature by creating a bridge between the matching estimator and the setting to which it is applied. To estimate the average treatment effect of the bonus at the boundary I utilize matching estimators. This paper utilizes estimators developed by Heckman, Ichimura, and Todd (1998) (implemented by Jann, 2017) and Abadie and Imbens (2011). The use of a matching estimator in a BDD is influenced by Gelman and Imbens (2018) recommendation to employ a non-parametric approach in regression discontinuity designs instead of high-order polynomials.[6] Keele and Titiunik (2015) follow this implementation in a spatial setting. To complement the

[6] The use of polynomials in a spatial setting is examined by Dell (2010).

matching estimator approach I also conduct robustness check using a classic regression approach with location-specific fixed-effects similar to Black (1999).

The rest of the paper is organized as follows. Section 2 provides details on the institutional framework and the bonus schedule, as well as the data. Sections 3 presents the matching approach to boundary discontinuity design and Section 4 shows the results. Section 5 connects teacher retention to a discrete choice model to provide an explanation of the heterogeneous results. Section 7 concludes.

2 Institutional Setting and Data

In this section, I describe the institutional setting and document key facts of the Peruvian public education system. First, I describe how public education teachers are managed. Second, I describe the implementation of the location-specific bonuses. The bonus implementation divides schools into different groups and induces geographic boundaries that lead wages to change discontinuously across space. Finally, I describe the profile of schools and teachers in the different groups generated by the bonuses. As expected, teachers in more remote schools face a more challenging environment.

2.1 Peruvian public education system

The Peruvian public education system is in charge of serving almost six millions students in three education levels: Kindergarten, Elementary and High School.[7] The government establishes the personnel policies such that they can serve this sizable student body.[8] Public schools are geographically grouped into 228 school districts.[9] Peruvian schools districts are in charge of the local management of personnel and implement many of the regulations defined at the national level. In contrast to the U.S., school districts in Peru only administer the nationwide policies and do not have the ability to enact their own policies. School districts are grouped in 26 departments (*departamentos*)[10] which are the closest analog to a state in the U.S.

The education system has two contract types for teachers: permanent and temporary. Permanent teachers have tenure with full labor stability. Access to a permanent contract

[7] Kindergarten has two grades, Elementary has six grades and High Schools has five grades.
[8] The Ministry of Education is the institution in charge of public education.
[9] School district's legal name is Local Education Management Unit (*Unidad de Gestión Educativa Local* in Spanish.).
[10] Formally there are 24 departments and two provinces with special regimens (Lima and Callao).

position is accomplished through national entrance exams.[11] Permanent teachers have full flexibility to move within Peru as long as they find an available opening. Once permanent teachers move they are required to stay in the new location for at least two years or they lose tenure. Since permanent teachers have a high degree of freedom to move across schools, the system has to fill permanent teacher shortages with temporary teachers. Temporary teachers are hired to fill teaching positions when there are not enough permanent teachers. It is worth noting that temporary teachers cannot become permanent unless they pass the national entrance exams.

Both permanent and temporary teachers are required to be qualified (i.e., hold an educational college degree.) A requisite to become a permanent teacher is to be qualified, but the qualification requirement can be waived for temporary teacher when vacant openings have not been filled. The contract and qualification characteristic define three distinct groups of teachers: Permanent teachers, qualified temporary teachers and unqualified temporary teachers. All three types of teachers are eligible to receive the location-specific bonuses.

In Figure 1 I show recent trends in number of teachers for all three types. Permanent teachers represent the largest share of teachers in the public education system. With more than two hundred thousand teachers, they represent 58% of all teachers in 2017. The number of qualified temporary teachers has increased considerably in the last five years. The system went from 45 thousand qualified temporary teachers in 2013 to 138 thousand in 2017. There has also been an increase in the hiring of unqualified teachers, reaching a high of thirteen thousand in 2017.

2.2 Monthly Salaries and Location-Specific Bonuses

Public teachers' monthly salaries were stagnant for a long period of time but have been rapidly improving as part of the education reform implemented in 2015. Every teacher receives a base salary that is independent of school characteristics. The education reform set an eight step schedule for base salaries, with the base salary in each step tied to the lowest step, as shown in the third column of Table 1. In 2017, the lowest base salary starts at S/ 1,780.50, which is more than double the national minimum monthly salary of S/850.[12] Only permanent teachers are allowed to moved up the salary schedule, while all

[11] The most recent entrance exams have taken place in 2009, 2010, 2011 and 2015 (Majerowicz and Montero, 2018).

[12] The Peruvian monetary currency is the "sol" and is denoted by S/. The exchange rate has been around S/3.3 for $1 for the last four years.

temporary teachers are placed in the lowest step.[13] As of 2017, the highest step reached by teachers is the sixth step which is associated with a 50% increase over the first step base salary.

As part of the education reform, the government implemented location-specific bonuses focused on non-urban schools, frontier schools and the main coca-growing region (VRAEM region). The main intention of these bonuses is to provide financial support to teachers working in remote locations. The full salary is a combination of these location-specific bonuses and the base salary.[14]

Non-urban schools are classified by the government as Isolated, Rural, and Suburban.[15] The classification depends on the distance in minutes from the nearest provincial capital and the population size of the surrounding town.[16] If a school is more than 120 minutes from the provincial capital and if the surrounding town has less than 500 inhabitants, then the school is classified as Isolated. Teachers working at Isolated schools receive a S/500 monthly bonus, which represents 28.1% of the lowest base salary. All the other non-urban classifications are shown in Figure 2 and complete location-specific bonus schedule is described in Table 2. All three non-urban school types are spread across Peru while Urban schools are mostly concentrated along the coast. The spatial distribution of schools across Peru in terms of their urban/non-urban classification is shown in Figure 19 in Appendix D.

The other two bonuses have a narrower geographical scope. The first one is the VRAEM bonus which is applied to schools in the VRAEM region, the main coca-growing region in Peru. The bonus is S/ 300 and is approximately 17% of the first step base salary. The second one is the Frontier bonus for teachers working close to frontiers with Bolivia, Brazil, Chile, Colombia and Ecuador. This is a smaller bonus and is equivalent to 5.6% of the lowest base salary. In Figure 18 in Appendix D, I show the spatial distribution of both classifications.

2.3 Distribution of Resources

The implementation of the bonus schedule is partially motivated by the unequal distribution of resources that teachers and students face across space. To have a sense of the unequal situation of schools, I provide some summary statistics for Elementary schools

[13] Permanent teachers can move up the ladder through nationwide evaluations.
[14] Teachers can also take additional tasks that have a corresponding bonus.
[15] Isolated, rural and suburban are legally classified as rural 1, rural 2 and rural 3.
[16] Peru is divided in 196 provinces (*provincias*) which are usually are used to construct school districts.

in 2016 in Table 3. The statistics are shown for each classification in the bonus schedule, as well as urban schools in cities with less than 50 thousand inhabitants.

Consider first the environment outside of schools. By construction, the town size decreases as we move towards more remote locations. In terms of the three main public services (electricity, water and sanitation), we can clearly observe the unfavorable environments students and teachers face in remote locations.

In terms of student outcomes, the standardized reading and math scores of students decrease as we depart from Urban locations. The results are significant if we notice that Isolated schools are around a standard deviation away from Urban schools in both subject areas. The percentage of students who repeat a grade is between 9 and 11%, but there is no clear pattern across schools.

If we look at the contract type composition between schools, the results show an expected pattern. As we move from Urban to Isolate schools, the fraction of permanent teachers decrease from 78% on average to 48% on average. In the same manner, the fraction of unqualified temporary teachers increases from 2% to 5%.

With the education reform the government increased its efforts in filling more teaching positions, particularly in non-urban locations, by hiring more temporary teachers. The rise in temporary teacher hiring reached all geographic locations. I show the aggregate distribution of Temporary teachers across different school classifications in Table 26 of Appendix G. Consistent with previous statistics, the more remote locations always have relatively more temporary teachers. The share of temporary teachers within each geographical location decreases as we move towards remote locations: from 27% in Urban locations to 63% in Isolated locations in 2017.

Since the personnel system is tailored to resolve shortages, the number of students is the main force that drives the number of teachers (Permanent + Temporary) at each school. As can be seen in Table 3, the student to teacher ratio is constant across different types of schools. On average, schools do not deviate from the student to teacher ratio and teacher shortage is not the main issue. Instead, the main issue seems to be the distribution of teachers.

The personnel system gives priority to permanent teachers to move across space, lower priority to qualified temporary teachers and lowest to unqualified temporary teachers. This obviously shapes the geographic distribution of human capital, resulting in an uneven distribution of permanent, qualified temporary, and unqualified temporary teachers. The retention rates by teacher type in Table 3 clearly show this pattern across different school classifications. The retention rate of permanent teacher decreases from 95% in Urban schools to 89% in Isolated schools. The retention rate of qualified temporary teachers does not present this monotonic behavior in terms of isolation and fluctuates around 25%, while unqualified temporary teachers show very low retention rates. The priority order induces permanent teachers to work at urban locations, the qualified temporary teachers work at more isolated schools and unqualified temporary teachers work at the most isolated locations. The manner by which the distribution is shaped has a strong life-cycle component that is explored in Appendix B.

If we focus on temporary teachers, the previous empirical facts imply that the share of qualified temporary teachers (in term of all temporary teachers) decreases as we move to more isolated schools. In Table 4, I show the share of qualified teachers (in term of the total number of temporary teachers) by geographical regions. In 2017, the percentage of qualified temporary teachers at Isolated schools is 84%. This percentage increases to 89% in Rural, 93% in Suburban and 95% in Urban. This structure is persistent and has worsened in recent years, to some extent driven by the efforts to fill teaching positions without increasing the number of permanent teachers. To be sure that this is not driven by certain school districts with a higher level of rural schools, I carry out a regression analysis controlling for school districts fixed effects. The results are shown in Table 25 of the Appendix.

2.4 Teacher Composition and Students Outcomes

To show the importance of teacher composition, I first show the relation between the share of Qualified Temporary and Unqualified Temporary and student outcomes. I regress standardized reading and math test scores of second grade students on the share of Qualified and Unqualified Temporary teachers at the school-year level. I include school fixed-effects as well as time fixed-effects. The results show that teacher composition is related to lower scores, but this relation does not hold across different school types. The results are shown in Table 5 and 6. The first column shows the result for all schools pooled together and the other six columns focus on schools depending on their non-urban classification, VRAEM schools and Frontier schools.

The presence of unqualified temporary teachers is strongly associated to low reading and math scores. This result is especially true in Isolated Schools and to a lesser degree in VRAEM and Frontier schools, where results are less precise. To understand the magnitude of the effect, any given year 5 percent of Isolated schools experience a 33 percentage point increase in the fraction of unqualified temporary teachers. An increment of such magnitude is associated to a decrease in reading and math scores equivalent to .12 and .65 standard deviations, respectively. Qualified temporary teachers appear to not be strongly related to higher or lower scores. The only exception to these relations appears in Urban schools, where the fraction of qualified temporary teachers is correlated with lower reading and math scores.

Even though the results do not provide conclusive evidence of a causal relation between teacher composition and student outcomes, they do show that unqualified temporary teachers can potentially contribute to lower academic outcomes in remote schools.

2.5 Data

I employ two datasets in this paper. First, I construct a school panel dataset that provides information on schools as well as the town surrounding it. This data is crucial to calculate proximity between schools and teachers and allows the use of a boundary discontinuity design. Second, I assemble a teacher panel dataset which, coupled with the school data, allows me to follow teachers across schools and space.

2.5.1 School Data

The school data uses information from different sources. The geolocation, as well as, the bonuses classification of each school is obtained from public records posted online by the Ministry of Education. Most school characteristics are obtained from the School Census (Censo Escolar) which is answered by school principals. This source has information on student enrollment, number of students repeating grade, and town characteristics.

I complement this data with average student test scores from the Students Census Evaluation (Evaluacion Censal de Estudiantes). This evaluation is a standardized reading and math test applied every year to second grade students in elementary schools.

2.5.2 Teacher Data

I use administrative data to assemble a panel dataset of teachers from 2013 to 2017. On one hand, I use a job position dataset to identify the schools where teachers worked as well as obtain their contract type (permanent or temporary). This dataset also contains information on age, gender, and experience. I complement this data with salary information that has details on the base monthly salary and all the bonuses received. Finally, I use information from the School Census to recover teacher qualifications.

2.5.3 Openings Data

For this project I collected data on job openings from school districts between 2015 and 2017. In the Peruvian public education system, each teaching position has a unique code. When a teacher leaves a position an opening is generated. Each school district is in charge of publishing and allocating these job openings. The central government is not directly involved and does no gather information on the process of allocating teachers to job openings. The job openings are published at the end of the school year by each school district.

I collected this data by contacting school district officials. I was able to recover a considerable number of job openings. The exact number of data recovered is shown in Table XX. The data is not representative for two reasons. First, we were unable to recover data from all school districts. Second, the job openings are concentrated in remote locations. This makes sense, because the openings will appear in less desirable locations. This will not be problematic for my estimates because I comparing job openings on two sides of a bonus boundary. Hence, to some extent the fact that I oversample openings in remote locations is beneficial for my identification strategy because this increases the likelihood of observing openings on both sides of the boundary.

The specific manner in in which job openings are assigned to teachers is described in detail the appendix. Nevertheless, it is important to understand three main points. First, schools with job openings have no say on what teacher can fill the position. teachers have all the power to decide where to move if there is an opening. Second, permanent teachers have priority over temporary teachers to choose an opening. Third, the job openings positions have three possible outcomes: they are filled by a permanent teacher, they are filled by a temporary teacher, or they remain vacant.

3 Boundary Discontinuity Design

In this section I provide the framework that combines matching estimators with the BDD. First, I describe why the bonus schedule induces a BDD that is unobserved. Second, I present the potential outcome framework. Finally, I describe the matching estimators I employ to recover the unobserved boundary and estimate the average treatment effect at the boundary.

3.1 Unobserved Boundaries

The border of non-urban salary bonuses are defined at the school level by two variables: distance to province capital and the population of the surrounding town, as already shown in Figure 2. This rule implicitly defines a border line between different bonus regions.[17] The boundaries can appear between Isolated, Rural, Suburban and Urban schools. This boundary is not to be confused with a fuzzy discontinuity design because, in this case, the treatment and control groups are well defined.

The geographical boundary between bonuses defines a two-dimensional discontinuity on the longitude-latitude space. In figure 3, I show a two dimensional hypothetical example of the implicit boundary between Isolated and Rural schools. I provide two examples from the data showing 20 km^2 blocks distinguishing school classification across space in Figure 16 in Appendix D.

A crucial issue I need to address is the presence of unobserved amenities which correlate with the bonus salary. The most remote schools have higher bonuses but are also in less desirable locations. If I do not control for such unobserved amenities, my estimates will suffer from omitted variable bias.

To address the omitted variable bias, I identify the effect of the bonus by employing the boundary discontinuity design assumption. I assume that two teachers working on opposite sides of the boundary face the same unobserved amenities, so that the only difference is the salary bonus. Formally, I assume the unobserved amenities change continuously across space.

The previous literature that incorporates a boundary discontinuity design has made use of a regression framework to estimate the average treatment effect. To control for unobservable spatial variables the literature uses geographical fixed effects (Bayer, Ferreira, and McMillan, 2007; Black, 1999) or flexible polynomials of longitude and latitude (Dell, 2010). Shapiro (2018) uses a regression framework but focuses on counties matched on both sides of a border. To some extent, he is embedding nearest-neighbor matching in a regression framework.

[17] Computationally, it is possible to calculate the bonus for each coordinate to uncover the true boundary but it is useless and time-costly to calculate from the government's perspective.

In this paper, I use matching estimators to recover the average treatment effect. Gelman and Imbens (2018) state that the main difference between various average treatment estimators is the weighting structure between treated and control units. Each matching estimator provides another a specific weighting structure. For example, kernel matching estimators weight comparable units by their proximity while the nearest-neighbor matching creates a discrete weight depending on the number of neighbors used.

I use a kernel and nearest-neighbor matching estimator for three reasons. First, I do not observe the exact location of the boundary. I only know the treatment status. By applying a matching estimator I am able to keep pairs (or groups) with both levels of treatment that are physically close. By doing so I can be confident that a boundary crosses between them as in figure 3. Second, kernel matching estimators allow me to gather more comparable units than would be possible using a boundary fixed-effects regression framework. The matching estimators find local neighborhoods for each treated unit while boundary fixed-effects regressions divide space in blocks, not taking advantage of the full potential of grouping comparable units. Third, the matching estimator allows me to give the most weight to the geographically closest units without throwing away information from more distant units.

3.2 Potential outcome framework

To describe a canonical matching estimator, I borrow the notation from Abadie and Imbens (2011) and Cameron and Trivedi (2005) to present the potential outcome framework developed originally by Rubin (1974). I define $W_i = 1$ as the teacher i receiving a bonus (treatment) and $W_i = 0$ as not receiving the bonus or receiving a lower bonus (control)[18]. With an abuse of notation, let the number and set of control units be denoted by N_0. Analogously, the number and set of treatment units is N_1. The outcome variable is the indicator variable equal to one if teacher i stays one more year at the same school. The potential outcome variable without treatment is $Y_i(0)$ and with treatment is $Y_i(1)$. The treatment effect for teacher i is $Y_i(1) - Y_i(0)$.

In the case of a unidimensional RDD where treatment is defined by a cutoff c in terms of some variable X_i (i.e., $W_i = \mathbb{1}\{X_i \geq c\}$), the statistic of interest is (Imbens and Lemieux,

[18] This can be expanded to comparing two regions with different bonuses.

2008):

$$\begin{aligned}\text{ATE}_c^{RDD} &= \lim_{\varepsilon\downarrow 0}\mathbb{E}\left[Y_i|\, X_i = c+\varepsilon\right] - \lim_{\varepsilon\downarrow 0}\mathbb{E}\left[Y_i|\, X_i = c-\varepsilon\right]\\ &= \mathbb{E}\left[Y_i\left(1\right) - Y_i\left(0\right)|\, X_i = c\right]\end{aligned}$$

The BDD expands this concept to a two-dimensional space, where each point of the the boundary can be thought of as specific RDD cutoff.[19] Therefore, the statistics of interest in the two-dimensional setting is the average across all the local RDD:

$$\begin{aligned}\text{ATE}_{\mathcal{B}} &= \int \mathbb{E}\left[Y_i\left(1\right) - Y_i\left(0\right)|\, \boldsymbol{X}_i = \boldsymbol{c}\right] dF\left(\boldsymbol{c}|\mathcal{B}\right) &(1)\\ &= \mathbb{E}\left[Y_i\left(1\right) - Y_i\left(0\right)|\, \mathcal{B}\right] &(2)\end{aligned}$$

where $\mathcal{B}$ is the set of boundary units. I will focus on $\text{ATET}_{\mathcal{B}}$, the average treatment effect on the treated at the boundary, which changes equation (2) slightly:

$$\text{ATET}_{\mathcal{B}} = \mathbb{E}\left[Y_i\left(1\right) - Y_i\left(0\right)|\, \mathcal{B}, W_i = 1\right]. \tag{3}$$

3.3 Estimator

The matching estimator used to calculated $\text{ATET}_{\mathcal{B}}$ creates a comparable group for each treated unit i at the boundary $\mathcal{B}$. For each treated unit we would like to calculate $Y_i\left(1\right) - Y_i\left(0\right)$, but only $Y_i\left(1\right)$ is observable and the main goal is to approximate $Y_i\left(0\right)$. Matching

[19] Formally, we can think of the BDD as averages of RDD estimates. Let $\mathcal{B} \subset \mathbb{R}^2$ be the set of boundary points where $\boldsymbol{c} \in \mathcal{B}$and $\boldsymbol{v_c}$ is a vector that crosses the boundary at $\boldsymbol{c}$. A graphical representation is shown in Figure 17 in Appendix D. If we only studied the direction of $\boldsymbol{v_c}$ at point $\boldsymbol{c}$, the estimate of interest would be

$$\text{ATE}_{\boldsymbol{c}}^{RDD} = \lim_{\varepsilon\downarrow 0}\mathbb{E}\left[Y_i|\, \boldsymbol{X}_i = \boldsymbol{c}+\boldsymbol{v_c}\varepsilon\right] - \lim_{\varepsilon\downarrow 0}\mathbb{E}\left[Y_i|\, \boldsymbol{X}_i = \boldsymbol{c}+\boldsymbol{v_c}\varepsilon\right]$$

Since we have many points then the statistic of interest in a BDD would be the average across the boundary points $\mathcal{B}$:

$$\begin{aligned}\text{ATE}_{\mathcal{B}} &\equiv \int \text{ATE}_{\boldsymbol{c}}^{RDD} dF\left(\boldsymbol{c}|\mathcal{B}\right) = \int \lim_{\varepsilon\downarrow 0}\mathbb{E}\left[Y_i|\, \boldsymbol{X}_i = \boldsymbol{c}+\boldsymbol{v_c}\varepsilon\right] - \lim_{\varepsilon\downarrow 0}\mathbb{E}\left[Y_i|\, \boldsymbol{X}_i = \boldsymbol{c}+\boldsymbol{v_c}\varepsilon\right] dF\left(\boldsymbol{c}|\mathcal{B}\right)\\ &= \int \mathbb{E}\left[Y_i\left(1\right) - Y_i\left(0\right)|\, \boldsymbol{X}_i = \boldsymbol{c}\right] dF\left(\boldsymbol{c}|\mathcal{B}\right)\end{aligned}$$

In my case I will be interested in using the distribution of $\boldsymbol{X}_i$ on the treatment side, hence, I only need to change the conditioning of the integral to obtain $\text{ATET}_{\mathcal{B}} = \int \mathbb{E}\left[Y_i\left(1\right) - Y_i\left(0\right)|\, \boldsymbol{X}_i = \boldsymbol{c}\right] dF\left(\boldsymbol{c}|\mathcal{B}, W_i = 1\right)$. This derivation is similar to Keele and Titiunik (2015).

estimators approximate $Y_i\left(0\right)$ by

$$\hat{Y}_i\left(0\right) \equiv \sum_{j \in A_i \cap \mathcal{B}} w\left(i,j\right) Y_j$$

where $A_i \subset N_0$ is the comparison set in the control group for treated unit i and $w\left(i,j\right)$ is an i-specific weighting function so that $\sum_{j \in A_i} w\left(i,j\right) = 1$ and $w\left(i,j\right) \in [0,1]$. The manner in which A_i is constructed and the functional form of $w\left(\cdot,\cdot\right)$ define different matching estimators. The ATET$_{\mathcal{B}}$ estimate is the aggregation across treated observations at the boundary

$$\hat{\text{ATET}}_{\mathcal{B}} \equiv \frac{1}{N_1} \sum_{i \in N_1 \cap \mathcal{B}} \left(Y_i - \sum_{j \in A_i} w\left(i,j\right) Y_j \right)$$

which is the analog of equation (3).

Now lets focus on nearest-neighbor matching and kernel matching. The matching is done comparing a two dimensional vector $\boldsymbol{x}_i = \left(x_i^{lon}, x_i^{lat}\right)$ that represent teacher i's workplace longitude and latitude coordinates. The M-nearest-neighbor matching estimator chooses M control units closest (with respect to some norm $\|\cdot\|$) to i . Formally, the nearest-neighbor comparison set A_i^M is

$$A_i^M = \left\{ j \in N_0 : \left| \left\{ k \in N_0 : \|\boldsymbol{x}_i - \boldsymbol{x}_k\| \leq \left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\| \right\} \right| \leq M \right\}$$

where I assume that there are no ties. The weighting function is

$$w^M\left(i,j\right) = \frac{1}{M}.$$

In the case of the kernel matching estimator, the comparison set is the whole control set, i.e. $A_i^K = N_0$, and the weighting function is

$$w^K\left(i,j\right) = \frac{K\left(\frac{\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|}{b}\right)}{\sum_{p \in N_0} K\left(\frac{\|\boldsymbol{x}_i - \boldsymbol{x}_p\|}{b}\right)}$$

where K is a Epanechnikov kernel, $\|\cdot\|$ is the same norm used for nearest-neighbors and b is the half-bandwidth. This methodology is analogous to propensity score matching, but instead of using distance between probabilities, it uses the physical distance between units.

For all my estimates I use the Euclidean norm and for the kernel matching estimator I apply the Epanechnikov kernel. Usually the bandwidth does not have an intuitive interpretation but in the case of distance-based matching, the bandwidth has a spatial interpretation. The Epanechnikov kernel creates a caliper with respect to distance as shown in Figure 4. As the bandwidth increases, the implied caliper also increases. For example, when using a $b = 5$, each treated unit is matched with observations within a 5 km radius[20] and closer units receive more weight.

Since my matching strategy only uses a distance matching procedure, I need to control for other observable characteristics such as age and gender. To do so, I employ a bias-adjustment procedure proposed by Abadie and Imbens (2011) such that

$$\hat{Y}_i(0) \equiv \sum_{j \in A_i \cap \mathcal{B}} w(i,j) \left(Y_j + \hat{\mu}_0(X_i) - \hat{\mu}_0(X_j)\right)$$

where $\hat{\mu}_0$ is the predicted value of a regression of Y_j on X_j restricted to the control group. The intuition behind this correction is straightforward. Since the matching procedure will gather teachers of different age groups, I need to subtract the age effect of the matched control unit $\hat{\mu}_0(X_j)$ and add the average outcome related to treated unit i's characteristics, $\hat{\mu}_0(X_i)$.

I choose not to include gender and age as matching variables for two reasons. First, the discontinuity is spatial and I match observations in the same dimension. As in a propensity score matching, I do not include irrelevant variables in the logit function that do not affect the propensity to be treated. Second, the interpretation of the bandwidth is intuitive and transparent. By including gender or age, I would lose these two properties.

A regression framework does not encounter these issues, because all the variables would be included as regressors. In the regression framework, I use 5 and 10 km^2 fixed-effects (which I will call block fixed-effects) to approximate the local amenities and local markets. Also, I only focus on treated and control observations that have at least one observation from the other group in a 10 km radius[21]. The regression I estimate is given by

$$Y_i = Z_i \alpha' + \beta W_i + \phi_{b(i)} + \varepsilon_i \tag{4}$$

where Z_i are control variables, W_i is the treatment variable, and $\phi_{b(i)}$ is the block fixed-effect for unit i. The ATET$_{\mathcal{B}}$ estimator in this framework is $\hat{\beta}$.

[20] Any observation outside of the 5 km radius receives a weight of zero.

[21] The results are robust to the radius size (1 km and 5 km radii), but I do lose power when using smaller radii.

4 Results

I aim to quantify the effect of the bonus on teacher retention and teacher composition within schools. I focus my analysis on the three boundaries that provide the steepest salary changes: Isolated/Rural boundary, VRAEM boundary and Frontier boundary.[22] The estimates show three patterns. First, temporary teachers do show an increase in retention rates across the different bonuses, while permanent teachers are unaffected by them. Second, the effect of the bonuses on teacher retention is not monotonic in terms of its monetary value. For example, the effect of S/100 bonus (Frontier bonus) can have a similar impact compared to a S/400 bonus increase (Isolated/Rural boundary). Third, the bonus appear to not have major effects on teacher composition within schools.

I first show graphically the monthly salary discontinuities induced by the bonus schedule. These bonuses cover different areas of Peru as shown in Figures 18 and 19 in Appendix D. Then, I provide evidence for the validity of the BDD across the three boundaries. Finally, I present the ATET$_{\mathcal{B}}$ estimates on teacher retention and teacher composition within schools.

4.1 Monthly Salary Discontinuities

To fix ideas and show that the bonus schedule does generate salary discontinuities, I first show figures in the regression discontinuity tradition as in Bayer, Ferreira, and McMillan (2007). I do so with the monthly salary at the three boundaries. To collapse the information into a unidimensional space I keep pairs of treated and control units such that they are less than 10 km apart. Since I am unable to know where the geographical boundary is located, I assume the boundary is exactly between the two observations. Treated observations are placed to the right of zero, and control observations to the left. I also draw a line for the average salary in each group with its corresponding confidence interval.[23] All the figures are plotted separately by permanent, qualified temporary and unqualified temporary.

The figures for the Isolated/Rural boundary in 2013 and 2016 are shown Figures 5 and 6. As expected, in 2013 the difference is tiny. In 2016, the gap of the monthly salaries at the discontinuity should be around S/400 and the magnitude calculated in the graph closely approximates this value.

22 The salary gaps for the Isolated/Rural boundary, the VRAEM boundary and the Frontier boundary are S/400, S/300 and S/100.

23 To avoid confusion, I do not fit a polynomial as in classic RDD.

For the VRAEM and Frontier boundaries I show the results only for 2016. In each case, the gap should be around S/300 and S/100, respectively. The magnitude of the gap at the discontinuities does roughly match the value established by the bonus schedule.

4.2 Validity of Boundary Discontinuity Design

To confirm the validity of the BDD, I use the a kernel matching estimator with a 5 km half-bandwidth to test for potential discontinuities on observable variables at the school level in the base year of 2013. The results are shown in Table 27 of the Appendix G. The reading and math scores are not significantly different at the three boundaries. Nevertheless, the Isolated/Rural and the Frontier boundary shows different averages for the presence of public utility in the surrounding town (potable water, electricity and sanitation). The VRAEM boundary does not show the discontinuities in public utility provision.

To alleviate that towns are not completely similar on both sides of the boundary, I control for the presence of public utilities by including these variables in the bias-adjustment covariates.

4.3 Teacher Retention

I estimate $\text{ATET}_{\mathcal{B}}$ on teacher retention which is an indicator variable equal to 1 if teacher stays at the same school one more year. I employ two matching estimators and a regression with block fixed-effects. Specifically, I use two and four-nearest-neighbors matching estimators and a Epanechnikov kernel matching estimator with 5 and 10 km half-bandwidth. In both cases the matching is based on as-the-crow-flies distance.[24] For both matching estimators I apply the bias-adjustment procedure to control for age, gender, education level of school and characteristics of the surrounding town (water, electricity and

[24] Since the matching estimator software does not calculate the geodesic distance between two geographical locations, I approximate longitude and latitude distances. Peru is close to the equator and I can approximate 1° latitude and 1° longitude by 110.57 kilometers and 111.32 kilometers. The approximation is quite good and the results do not change when using the correct geodesic calculations. The differences between the approximate distance and the distances using geodesics are negligible, especially in short distances.

sanitation services). The bias-adjustment procedure has the following form

$$\hat{Y}_i(0) = \sum_{j \in A_i} w(i,j) \Bigg(Y_j - \hat{\alpha}^0 \text{Age}_j - \hat{\beta}^0 \mathbb{1}\{\text{Gender}_j\text{=Female}\} - \sum_{l=1}^{3} \hat{\gamma}_l^0 \mathbb{1}\{\text{Level}_j\text{=}l\} - \sum_{c=1}^{3} \hat{\phi}_l^0 \mathbb{1}\{\text{Town Characteristic}_j\text{=}c\}$$

$$+ \hat{\alpha}^0 \text{Age}_i + \hat{\beta}^0 \mathbb{1}\{\text{Gender}_i\text{=Female}\} + \sum_{l=1}^{3} \hat{\gamma}_l^0 \mathbb{1}\{\text{Level}_i\text{=}l\} + \sum_{c=1}^{3} \hat{\phi}_l^0 \mathbb{1}\{\text{Town Characteristic}_i\text{=}c\} \Bigg)$$

where $\hat{\alpha}^0, \hat{\beta}^0, \hat{\gamma}_l^0$ and $\hat{\phi}_l^0$ are estimates obtained from regressing Y_j on the bias-adjustment covariates. The superscript "0" refers to the fact that this regression is evaluated only within the control group. I also estimate the regression in (4) with 5 km^2 and 10 km^2 block fixed-effects whenever they are feasible.

For each boundary I present two tables of results. The first table only shows the kernel matching estimators with a 5 km half-bandwidth pooling across all education levels. The second table shows the estimates separately for each education level and for every estimator.

4.3.1 Isolated vs Rural

The results on teacher retention, pooling across schools at the Isolated/Rural border, is shown in Table 8. I provide a graphical representation of the results for 2013 and 2016 in Figures 9 and 10, respectively.[25] Coefficients of Table 8 are represented by the bold black vertical lines with the **ATET** label.[26] The complete set of estimates is shown in Table 9.

My estimates show heterogeneous effects of the bonus on retention rates. First, temporary teachers show a positive effect in retention rates in the post-reform years. In Panel B. of Table 9, I show the results for qualified temporary teacher across different matching estimators and separately by levels of education. I quantify an increase of yearly retention rates of 4 to 8 percentage points from a base rate of approximately 25%. The regression block fixed-effects estimates show similar magnitudes but they are not statistically significant. This could be due to the fact that the matching estimators "tailor" fixed-effects for each treated observation.

[25] The graphical representation does not include unqualified temporary teachers because there results are non-significant in all cases.

[26] It is worth noting that reallocation restrictions where in place for permanent teachers during 2013 and 2014. That explains the null effect and very high retention rates, and it also explains the deep in retention in 2015.

Second, I find no significant results for permanent teachers. Even though the permanent teachers' results in Table 8 show a slight positive effect in 2016, the results are not robust. The estimates in Panel A. of Table 9 show that the effect is not consistent across education levels and estimators. The only exception is Elementary school teachers.[27]

It is important to note that the base rate for permanent teachers is much higher, above 90%. At most retention can increase in 10% percentage points. I will examine this idea with a discrete choice model. The main intuition is that this high retention level is revealing a strong preference for the teachers that stay at remote locations bonus. Given our previous statistics, there are few permanent teachers that want to move and the ones that move are not deterred by the bonus.

Finally, since the reform was implemented in 2015, the previous year provides a suitable placebo test. The estimates for 2014 are non-significant in all cases, reassuring the findings for this boundary.[28]

4.3.2 VRAEM vs non-VRAEM

The VRAEM boundary presents the second steepest (S/300) salary gap. The pooled results for Permanent and Qualified temporary teachers are shown in Table 10.[29]

This boundary is interesting because it shows the weakest effect on both type of teachers. Permanent teachers are not affected and these results are confirmed by the different specifications in Table 11. On the other hand, the estimates for qualified temporary are only show significant and robust results for for Kindergarten students after the reforms. The placebo test are mainly non-significant. The graphical representation for 2013 and 2016 are shown in Figures 22 and 23 of the Appendix.

The regression framework with 5 and 10 km^2 block fixed-effects where infeasible to estimate in many cases so I omit them from the results. This is another nice feature of matching estimators in comparison to a more constrained regression. The matching estimator is able to "search" for comparable units to create a counterfactual $\hat{Y}_i(0)$ for treated unit i. The matching estimator "creates" suitable fixed-effects for each observation.[30]

[27] The results for unqualified temporary teachers are noisy and do not show any clear patterns (Panel C. of Table 9). The smaller size of their sample could be reducing the power of the tests, but I do not see any strong evidence in the robustness check that leads to belief that unqualified teachers retention rate was altered.

[28] The samples for 2013 were too small to be estimated separately by levels of educations. That is why I only present results pooled results for 2013.

[29] The unqualified temporary teacher sample is too small to calculate $\text{ATET}_{\mathcal{B}}$ along the VRAEM boundary.

[30] This equivalence between matching estimators and fixed-effects has been examined by Imai and Kim

4.3.3 Frontier vs non-Frontier

The Frontier boundary has the smallest increment (S/100) out of the three boundaries but shows the most consistent effect on qualified temporary teachers.[31]

The results in the Frontier boundary reassure the inability of the current bonus schedule to modify permanent teachers' retention rate. The pooled results are shown in Table 12 and the robustness checks in Table 13. The graphical representation for 2013 and 2016 are shown in Figures 20 and 21 of the Appendix.

For qualified temporary teachers, the effects are strong in 2016. This effect is spread across all levels of education and is captured throughout all the matching estimators. The S/100 bonus is related to an increase in retention between 5% to 13% from a base retention rate ranging from 29 to 38%. For 2015, I only find a significant effect for Kindergarten teachers. The placebo tests in 2014 are inconsistent and change from one specification to the other. Given this result we can have some confidence that the post-reform results are capturing the effect of the bonus. Compared to the Isolated/Rural boundary, this estimates show that the effect of the bonus on teacher retention do not necessarily behave monotonically.

4.4 Teacher Retention at the School Level

I apply the same method but instead of focusing on teachers, the analysis is carried out at the school level. To make results comparable with the previous results, the outcome variable is the fraction of teachers that stay at the same school an additional year.[32]

The results for the Isolated/Rural and VRAEM boundaries are similar to those obtained at the teacher level. The results are shown in Tables 18 and 19 in Appendix E. As noted before, qualified temporary teachers react significantly at the Isolated/Rural boundary, while permanent and unqualified temporary teachers appear unaffected by the bonus. The regression framework provides similar estimates but the magnitudes are smaller

(2018).

31 This boundary does not have enough unqualified temporary teachers to estimate their $\text{ATET}_{\mathcal{B}}$.

32 The new bias-adjustment procedure has the following form

$$\hat{Y}_i(0) = \sum_{j\in A_i} w(i,j)\left(Y_j - \sum_{l=1}^{3}\hat{\gamma}_l^0 \mathbb{1}\{\text{Level}_j{=}l\} - \sum_{c=1}^{3}\hat{\phi}_l^0 \mathbb{1}\{\text{Town Characteristic}_j{=}c\}\right.$$
$$\left. + \sum_{l=1}^{3}\hat{\gamma}_l^0 \mathbb{1}\{\text{Level}_i{=}l\} + \sum_{c=1}^{3}\hat{\phi}_l^0 \mathbb{1}\{\text{Town Characteristic}_i{=}c\}\right)$$

and statistically not significant. The placebo test for 2013 and 2014 provide robust non-significant results, as desired. The VRAEM boundary shows mainly non-significant effects across all groups and estimators for all years.

The results for the Frontier boundary are less robust than when estimated at the teacher level. Permanent teacher do not show any sign of being affected by the bonus, while qualified temporary teachers still show strong effects for Kindergarten in 2015 and Elementary in 2016. The results do not show statistical significance for Kindergarten and High-school schools in 2016. The treatment effects are shown in Table 20 in Appendix E. The placebo test for Kindergarten schools in 2014 also fail, but apart from this anomaly, all other placebo test do not show effects before the implementation of the reform.

4.5 Teacher Attraction

To study teacher attraction I use the job openings data. After a job opening three potential events can happen: it is covered by a permanent teacher, it is covered by a temporary teacher or it is left vacant. I describe the job openings in terms of this three potential outcomes in Table XX. The data shows clearly this job openings are mainly covered by temporary teachers, then come permanent teachers and finally a small portion of positions are unable to attract teachers. From the Tables two clear patterns appear. First, openings in urban towns are more attractive to permanent teachers. The share of permanent teachers that fill this positions increase as we get closer to urban areas. Second, the vacancy rate is uniform across independent of the isolation of the school.

As should be expected, openings usually do not attract tenured teachers. The most wanted positions will not be available because tenured teachers will hold them for long periods of time. Hence, the openings left will usually be undesirable. I Even though the positions are undesirable, the government is able to fill them.

To show how undesirable this positions are, I show what is the share of the openings that get covered by tenured, temporary and

The job opening data is not centralized and the process is carried out by school districts. Job openings are first available to tenured teachers and if they are empty they are available for temporary teaches.

4.6 Teacher Composition

An objective of the bonus schedule was to affect teacher composition of schools. Given the results in section 2.4 that show a relation between teacher composition and standardized tests, a first effect needed to potentially affect education outcomes would require modifying the teacher composition in schools. This is the “first stage” of the channel through which bonuses potentially affect educational outcomes.

I focus on two outcome variables: the fraction of permanent teachers and the fraction of qualified temporary teachers at the school level. In general, my results show that even though the bonus has some effect on teacher reaction, it does not affect the teacher composition at schools. The results for the Isolated/Rural, VRAEM and Frontier boundary are shown in Tables 21, 22 and 23 in the Appendix.

The treatment effects estimated for the VRAEM boundary do not show significant effects across levels of education. This result is expected because the VRAEM bonus did not even have an effect on teacher retention at the boundary.

In the case of the Frontier bonus, I also do not see significant effects on teacher composition: no robust and significant effect on fraction of permanent or qualified temporary teachers. The estimated effect of the Frontier bonus on qualified temporary retention from the previous section could have potentially modified the composition of teachers, but this does not seem to be the case.

In both boundaries, VRAEM and Frontier, the placebo test for 2013 and 2014 are non-signifcant. This does not happen in the case for the Isolated/Rural boundary. In this boundary I consistently find more permanent teacher on the Rural side (control group) before the bonus implementation. This is problematic for my estimates, because I am unable to rule out if the bonus had any effect on teacher composition. Nevertheless, given that the difference in shares of permanent teacher is constant across periods, it seems reasonable that the bonus did not have an effect on teacher composition.

4.7 Limitations

The main limitation of these results is that the stable unit treatment value assumption (SUTVA) may not hold. This assumption requires that control units are not affected by the bonus schedule at the boundary. It can be the case that the retention rates are affected on the control side of the boundary, because they now want to also obtain a higher bonus. In that sense, the $\text{ATET}_{\mathcal{B}}$ estimates could be recovering the net effect of the presence of the boundary. This issue has not been dealt with in previous research on teacher retention, and could be a possible venue of investigation.

5 Interpretation of Teacher Retention under a Discrete Choice Model

In this section I interpret the teacher retention results under a discrete choice model and provide a potential explanation why I observe no effect on permanent teachers, while I do find effects on qualified temporary teachers. First, I will present the framework to think about the teacher's migration problem. Second, I will apply this framework to my specific variable of interest which is teacher retention. Finally, I provide an interpretation of the $\text{ATET}_{\mathcal{B}}$ estimates.

5.1 Framework

Consider teacher i working at origin o at date $t-1$. At the end of period $t-1$ she can decide to stay or move to another location. The teacher faces a location choice set D and the utility of working/living at destination d at time t is denoted by[33]

$$U_d = u_d + \beta B_d - c_{od} + \varepsilon_d$$

where $u_d + \beta B_o + \varepsilon_d$ is the utility of living at destination d , receiving a bonus B_d and β is the marginal utility of the bonus. The shock ε_d is a independent and identically distributed random variable across destinations, and c_{od} is the cost of moving from origin o to destination d.[34] The optimal migration problem is given by

$$d^* \equiv \arg\max_{d \in D} \{u_d + \beta B_d - c_{od} + \varepsilon_d\}$$

Assuming ε_d has a type-I extreme value distribution then the likelihood of choosing location d is[35]

$$\mathbb{P}\left[d^* = d \mid o\right] = \frac{\exp\left(u_d + \beta B_d - c_{od}\right)}{\sum_{d'} \exp\left(u_{d'} + \beta B_{d'} - c_{od'}\right)} \tag{5}$$

33 All the variable have an implicit i.

34 $C_{it,d} = 0$ if $d = o_{it}$.

35 I can allow for a more flexible specific such as nested logit outside the origin o, but the same analysis carries through.

Using this model, I can focus on the decision to stay and rewrite equation (5) only for the case where $d^* = o$:

$$\mathbb{P}\left[d^* = o|o\right] = \frac{\exp\left(u_o + \beta B_o\right)}{\Phi_o + \exp\left(u_o + \beta B_o\right)} \tag{6}$$

Where

$$\Phi_o \equiv \sum_{d' \in D, d' \neq o} \exp\left(u_{d'} + \beta B_{d'} - c_{od'}\right).$$

I rescale the utility by $\ln\left(\Phi_o\right)$ to obtain a logit formulation

$$\mathbb{P}\left[d^* = o|o\right] = \frac{\exp\left(\tilde{u}_o + \beta B_o\right)}{1 + \exp\left(\tilde{u}_o + \beta B_o\right)} \tag{7}$$

Notice that I have not imposed any structure on $\tilde{u}_o$ and it can be thought of as a function of origin o. The variable $\tilde{u}_o$ includes the local amenity value u_o and the inclusive value of the outside opportunities with respect to o: Φ_o.

Imagine now that two identical agents work at location o and one has a bonus of value B for working at location o while the other agent does not have the bonus. The agent with bonus has a probability of staying at location o given by (7), while the agent with no bonus has the probability of staying

$$\mathbb{P}\left[d^* = o|o\right] = \frac{\exp\left(\tilde{u}_o\right)}{1 + \exp\left(\tilde{u}_o\right)} \tag{8}$$

If we think of o as a longitude-latitude coordinate, it is impossible for two teachers to work at that same location and have different bonuses. Nevertheless, for a location o that is close to the boundary and close to to a control unit at location o' on the other side of the boundary, we can assume that u_o and $u_{o'}$ are very similar (e.g. $u_o \approx u_{o'}$.)

5.2 Relation between Average Treatment of Retention and Discrete Choice

The outcome variable I am have focused mostly is teacher retention or the decision to stay at the origin location o:[36]

$$Y_o = \mathbb{1}\left\{d^* = o\right\} \tag{9}$$

where $\mathbb{1}$ is the indicator function. Note the LHS of (9) is the outcome variable from the BDD framework, while the right hand side is the event inside the likelihood of (8). This simple equivalence is key to interpret the $\text{ATET}_{\mathcal{B}}$ estimates.

[36] I changed the subscript from i to location o to emphasize the local unobservables, $\tilde{u}_o$.

To simplify the explanation, lets focus on the border between a region with bonus B and another region with no bonus. The treatment group is the school with the salary bonus and the utility of the bonus B at location o is $\tilde{u}_o + \beta B$.

As I mentioned before, the treatment effect of the bonus on choosing to stay is $Y_o(1) - Y_o(0)$. Taking expectations conditional on location o and $W = 1$:

$$\begin{aligned} & \mathbb{E}\left[Y_o(1) - Y_o(0) \mid o, W = 1\right] & (10) \\ & = \mathbb{P}\left[d^*(1) = o \mid o, W = 1\right] - \mathbb{P}\left[d^*(0) = o \mid o, W = 1\right] \\ & = \frac{\exp(\tilde{u}_o + \beta B)}{1 + \exp(\tilde{u}_o + \beta B)} - \frac{\exp(\tilde{u}_o)}{1 + \exp(\tilde{u}_o)} & (11) \end{aligned}$$

Let $\mathcal{B}$ be the set of teachers on the treated side of the boundary. I can then integrate (10) across location origins to obtain:

$$\text{ATET}_{\mathcal{B}} \equiv \int_{\mathcal{B}} \mathbb{E}\left[Y_o(1) - Y_o(0) \mid o, W = 1\right] dF(o|\mathcal{B}, W = 1)$$

where $\text{ATET}_{\mathcal{B}}$ is the average treated effect on the treated at the boundary and which is the estimate from section 3. From a discrete choice perspective the same integral is equal to

$$\text{ATET}_{\mathcal{B}} = \int_{\mathcal{B}} \frac{\exp(\tilde{u}_o + \beta B)}{1 + \exp(\tilde{u}_o + \beta B)} - \frac{\exp(\tilde{u}_o)}{1 + \exp(\tilde{u}_o)} dF(o) \tag{12}$$

The estimation of $\text{ATET}_{\mathcal{B}}$ has a discrete choice representation, but it does not allow to examine counterfactuals because to do so I would need to recover $\tilde{u}_o$ individually.[37]

5.3 Interpretation of heterogeneous effects

Lets focus on the Isolated/Rural boundary. From Figures 9 to 23 it is obvious that the level of retention of permanent teachers is much higher than that of qualified temporary teachers. Any given year, almost 90% of them will stay at the same school. The retention rate is one of the migration choice probabilities and can be mapped to $\tilde{u}_o$.

Using the retention rates of 2016, assume retention in the control group is 90% for all permanent teachers and 25% for all qualified temporary. Using equation (8), the implied value of $\tilde{u}_0$ for permanent and qualified temporary teachers would be 2.2 and -1.1, re-

[37] In Appendix C I provide a potential way to recover $\tilde{u}_o$.

spectively.[38] From the qualified temporary teachers' $\hat{\text{ATET}}_{\mathcal{B}}$, I estimate that a S/400 wage bonus increases retention by 4 and 8 percentage points. Assuming the increase in retention rate is of 6 percentage points, I can obtain a back-of-the-envelope calculation for β by assuming away the integral of equation (12) and solving for β.[39] Using this value, I can then predict the increase in permanent teachers' retention rate in about 2.4 percentage points. Hence, even if the effect was present in permanent teachers, it would be harder to distinguish it from zero because its magnitude in terms of retention rate would be 1/3 of that of qualified teachers.

I have not explained why permanent teacher have a higher value of $\tilde{u}_0$ for remote locations. One possible answer is that selection is occurring only for permanent teachers. Permanent teacher have priority to move across schools and having stayed at remote locations reveals their strong preference. A complementary answer is that permanent teachers solve a forward-looking problem (which I have not address in this paper) while temporary teachers only face a one-shot problem. As mentioned before, tenured teachers have to stay two years at the destination school or they lose tenure. Temporary teachers do not face this problem, they can quit whenever they want and not loose future earnings. Hence, the combination of selection with forward-looking behavior could potentially explain the differential effect between permanent and qualified temporary teachers.

6 Conclusions

In this paper, I estimated the effect of a location-specific bonus schedule that tried to alleviate the unequal distribution of teachers. More specifically, I focused on how this bonus schedule affected teacher retention rates and the composition of teachers within schools.

My estimates show that the effect of the bonus schedule interacts with the teachers' contract type. Permanent teachers seem to be unaffected by the bonus, or in light of the discrete choice model, require a much higher bonus to raise their retention rates. Qualified temporary teachers responded the most, with their retention rate by between 4 and 12 percentage points depending on the bonus. Relative to the original retention rates, the increase was almost 33%. I also find that bonuses do not have a monotonic effect on retention rates.

[38] I obtain these values by equating $\frac{\exp(\tilde{u}_o^p)}{1+\exp(\tilde{u}_o^p)} = 0.9$ and $\frac{\exp(\tilde{u}_o^q)}{1+\exp(\tilde{u}_o^q)} = 0.25$.

[39] The values would be of 0.0000746.

The contract types seem to induce very different incentive problems for both types of teacher, maybe because permanent teachers that choose remote location are typically retained even without the bonus. Also, permanent teachers face a forward-looking problem, while temporary teachers play a one-shot problem every period. This could be a potential driver of the difference in the effect of the bonus schedule on retention rates, and maybe more generally of all the migration decisions.

Methodologically, I apply a matching estimator approach to a boundary discontinuity design. This approach allows to estimate the average treatment effect at the boundary without directly observing the boundary. The matching approach seems to perform better than the classic fixed-effect approach because it searches and tailors comparable observations for treated units.

References

ABADIE, A., AND G. W. IMBENS (2011): "Bias-Corrected Matching Estimators for Average Treatment Effects," *Journal of Business & Economic Statistics*, 29(1), 1–11.

ALBOUY, D. (2009): "The Unequal Geographic Burden of Federal Taxation," *Journal of Political Economy*, 117(4), 635–667.

ALVA, C., M. BOBBA, G. LEON, C. NEILSON, AND M. NIEDDU (2017): "Teacher Wages, Student Achievement and the Recruitment of Talent in Rural Peru," Discussion paper, Unpublished Manuscript.

BAYER, P., F. FERREIRA, AND R. MCMILLAN (2007): "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," *Journal of Political Economy*, 115(4), 588–638.

BAYER, P., R. MCMILLAN, A. MURPHY, AND C. TIMMINS (2016): "A Dynamic Model of Demand for Houses and Neighborhoods," *Econometrica*, 84(3), 893–942.

BIASI, B. (2018): "The Labor Market for Teachers Under Different Pay Schemes," NBER Working Papers 24813, National Bureau of Economic Research, Inc.

BISHOP, K. C., AND A. D. MURPHY (2011): "Estimating the Willingness to Pay to Avoid Violent Crime: A Dynamic Approach," *American Economic Review*, 101(3), 625–29.

BLACK, S. E. (1999): "Do Better Schools Matter? Parental Valuation of Elementary Education," *The Quarterly Journal of Economics*, 114(2), 577.

CABRERA, J. M., AND D. WEBBINK (2018): "Do higher salaries yield better teachers and better student outcomes?," Discussion Paper 86972, MPRA, April 2017.

CAMERON, A. C., AND P. K. TRIVEDI (2005): *Microeconometrics: Methods and Applications*. Cambridge University Press.

CARD, D., AND A. B. KRUEGER (1994): "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania," *The American Economic Review*, 84(4), 772–793.

CASTRO, J. F., AND B. ESPOSITO (2018): "The Effect of Bonuses on Teacher Behavior: A Story with Spillovers," Discussion Paper 104, Peruvian Economic Association, August.

CLOTFELTER, C., E. GLENNIE, H. LADD, AND J. VIGDOR (2008): "Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina," *Journal of Public Economics*, 92(5), 1352 – 1370.

DELL, M. (2010): "The Persistent Effects of Peru's Mining Mita," *Econometrica*, 78(6), 1863–1903.

DIAMOND, R. (2016): "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," *American Economic Review*, 106(3), 479–524.

DOLTON, P., AND W. V. D. KLAAUW (1999): "The Turnover of Teachers: A Competing Risks Explanation," *The Review of Economics and Statistics*, 81(3), 543–550.

DOLTON, P., AND W. VON DER KLAAUW (1995): "Leaving Teaching in the UK: A Duration Analysis," *The Economic Journal*, 105(429), 431–444.

DUBE, A., T. W. LESTER, AND M. REICH (2010): "Minimum Wage Effects Across State Borders: Estimates Using Contiguous Counties," *The Review of Economics and Statistics*, 92(4), 945–964.

GELMAN, A., AND G. IMBENS (2018): "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs," *Journal of Business & Economic Statistics*, 0(0), 1–10.

GILPIN, G. A. (2011): "Reevaluating the effect of non-teaching wages on teacher attrition," *Economics of Education Review*, 30(4), 598 – 616.

HANUSHEK, E. A., J. F. KAIN, AND S. G. RIVKIN (2004): "Why Public Schools Lose Teachers," *The Journal of Human Resources*, 39(2), 326–354.

HECKMAN, J. J., H. ICHIMURA, AND P. TODD (1997): "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," *Review of Economic Studies*, 64(4), 605–54.

——— (1998): "Matching As An Econometric Evaluation Estimator," *The Review of Economic Studies*, 65(2), 261–294.

HOLMES, T. (1998): "The Effect of State Policies on the Location of Manufacturing: Evidence from State Borders," *Journal of Political Economy*, 106(4), 667–705.

IMAI, K., AND I. S. KIM (2018): "When Should We Use Unit Fixed Effects Regression Models for Causal Inference with Longitudinal Data?," Discussion paper, Harvard University.

IMAZEKI, J. (2005): "Teacher salaries and teacher attrition," *Economics of Education Review*, 24(4), 431 – 449.

IMBENS, G. W., AND T. LEMIEUX (2008): "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 142(2), 615 – 635, The regression discontinuity design: Theory and applications.

JANN, B. (2017): "kmatch: Stata module for multivariate-distance and propensity-score matching," https://ideas.repec.org/c/boc/bocode/s458346.html.

KEELE, L. J., AND R. TITIUNIK (2015): "Geographic Boundaries as Regression Discontinuities," *Political Analysis*, 23(1), 127–155.

KENNAN, J., AND J. R. WALKER (2011): "The Effect of Expected Income on Individual

Migration Decisions," *Econometrica*, 79(1), 211–251.

LANGROCK, R., N.-B. HEIDENREICH, AND S. SPERLICH (2014): "Kernel-based semiparametric multinomial logit modelling of political party preferences," *Statistical Methods & Applications*, 23(3), 435–449.

MAJEROWICZ, S., AND R. MONTERO (2018): "Can Teaching be Taught? Experimental Evidence from a Teacher Coaching Program in Peru (Job Market Paper)," Discussion paper, Harvard University.

MÜLLER, M. (2001): "Estimation and testing in generalized partial linear models—A comparative study," *Statistics and Computing*, 11(4), 299–309.

MURNANE, R., AND R. J. OLSEN (1990): "The Effects of Salaries and Opportunity Costs on Length of Stay in Teaching: Evidence from North Carolina," *Journal of Human Resources*, 25(1), 106–124.

POSNER, D. N. (2004): "The Political Salience of Cultural Difference: Why Chewas and Tumbukas Are Allies in Zambia and Adversaries in Malawi," *The American Political Science Review*, 98(4), 529–545.

RAFFO, J., AND S. LHUILLERY (2009): "How to play the "Names Game": Patent retrieval comparing different heuristics," *Research Policy*, 38(10), 1617–1627.

RUBIN, D. (1974): "Estimating causal effects of treatments in randomized and nonrandomized studies," *Journal of Educational Psychology*, 66(5), 688–701.

SEVERINI, T. A., AND W. H. WONG (1992): "Profile Likelihood and Conditionally Parametric Models," *The Annals of Statistics*, 20(4), 1768–1802.

SHAPIRO, B. T. (2018): "Positive Spillovers and Free Riding in Advertising of Prescription Pharmaceuticals: The Case of Antidepressants," *Journal of Political Economy*, 126(1), 381–437.

A Job Openings Data

Collecting the data was presented a couple challenges to overcome. First, the job openings data is managed by each school district and the central government does not have the information. To collect the data we had to personally contact 200 school district officials. Second, the data we recovered does not have a uniform formats: printed documents, pdf's, Word documents and Excel spreadsheets. Third, we digitize the printed documents and pdf's using an Optical Character Recognition software. Each job opening has a unique code and the OCR software sometimes recovered them with typos. To fix the typos and correctly match the job openings data to the administrative I used a fuzzy matching process implemented by Raffo and Lhuillery (2009). For those unmatched openings we would carry out a manual search in the administrative data.

B Spatial Life-Cycle

The combination of tenure with full mobility induces a strong spatial life-cycle pattern of permanent teachers. This spatial life-cycle appears as result of the combination two empirical facts. First, older teachers are less likely to change schools. Second, teachers from non-urban school are more likely to move to urban locations than the opposite way. These two facts result in a unequal distribution of experience and permanent/temporary teachers across space. Urban locations have older and permanent teachers while the isolated schools have young temporary teachers.

The spatial life-cycle pattern observed in the data is relevant from a methodological perspective, but thinking about policy. Methodologically, the life-cycle behavior of teachers can potentially induce a strong selection of teachers. The older permanent teachers established in isolated locations before the education reform did not require a bonus to work there and this reveals their strong preference for the location. From a policy perspective, the personnel policies implemented by governments can be undone by these transition patterns. Any policy which tries to induce a more uniform distribution of human resources needs to take into account the transition patterns of established permanent teachers. For example, a policy aimed to incentivize recently hired permanent teachers to hold non-urban positions would have to overcome the strong life-cyle patterns.

To simplify the exposition, I focus on the transitions between urban and non-urban schools. The main point is that the teacher drainage from non-urban location to urban schools is significant . Non-urban schools constantly loose teachers, but even more during their first years. Figure 11 shows that around 4% of teachers between the ages of 30 to 40 migrate from non-urban schools to urban schools.[40] A analogous figure for teachers working at urban schools is shown in Figure 12. Exit rates of the public system in both cases is below 1%.

To help understand the magnitude of this transitions, I use the transitions shown in Figures 11 and 12 to recreate a cohort of 100 permanent teachers at age 30. Of the 100 teachers, 70 of them work at non-urban schools and 30 work at urban schools. As the cohort ages, it moves between urban and non-urban regions and I follow the cohorts size in each location. The calculations are shown in Figure 13. After 10 years in the public system when teachers are 40 years old, the cohort is evenly divided between urban and non-urban schools. By the time teachers are close to retirement, 55 of them are in urban schools and 33 are in rural schools.

Until now I have only focused on permanent teachers because they exhibit a life-cycle patterns that favors urban locations. Given their contractual nature, qualified and unqualified temporary teachers move more between urban and non-urban schools and the preference for urban locations does not show as clear as with permanent teachers.

[40] Figure 11 shows the transition conditional on age. To understand the figure, fix the attention in 40 year old teachers. The figure shows that 4% of teachers at age 40 at non-urban schools move to an urban school and less than 1% exit the public system in one year.

C Recovering $\tilde{u}_0$

In this section I extend the program evaluation framework to a discrete choice setting. To recover $\tilde{u}_0$, I use a semiparametric approach proposed by Müller (2001) and extended by Langrock, Heidenreich, and Sperlich (2014). This approach is consistent with the BDD in the sense that the value of $\tilde{u}_o$ is explicitly modeled as a continuous function of space such that

$$u_o = m\left(\boldsymbol{x}_i\right) \equiv m\left(\text{latitude}_o, \text{longitude}_o\right).$$

The likelihood in (8) then becomes

$$\mathbb{P}\left[d^* = o|o\right] = \frac{\exp\left(m\left(\boldsymbol{x}_o\right) + \beta B\right)}{1 + \exp\left(m\left(\boldsymbol{x}_o\right) + \beta B\right)}. \tag{13}$$

The estimation of (13) follows the profile likelihood concept developed by Severini and Wong (1992) where function $m\left(\cdot\right)$ is estimated via kernel smoothing. As with the previous bandwidth-selection discussion, the bandwidth has the same spatial interpretation. The semiparametric estimation is the counterpart to the kernel matching estimator. At a first glance, one could think that in the same way, we could incorporate fixed effects and polynomials to recover u_o as a counterpart to the regression framework. This is not feasible in a logit framework because it suffers for the incidental parameter problem.

The idea behind the estimation procedure uses the same intuition as all kernel smoothing estimators. It applies a maximum likelihood approach but the semi-parametric nature of the estimator requires considering local log likelihoods. We start with a guess for β and $m\left(\cdot\right)$. Given the guess in β, I update function $m\left(\cdot\right)$ using local-information for each observation. The size and weighting of this local-information is given by the kernel. Once m is updated, I calculate the log–likelihood and derivatives which are then used in the next iteration to solve for a new β . The procedure ends when the parameter β converges. The reader can follow Müller (2001) and Langrock, Heidenreich, and Sperlich (2014) for a detailed exposition.

C.1 Results

I estimate the following specification

$$\mathbb{P}\left[d_i^* = o|o, i\right] = \frac{\exp\left(\alpha \text{Age}_i + \gamma \text{Sex}_i + \beta \text{Bonus}_i + m\left(\boldsymbol{x}_{i,o}\right)\right)}{1 + \exp\left(\alpha \text{Age}_i + \gamma \text{Sex}_i + \beta \text{Bonus}_i + m\left(\boldsymbol{x}_{i,o}\right)\right)}. \tag{14}$$

I estimate this model only using non-urban kindergarten qualified temporary teachers for 2014, 2015 and 2016. I use half-bandwidths of 5, 10 and 20 km. The results show the positive effect of the bonus after the reform and negative signs before the implementation of the bonus structure. In this setup, the discontinuity should be the main source of variation that identifies β. We should expect then that as the bandwidth increases we start to compare units farther apart and affecting the estimates. This does happen and as I increase the half-bandwidth, the estimate for β decreases.

To check compare my results, I provide a statistics that is comparable to the Isolated/Rural boundary ATET$_{\mathcal{B}}$estimates for Kindergarten teachers. To so I calculate the marginal effect of a S/400 bonus increase, which is the salary change at the Isolated/Rural boundary. The values are shown in the last row of the Table 17. This estimated marginal effect is consistent with the previous results using matching estimators. Using the 5 km half-bandwidth estimates in 2016, the probability of retaining a teacher an extra year increases by 5.4 percentage points.

I plot the estimated function $m(\cdot)$ on the "longitude-latitude-m" space. Since $\hat{m}$ allows me to recover $\tilde{u}_o$ for each coordinate I can calculate the changes in retention across space under a different bonus structure.

I show the use of this method by calculating the change in the retention rate under an increase of S/400 to Isolated areas. I show a the map of Peru and calculate the change of retention rate across space for each temporary kindergarten teacher, which I denote by $\Delta_{i,400}$, for Isolated schools. I calculate $\Delta_{i,400}$ as follows:

$$\begin{aligned}
&\Delta_{i,400} \\
&\equiv \hat{\mathbb{P}}\left[d_i^* = o | o, \text{New Bonus} = \text{Bonus}_i + 400, i = \text{Isolated}\right] \\
&\quad - \hat{\mathbb{P}}\left[d_i^* = o | o, i = \text{Isolated}\right]
\end{aligned}$$

The estimates show an upper bound for this increment of 2 percentage points in retention, but some teachers seem to be unaffected by the change. The national average change of $\Delta_{i,400}$ is 1.7 percentage points with a standard deviation of 0.4 percentage points.

C.2 Limitations

The results are no definite but instead provide a potential extensions to improve the study of teacher migration decisions in general. The current framework and its counterfactual are limited in scope for a couple reasons. First, the results only look at retention. From the discrete choice model, I derived $\tilde{u}_o$ as a function of the outside options. Any change in the the bonus schedule would have effects on $\tilde{u}_o$. The counterfactual results I provided assume that changes in B do not affect $\tilde{u}_o$. To have a correct assessment of retention requires estimating a full discrete choice model. This is done by expanding the implementation of Langrock, Heidenreich, and Sperlich (2014) to a multinomial setting with heterogeneous choice sets. The multinomial part is natural in migration problems. The heterogeneous choice set is need to make the the estimation feasible.[41] Second, even after implementing the correct discrete choice model, studying migration patterns in a closed system as the public education system requires taking into account "general equilibrium" effects. Theoretically, there is a value under which permanent teacher will start moving towards remote areas. This will create congestion at certain schools and automatically increase the turnover of temporary teachers because lower priority than permanent teachers.

D Figures

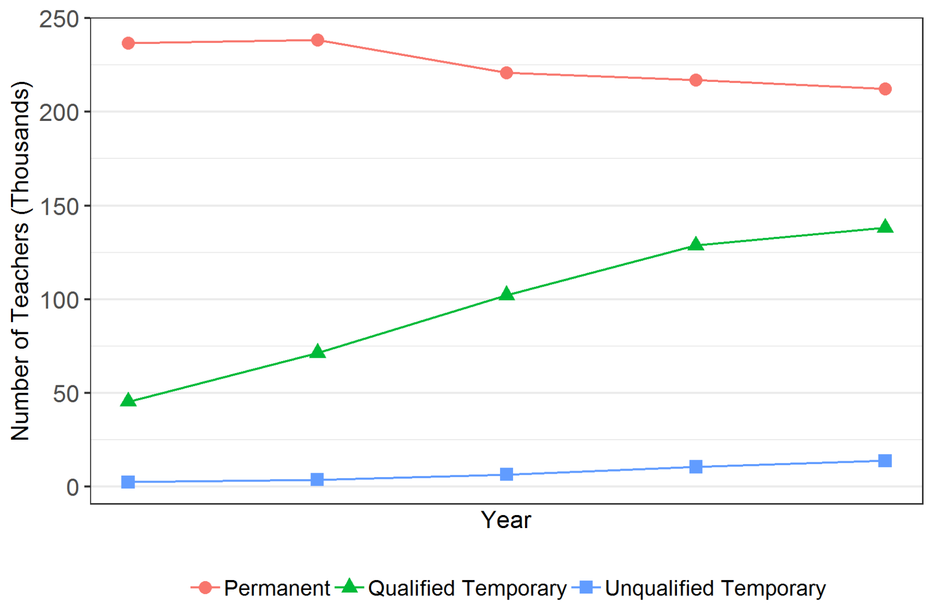
[41] Kennan and Walker (2011) provide a manner to reduce the choice set in a dynamic discrete choice setting.

E Teacher Retention at the School Level

F Teacher Composition

G Tables

Figure 1: Number of Teachers by Contract and Qualification



Source: Ministry of Education, administrative teacher data combined with School Census (Censo Escolar).

Figure 2: Rural Classification by Traveling Time to Provincial Capital and Town Population

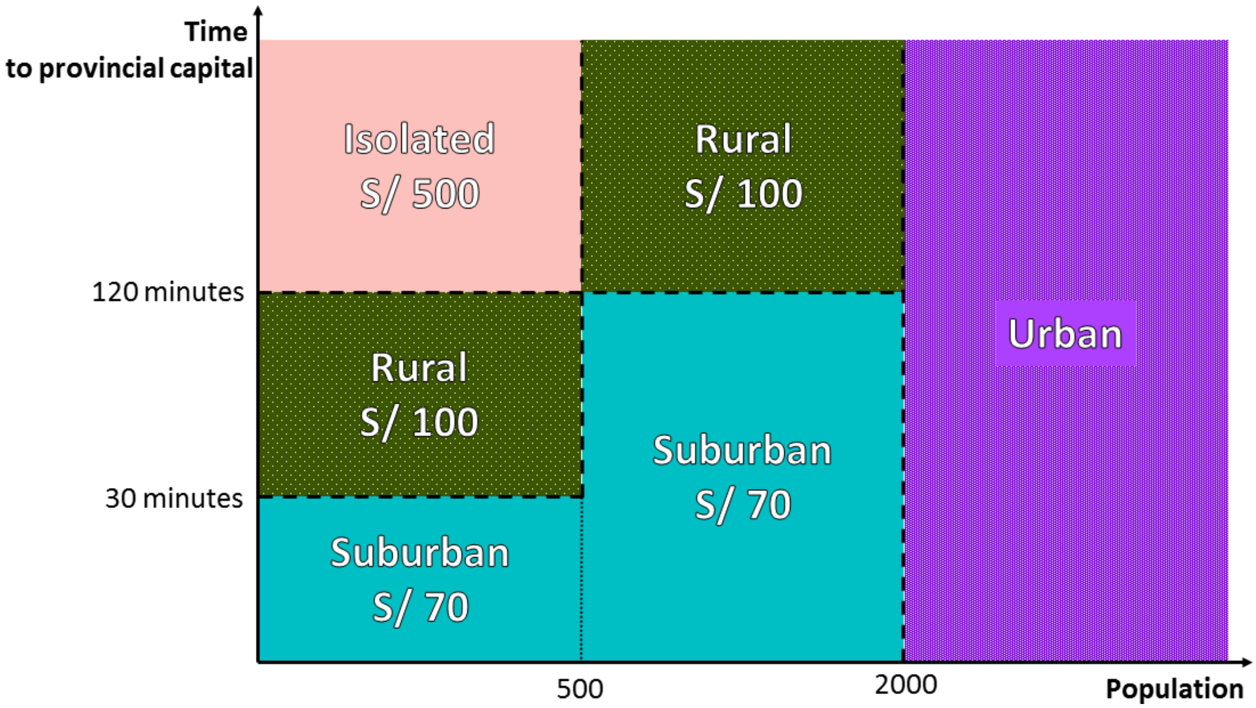
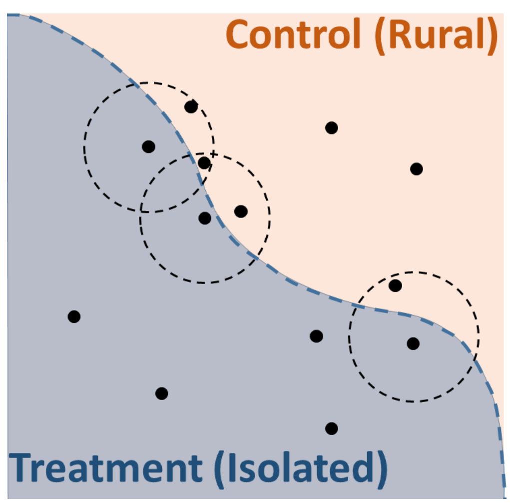


Figure 3: Multivariate-Matching under an unknown frontier



Note: Each dot represents a school. The radius of each circle represents the caliper or the half-bandwith when using a Epanechnikov kernel.

Figure 4: Epanechnikov kernel with 1 km, 5km and 10 km half-bandwidths

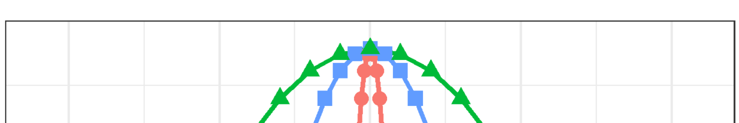
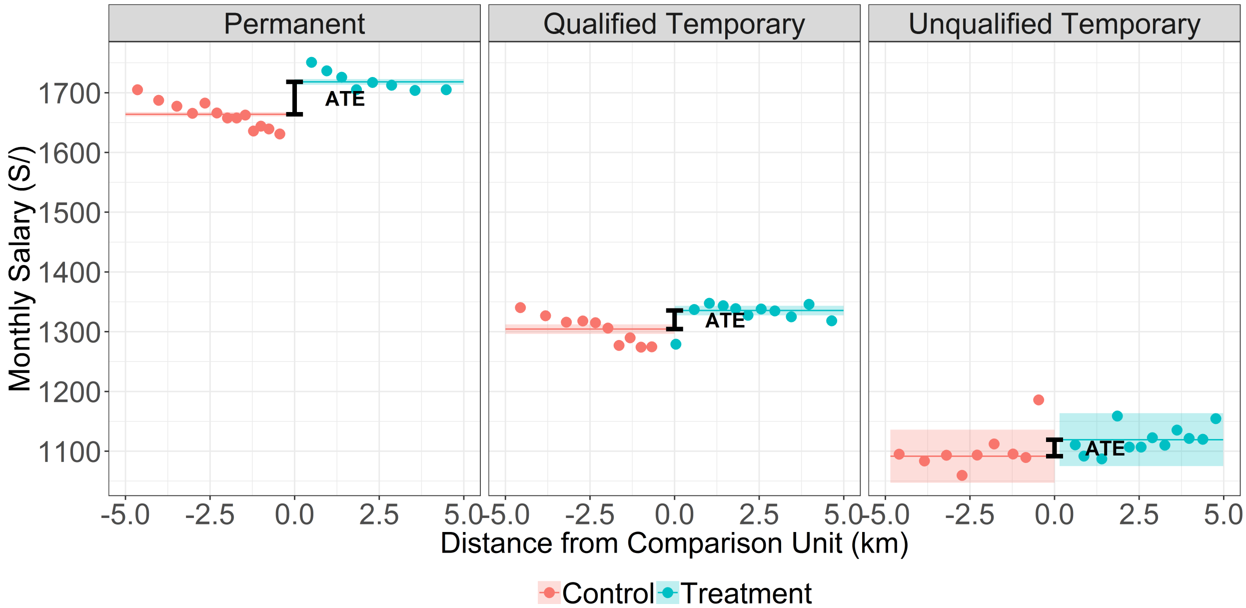


Figure 5: Monthly Salary at the Isolated/Rural Boundary Discontinuity in 2013 (Pre-reform)



Note: In this year there is no bonus schedule in place in 2013.

Figure 6: Monthly Salary at the Isolated/Rural Boundary Discontinuity in 2016

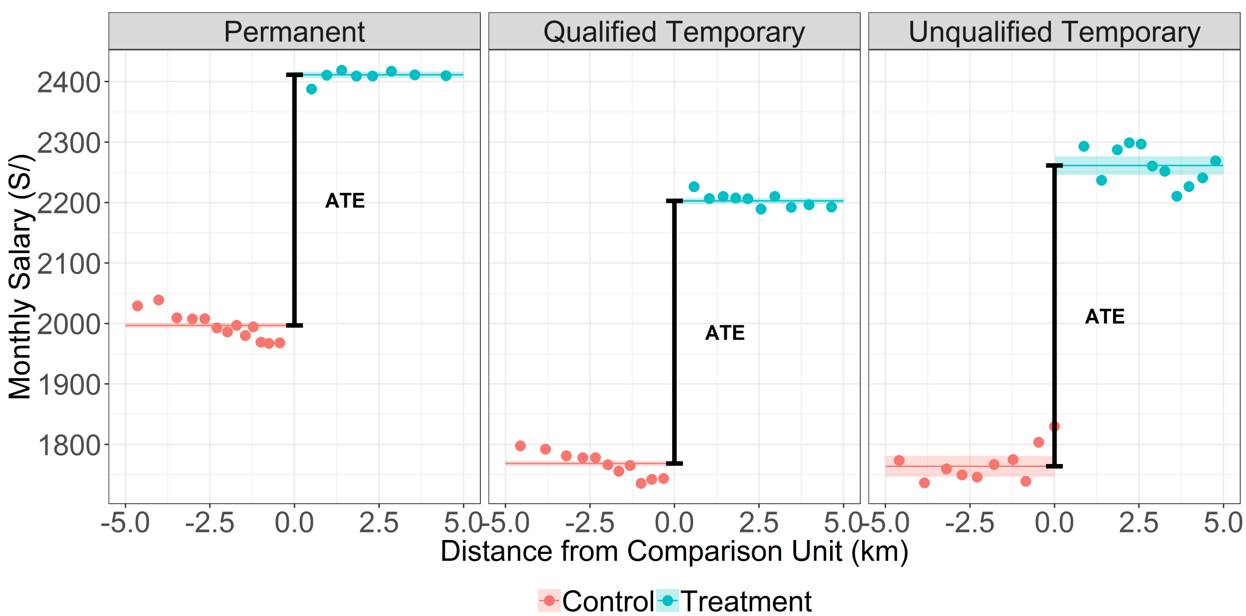


Figure 7: Monthly Salary at the VRAEM Boundary Discontinuity in 2016

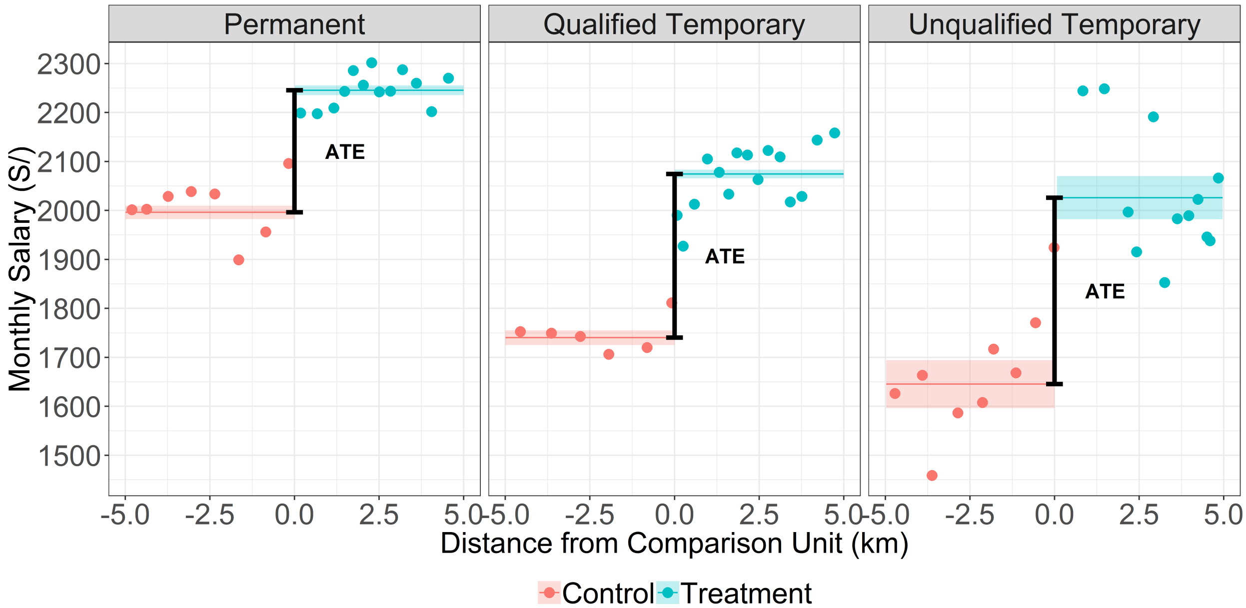


Figure 8: Monthly Salary at the Frontier Boundary Discontinuity in 2016

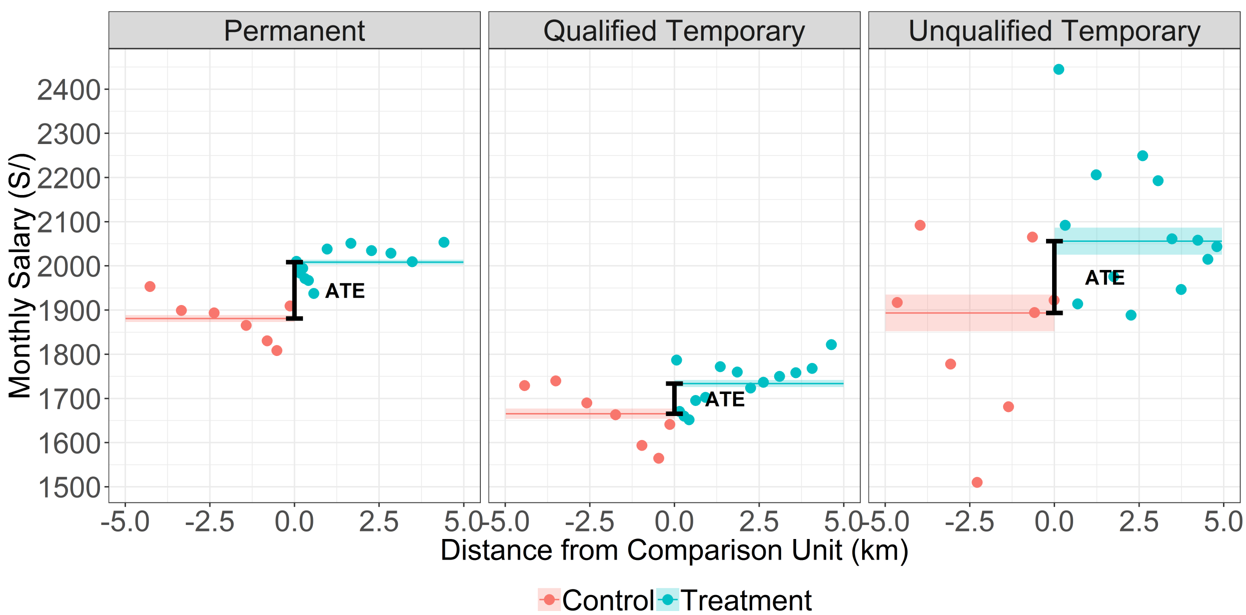


Figure 9: Teacher Retention at the Isolated/Rural in 2013

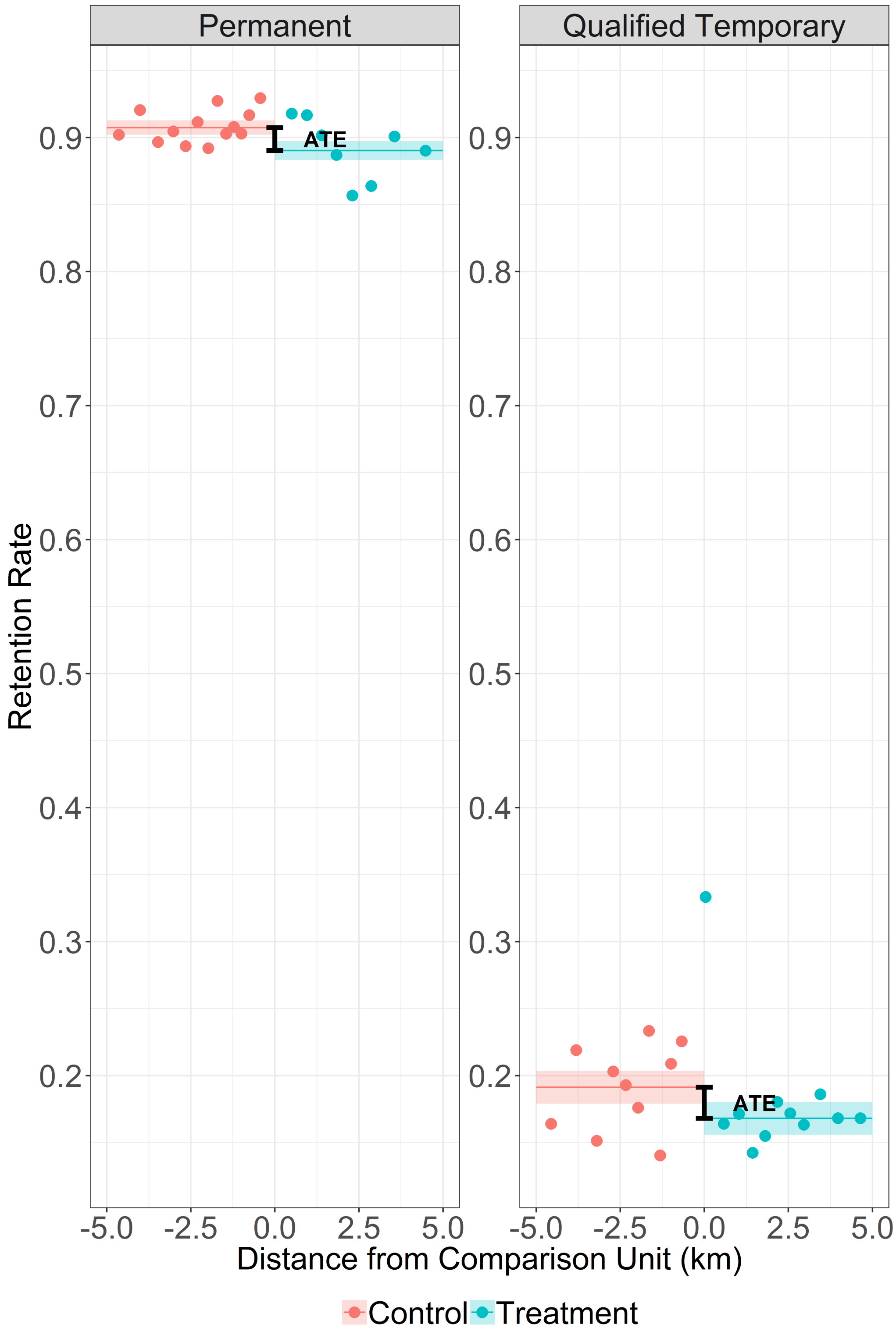


Figure 10: Teacher Retention at the Isolated/Rural in 2016

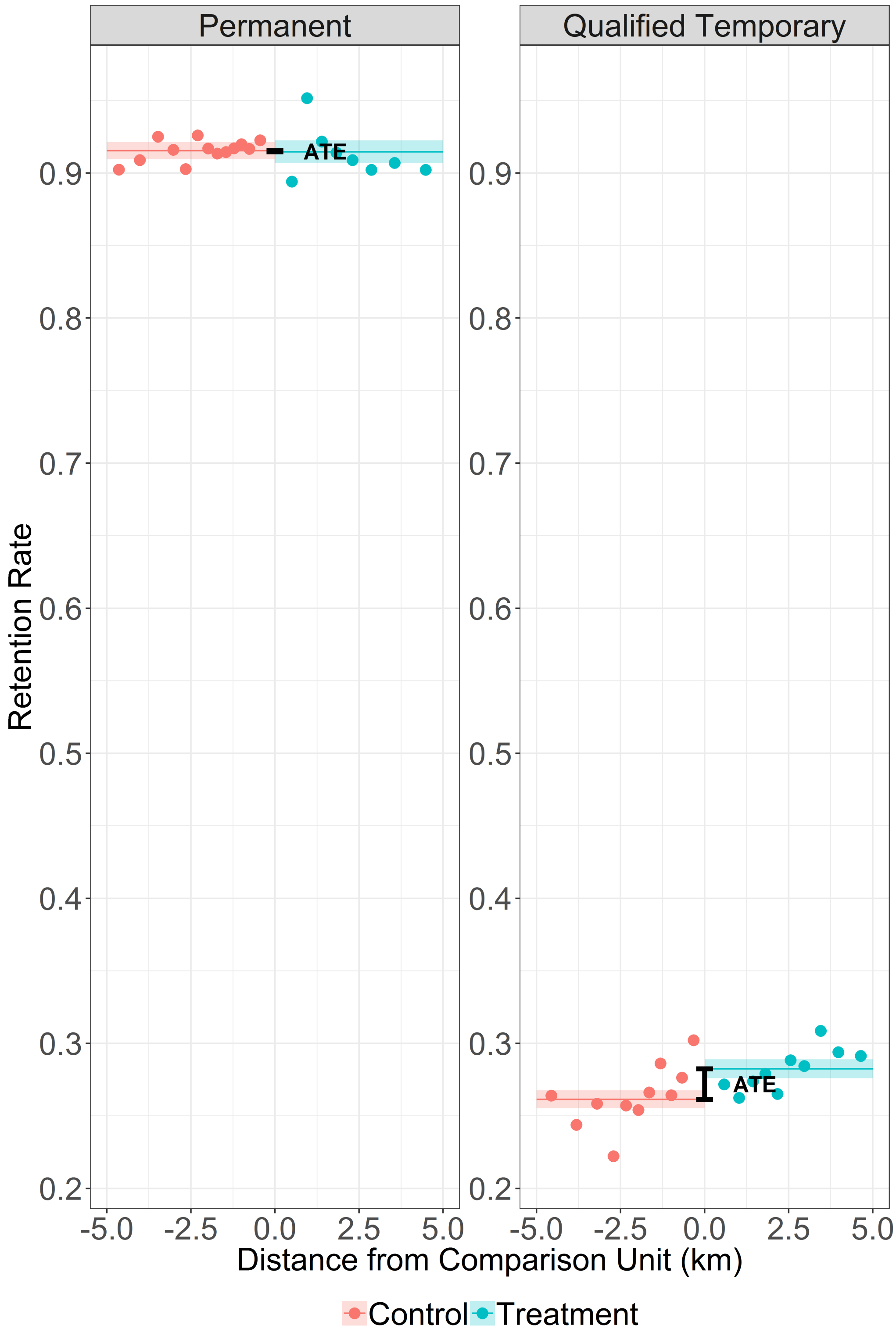


Figure 11: Transition from Non-urban Schools

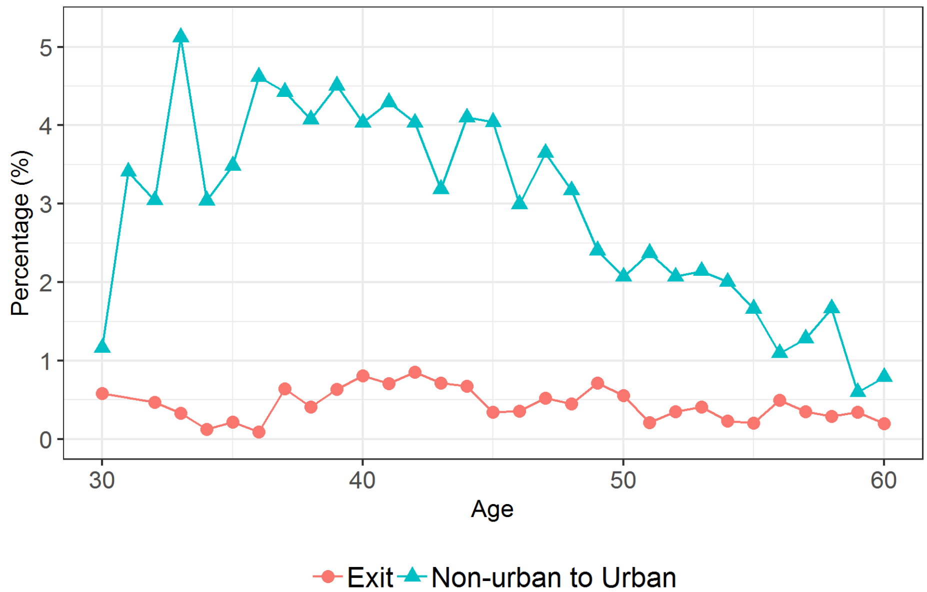


Figure 12: Transition from Urban Schools

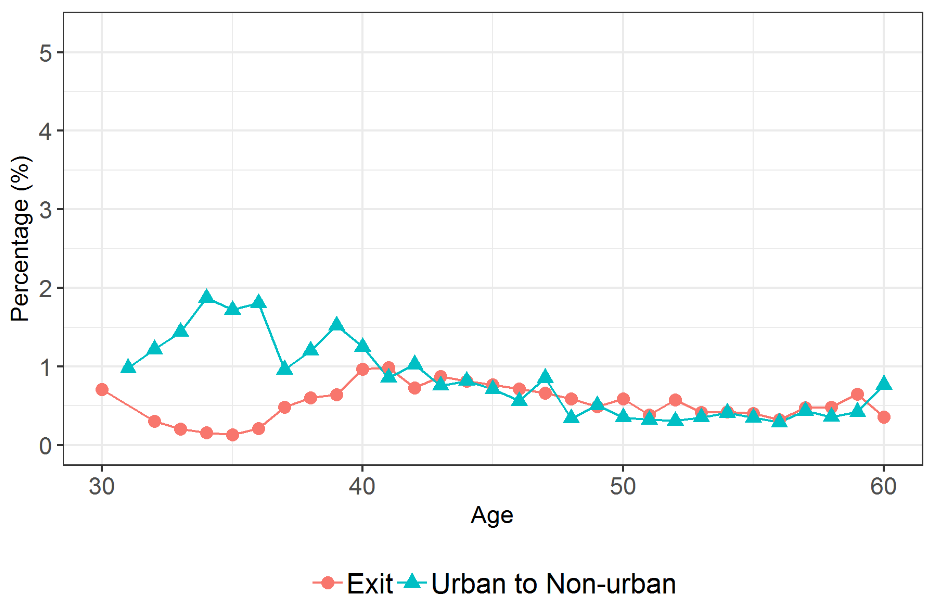


Figure 13: Transition from Urban Schools

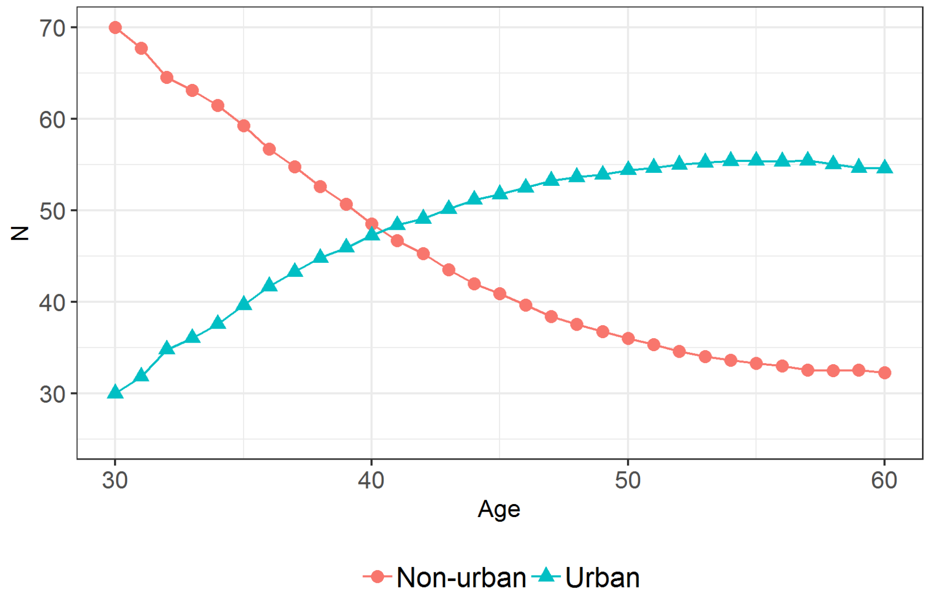
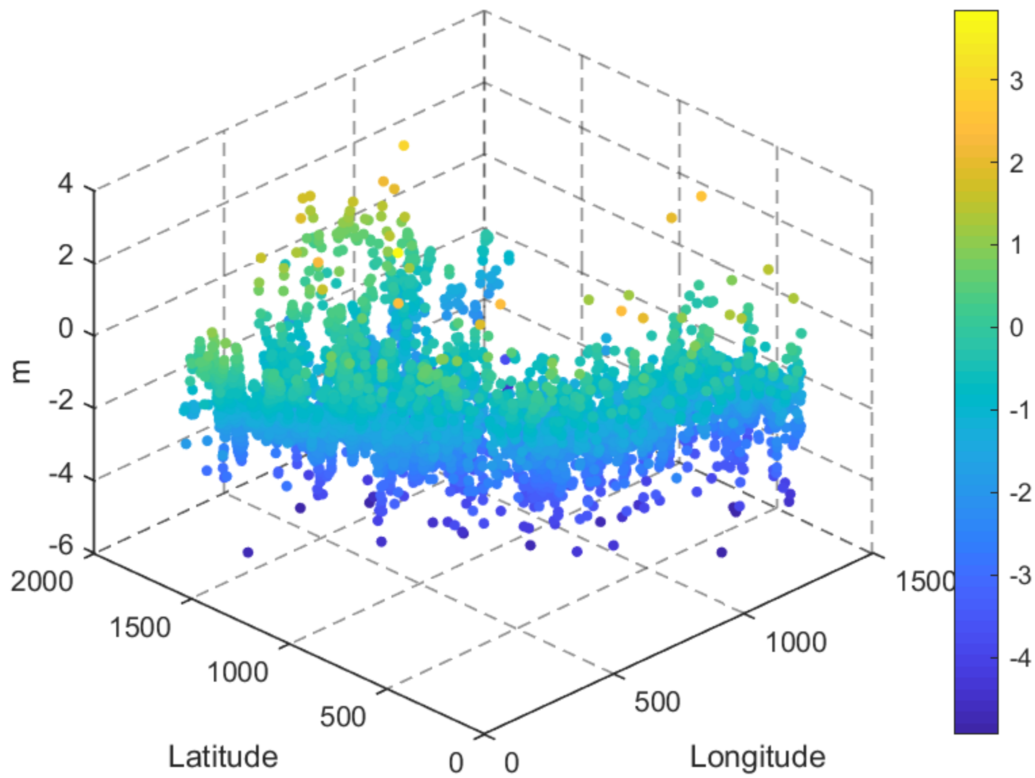
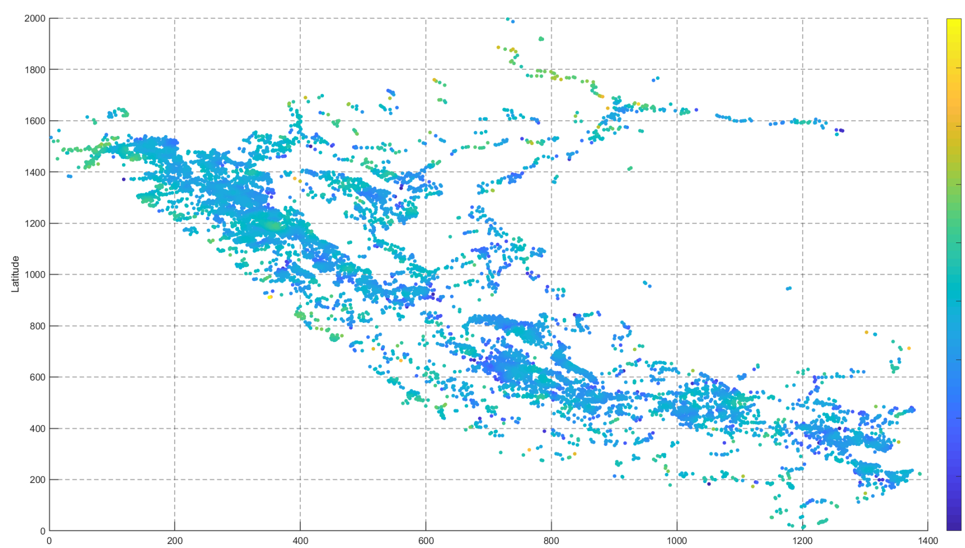


Figure 14: Visualization of estimated function $\hat{m}(\cdot)$



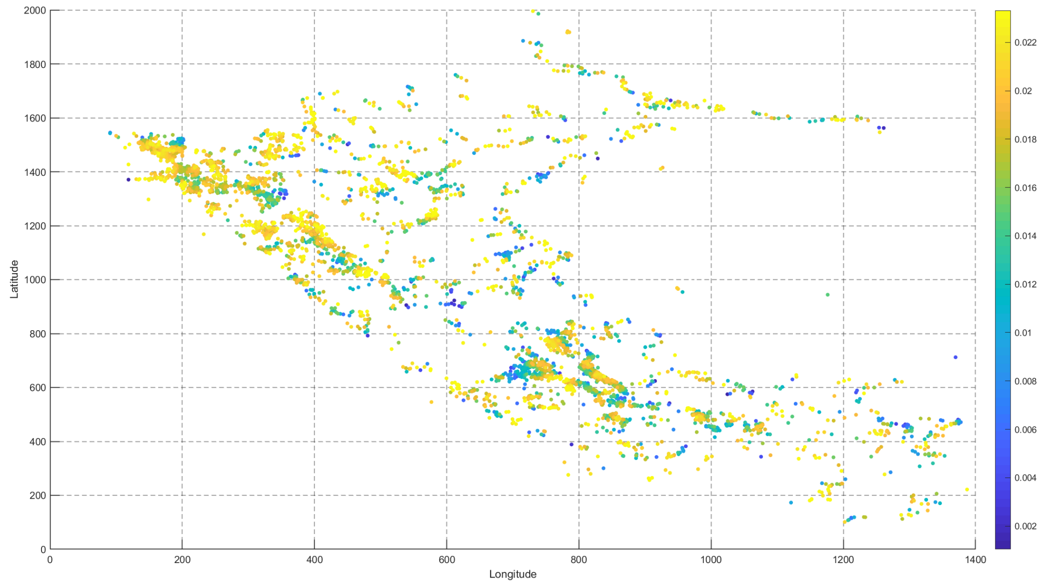
(a) Diagonal Perspective



(b) Top Perspective

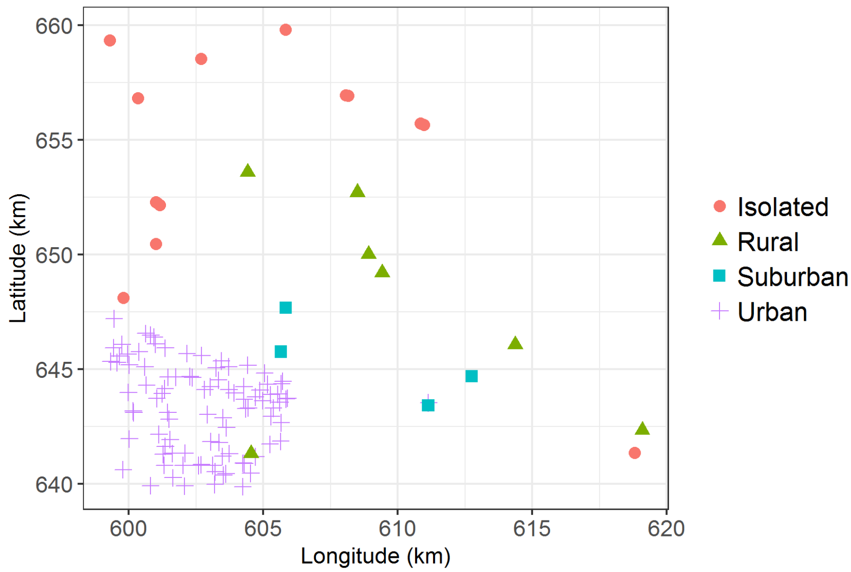
Notes: $\hat{m}$ was obtained by estimating equation (14) with a 5 km half-bandwidth and a Epanechnikov kernel. Each dot represents a non-urban school. The x and y-axis are latitude and longitude

Figure 15: Change in Retention rates in 2016

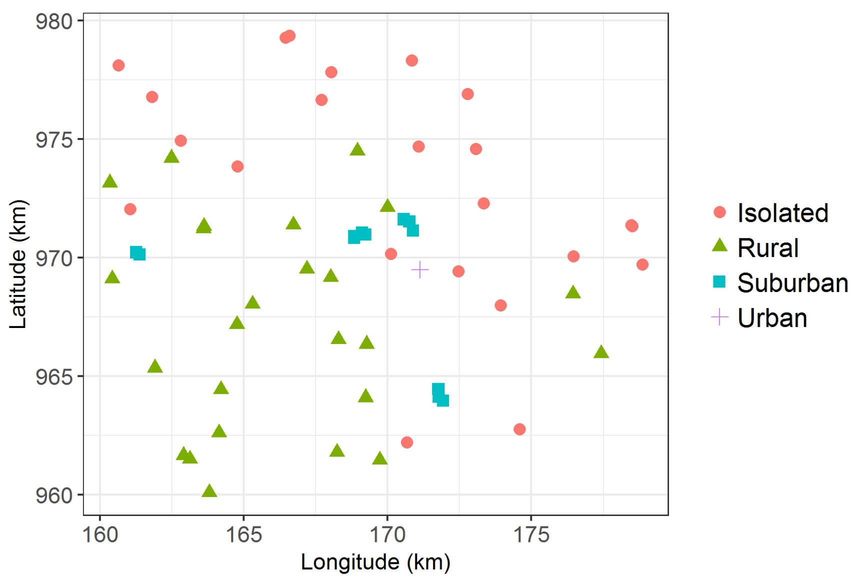


Notes: Each dot is a temporary kindergarten teacher (I take into account age and sex) and the color of the dot represents the estimated change in retention under a S/400 increment.

Figure 16: Examples of Non-Urban/Urban Bonus Classification



(a) First example



(b) Second example

Notes: Each symbol represents a school. The classification algorithm is not perfect and some schools might be misclassified (e.g., rural school surrounded by urban schools in the lower left part of the first figure), but this seems to be rare.

Figure 17: Graphical representation of vectors at $\mathcal{B}$

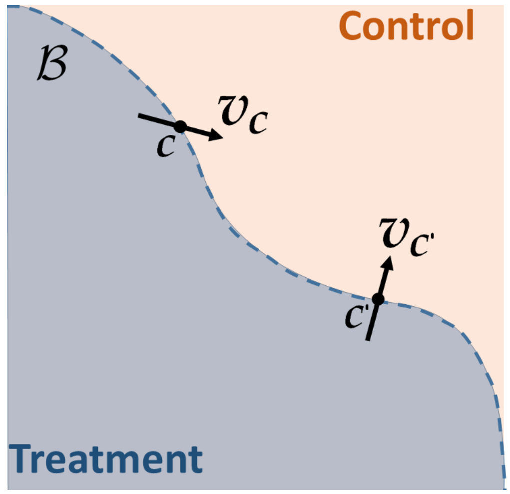
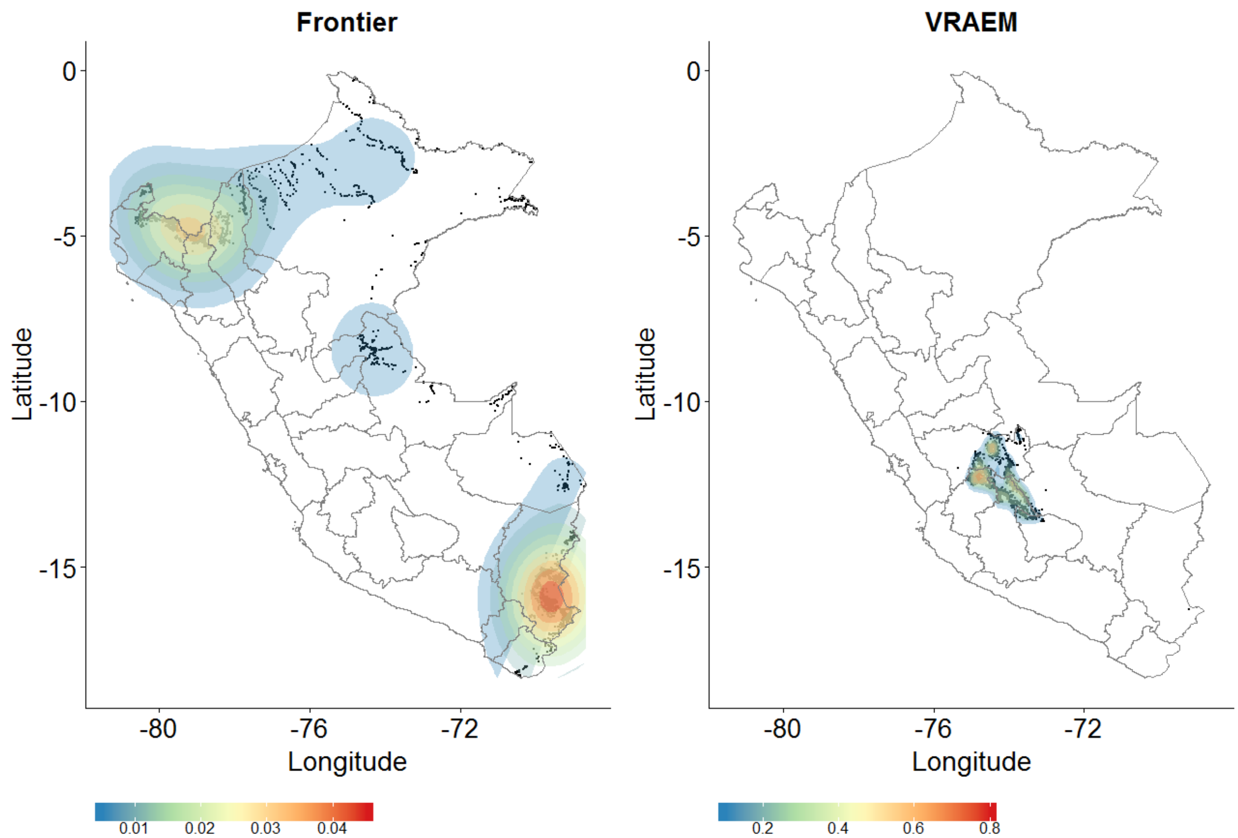
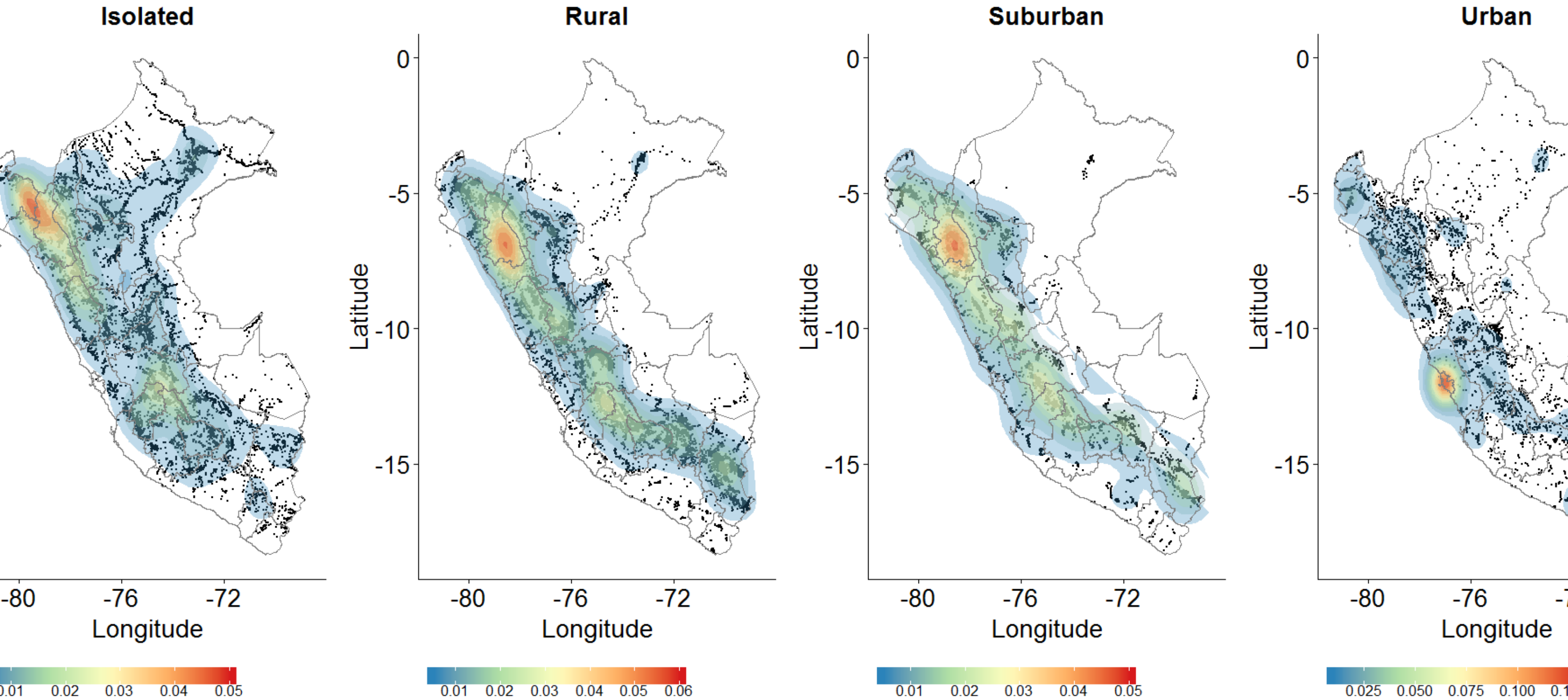


Figure 18: Spatial Distribution (Concentration) of Schools by Frontier and VRAEM classification



Source: Ministry of Education.
Note: Each dot represents a school. The heatmap represents the density of schools across space. Red represents high concentration of schools while light represents low concentration of schools.

Figure 19: Spatial Distribution (Concentration) of Schools by Non-urban Classification



Source: Ministry of Education. *Note*: Each dot represents a school. The heatmap represents the density of schools across space. Red represents high concentration of schools while light represents low concentration of schools.

Figure 20: Teacher Retention at the Frontier boundary in 2013

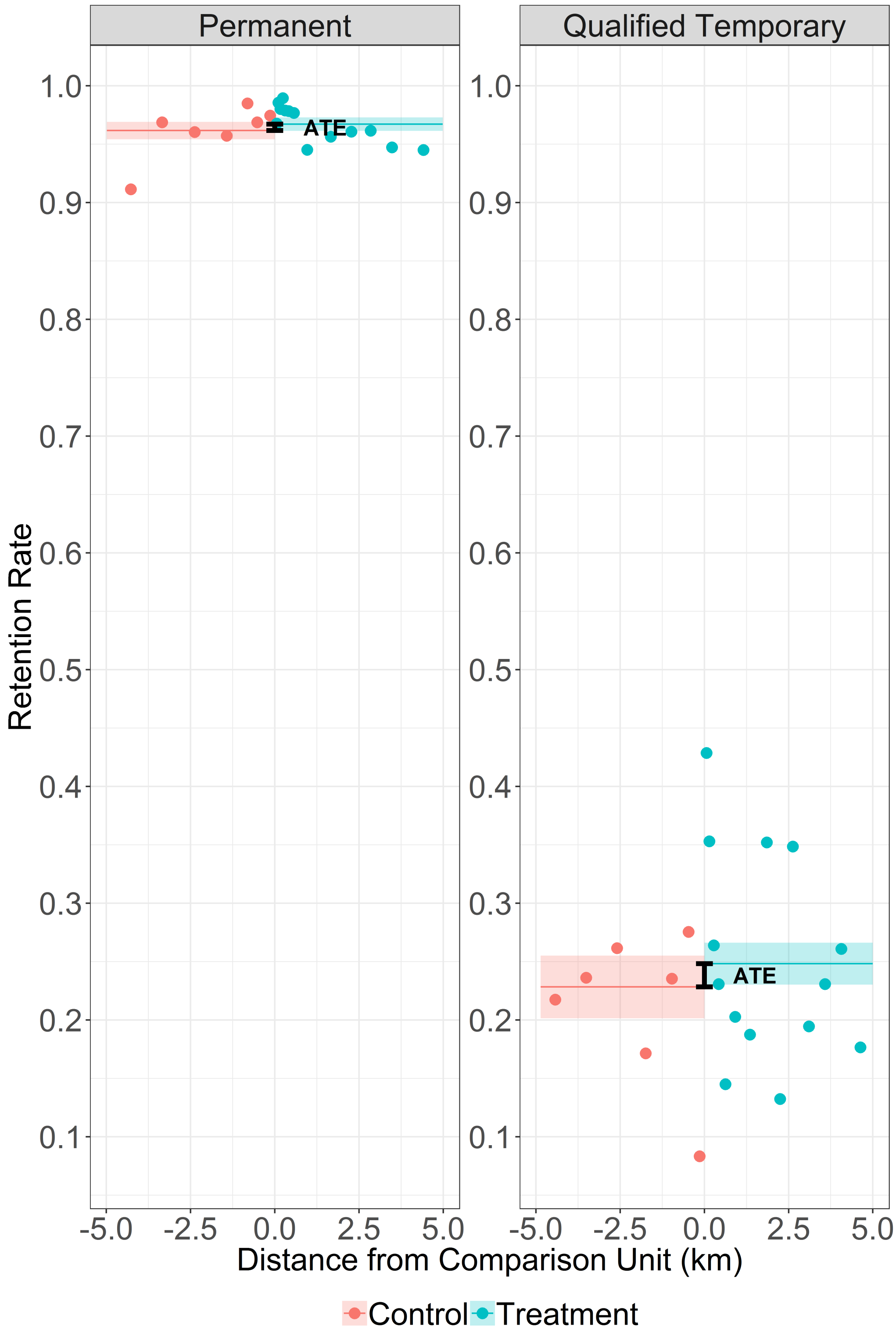


Figure 21: Teacher Retention at the Frontier Bonus boundary in 2016

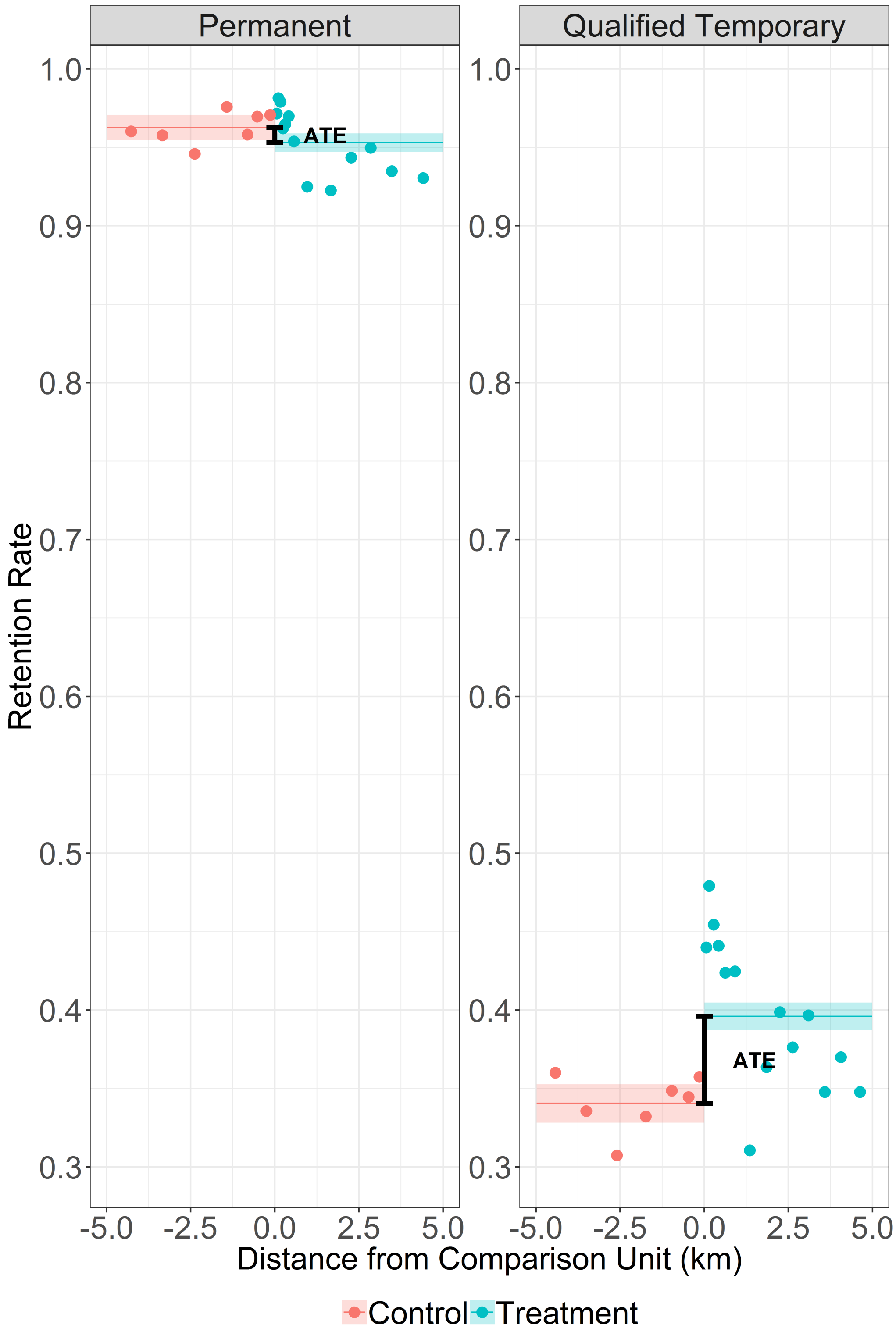


Figure 22: Teacher Retention at the VRAEM in 2013

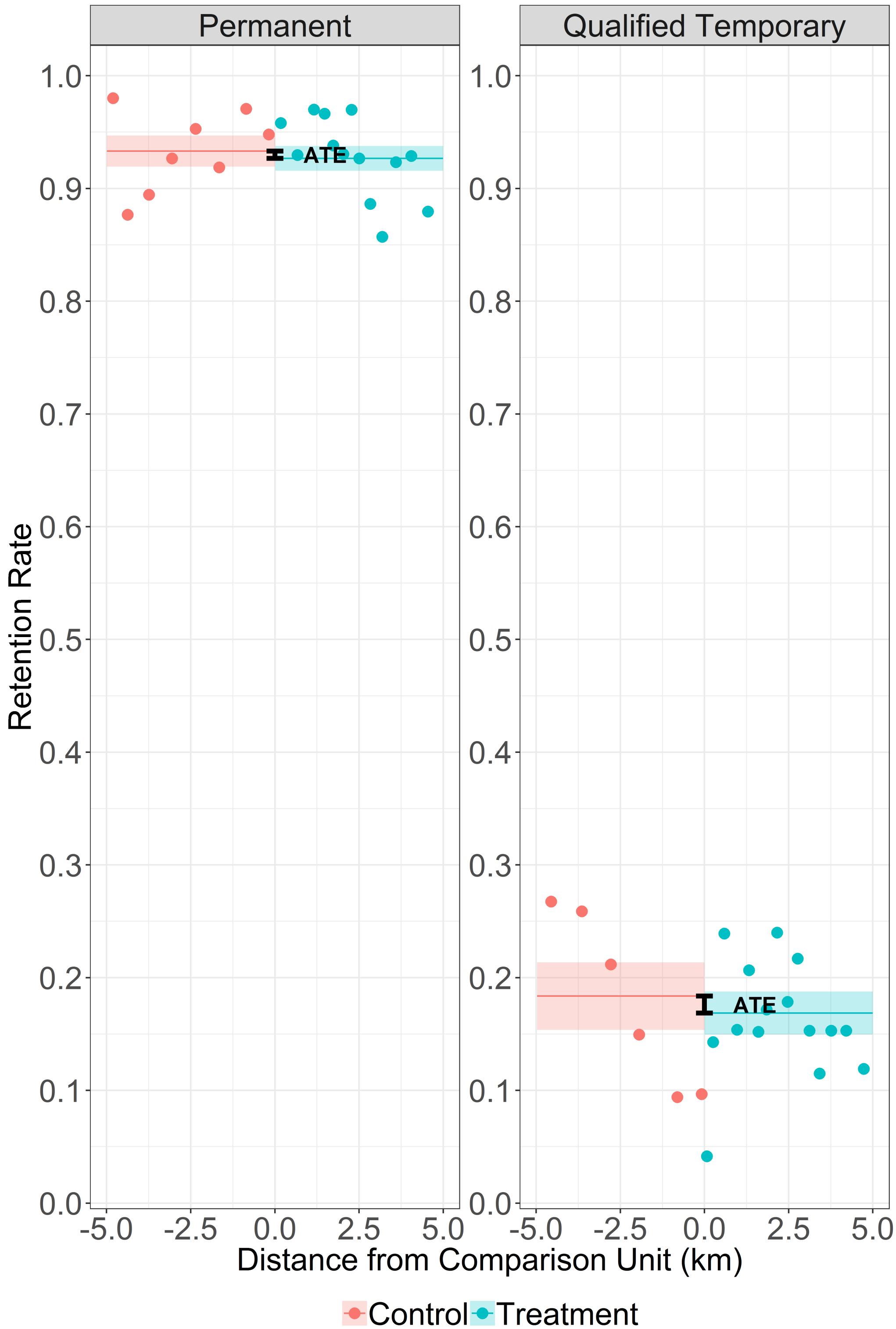


Figure 23: Teacher Retention at the VRAEM bonus boundary in 2016

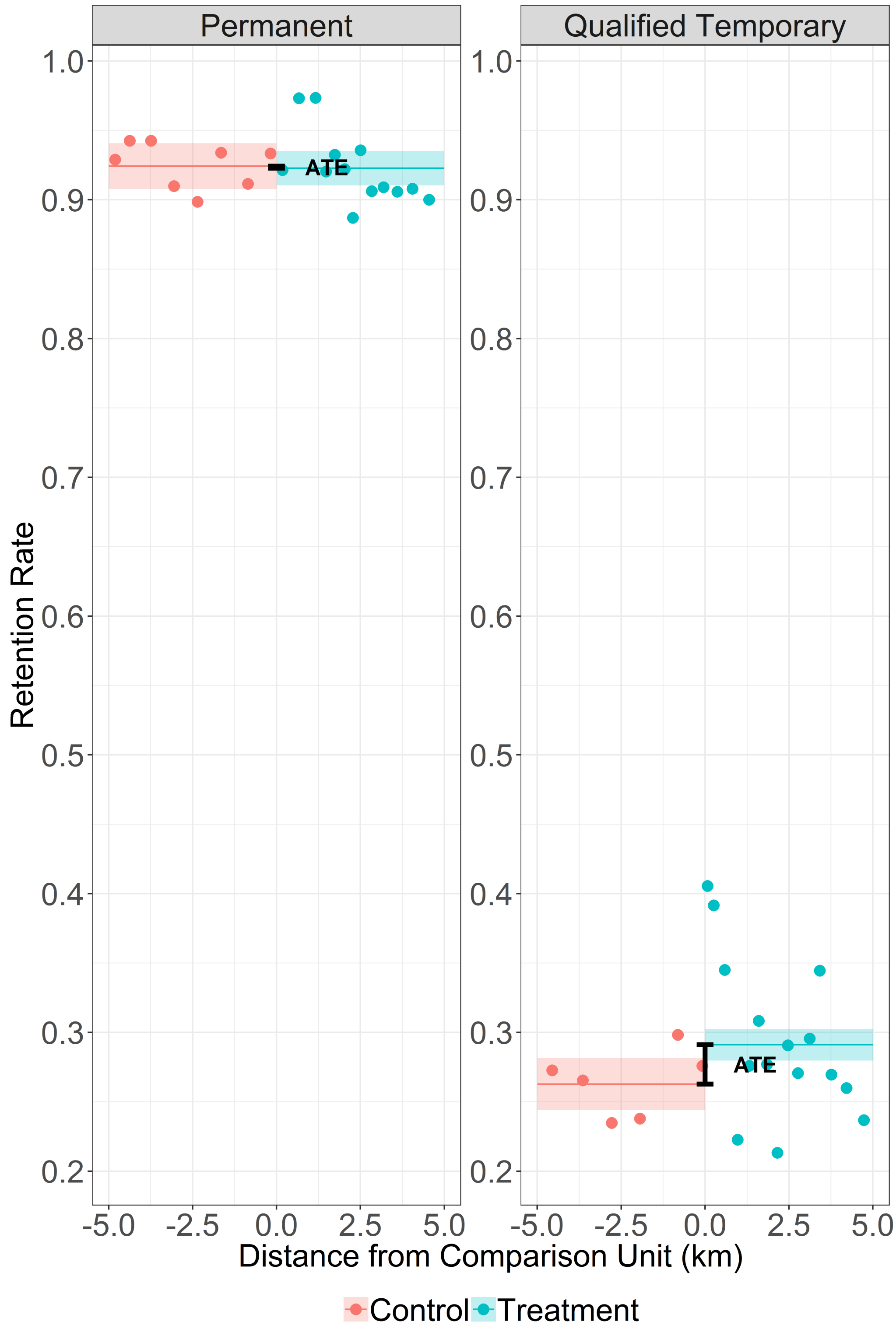


Table 1
Monthly Base Salary Schedule and Teacher Distribution as of 2017

(1) Step	(2) Monthly Base Salary (S/)	(3) % of First Step	(4) Distribution of Permanent Teachers	(5) Distribution of Temporary Teachers
First	1,781	100%	34%	100%
Second	1,959	110%	29%	0%
Third	2,137	120%	20%	0%
Fourth	2,315	130%	11%	0%
Fifth	2,671	150%	4%	0%
Sixth	3,116	175%	1%	0%
Seventh	3,383	190%	0%	0%
Eight	3,739	210%	0%	0%

Source: Ministry of Education, administrative teacher data combined with School Census (Censo Escolar).

Table 2
Monthly Salary Bonus Structure

Bonus	S/	% of First Step Base Salary	% of Fifth Step Base Salary
Non-urban Classification			
Isolated	500	28.1%	18.7%
Rural	100	5.6%	3.7%
Suburban	70	3.9%	2.6%
VRAEM	300	16.8%	11.2%
Frontiers	100	5.6%	3.7%

Table 3
Elementary School Summary Statistics - 2016

	(1) Isolated	(2) Rural	(3) Suburban	(4) Urban	(5) VRAEM	(6) Frontier
Town's population	191.25	294.09	602.76	96,862.75	664.72	6,241.82
	(637.6)	(1307.4)	(1104.5)	(171590.2)	(2829.3)	(23139.0)
Town with electric network	0.55	0.80	0.92	0.94	0.74	0.68
	(0.48)	(0.38)	(0.25)	(0.20)	(0.40)	(0.44)
Town with water network	0.44	0.62	0.74	0.88	0.52	0.51
	(0.46)	(0.44)	(0.39)	(0.26)	(0.42)	(0.46)
Town with sanitation network	0.14	0.23	0.40	0.79	0.23	0.23
	(0.32)	(0.40)	(0.45)	(0.35)	(0.37)	(0.39)
Fraction of Permanent Teachers	0.63	0.74	0.81	0.80	0.52	0.63
	(0.43)	(0.38)	(0.33)	(0.32)	(0.43)	(0.43)
Fraction of Qualified Temp. Teachers	0.31	0.25	0.19	0.19	0.43	0.28
	(0.41)	(0.37)	(0.33)	(0.31)	(0.42)	(0.39)
Fraction of Unqualified Temp. Teachers	0.06	0.01	0.00	0.00	0.06	0.09
	(0.22)	(0.096)	(0.036)	(0.047)	(0.21)	(0.27)
Retention Rate of Perm. Teachers	0.88	0.90	0.92	0.94	0.89	0.93
	(0.29)	(0.25)	(0.21)	(0.17)	(0.26)	(0.20)
Retention Rate of Qual. Temp. Teachers	0.12	0.12	0.12	0.16	0.12	0.16
	(0.30)	(0.28)	(0.27)	(0.30)	(0.28)	(0.32)
Retention Rate of Unq. Temp. Teachers	0.21	0.19	0.21	0.26	0.17	0.21
	(0.38)	(0.35)	(0.38)	(0.43)	(0.36)	(0.38)
Students to Teacher Ratio	18.87	18.28	20.13	24.81	18.38	19.44
	(11.8)	(10.6)	(10.8)	(11.2)	(9.82)	(11.9)
Number of Teachers	2.16	2.89	4.30	10.84	3.54	4.63
	(1.88)	(2.61)	(3.85)	(13.1)	(5.11)	(7.29)
Reading Score	480.21	498.33	511.96	543.04	489.17	500.72
	(71.1)	(60.8)	(55.4)	(52.4)	(59.0)	(71.2)
Math Score	479.22	496.40	507.51	534.30	481.89	499.87
	(84.0)	(80.4)	(75.0)	(69.0)	(74.1)	(84.6)
Fraction of Students Repeting Grades	0.14	0.14	0.14	0.15	0.14	0.14
	(0.14)	(0.14)	(0.15)	(0.16)	(0.14)	(0.15)
Schools	10,831	12,347	7,996	13,641	2,110	3,196
Permanent Teachers	14,850	27,262	28,983	131,038	4,475	11,687
Temporary Teachers	8,569	8,401	5,367	16,776	2,984	3,095

Source: Ministry of Education, administrative teacher dataset combined with School Census (Censo Escolar) data.
Notes: Unit of observation is a school. Each cell is an average. Standard deviations in parenthesis. Sample only includes schools at towns with less than 50,000 inhabitants. Town with electric, water of sanitation network are indicator variables equal to 1 when the service is available. Retention Rate is fraction of teachers that work at the same school in 2017. VRAEM and Frontier classification can overlap with Urban/Non-urban classification.

Table 4
Fraction of Temporary Teachers that are Qualified by Year

	(1) Isolated	(2) Rural	(3) Suburban	(4) Urban
2013	0.87	0.95	0.98	0.99
2014	0.88	0.95	0.98	0.98
2015	0.85	0.93	0.95	0.96
2016	0.84	0.90	0.93	0.95
2017	0.84	0.89	0.93	0.95

Source: Ministry of Education, administrative teacher dataset combined with School Census (Censo Escolar) data.
Notes: A qualified temporary teacher has an educational college degree. The numerator of each fraction is the number of qualified temporary teachers and the denominator is the total number of temporary teachers (qualified and unqualified).

Table 5
Teacher Composition on Standardized Reading Scores

	(1) All	(2) Isolated	(3) Rural	(4) Suburban	(5) Urban	(6) VRAEM	(7) Frontier
Fraction of Unq. Temp.	-12.9**	-17.7**	-9.75	-0.59	5.04	-24.1	-22.4*
	(4.60)	(5.50)	(10.9)	(16.5)	(13.7)	(17.4)	(10.9)
Fraction of Qual. Temp.	-2.74	1.54	-1.11	3.92	-9.99**	0.47	-2.06
	(1.75)	(3.18)	(2.74)	(4.26)	(3.28)	(9.98)	(8.33)
Constant	539.1***	505.0***	523.3***	539.5***	575.8***	524.8***	535.0***
	(0.34)	(1.12)	(0.54)	(0.52)	(0.42)	(3.03)	(2.03)
Obs	53802	11023	14811	10359	16968	1860	2994
R^2	0.71	0.64	0.61	0.62	0.77	0.65	0.75

Source: Ministry of Education, Students Census Evaluation, School Census, administrative teacher dataset.
Notes: School-district clustered standard errors. * p<0.05, ** p<0.01, *** p<0.001. Dependent variable: Standardized reading score of second grade students (school average). All regressions span from 2013 to 2016 and include school fixed-effects and department×year fixed-effects.

Table 6
Teacher Composition on Standardized Math Scores

	(1) All	(2) Isolated	(3) Rural	(4) Suburban	(5) Urban	(6) VRAEM	(7) Frontier
Fraction of Unq. Temp.	-19.7**	-17.3*	-13.7	-7.99	-8.88	-32.6	-19.6
	(6.83)	(8.36)	(16.3)	(25.3)	(19.0)	(29.0)	(17.3)
Fraction of Qual. Temp.	-7.21**	2.11	-3.72	1.30	-15.7**	-6.38	-2.78
	(2.47)	(4.28)	(4.19)	(6.35)	(5.00)	(13.4)	(13.8)
Constant	543.9***	505.3***	528.1***	546.6***	580.9***	528.9***	540.4***
	(0.49)	(1.54)	(0.83)	(0.77)	(0.64)	(4.14)	(3.34)
Obs	53801	11023	14810	10359	16968	1860	2994
R^2	0.66	0.61	0.59	0.60	0.73	0.62	0.70

Source: Ministry of Education, Students Census Evaluation, School Census, administrative teacher dataset. School-district clustered standard errors. * p<0.05, ** p<0.01, *** p<0.001. Dependent variable: Standardized math score of second grade students (school average) . All regressions span from 2013 to 2016 and include school fixed-effects and department×year fixed-effects.

Table 7

	2015		2016		2017	
	(1) Openings	(2) Teacher	(3) Openings	(4) Teacher	(5) Openings	(6) Teacher
Isolated	24.7	9.4	27.3	10.2	19.7	11.8
Rural	23.8	13.9	29.2	15.1	20.3	14.5
Suburban	14.9	13.3	13.8	12.8	10.7	11.8
Urban	36.6	63.4	29.7	62.0	49.3	61.9
Total	100	100	100	100	100	100

Table 17
Semi-parametric Logit Estimates

	2014			2015			2016		
	$b = 5$	$b = 10$	$b = 20$	$b = 5$	$b = 10$	$b = 20$	$b = 5$	$b = 10$	$b = 20$
α	0.020	0.019	0.017	0.011	0.011	0.011	0.017	0.015	0.015
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	0.000)
γ	-0.113	-0.132	-0.091	-0.160	-0.118	-0.073	-0.369	-0.370	-0.363
	(0.020)	(0.015)	(0.015)	(0.013)	(0.014)	(0.013)	(0.010)	0.009)	(0.009)
β	-0.017	-0.028	-0.030	0.027	0.019	0.020	0.031	0.016	0.001
	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Marginal Effect	-0.027	-0.049	-0.056	0.035	0.026	0.028	0.054	0.029	0.001

Notes: Standard errors in parentheses. I employ an epanechnikov kernel-based matching algorithm using euclidean distance with a half-bandwidth of b km. The marginal effect in the last row is the Average Marginal Effect of a S/400 bonus.

Table 8
Treatment Effect on Retention for Teachers at Isolated/Rural Boundary

	Pre-reform		Post-reform	
	(1) 2013	(2) 2014	(3) 2015	(4) 2016
Permanent	-0.01	0.00	0.02	0.03**
	(0.01)	(0.00)	(0.01)	(0.01)
Qualified Temporary	-0.00	0.02	0.05***	0.04***
	(0.02)	(0.02)	(0.01)	(0.01)
Unqualified Temporary	0.05	-0.04	-0.00	0.01
	(0.13)	(0.09)	(0.08)	(0.05)
Permanent				
Base Retention Rate	0.91	0.97	0.82	0.91
Obs.	13,442	12,483	12,130	10,676
Qualified Temporary				
Base Retention Rate	0.19	0.28	0.16	0.25
Obs.	2,875	6,168	8,292	11,628
Unqualified Temporary				
Base Retention Rate	0.32	0.31	0.29	0.27
Obs.	114	215	401	727

Notes: Bootstrapped standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. Each coefficient is a ATE estimator. The dependent variable is an indicator variable equal to 1 if teacher stayed an extra year at the same school the same school next year. Treatment unit: teacher working at Isolated school (S/500 bonus), Control unit: teacher works at Rural school (S/100 bonus). I employ an epanechnikov kernel-based matching algorithm using euclidean distance with a 5 km half-bandwidth. I use a bias-adjustment procedure that controls for age, gender, education level of school and town characteristics (water, electricity and sanitation services).

Table 9
Treatment Effect on Staying at School for Teachers at Isolated/Rural Boundary

	2014			2015			2016		
	(1) Kinder	(2) Elem.	(3) H-S	(4) Kinder	(5) Elem.	(6) H-S	(7) Kinder	(8) Elem.	(9) H-S
Panel A. ***Permanent***									
Kernel($b = 5km$)	-0.03	0.00	0.01	0.09	0.02	0.03	-0.00	0.03**	0.03
	(0.03)	(0.01)	(0.01)	(0.08)	(0.01)	(0.03)	(0.05)	(0.01)	(0.02)
Kernel($b = 10km$)	-0.02	0.00	-0.00	0.04	0.00	0.02	0.03	0.02*	0.02
	(0.02)	(0.00)	(0.01)	(0.06)	(0.01)	(0.02)	(0.04)	(0.01)	(0.01)
Kernel($b = 10km$) + Popul. $\leq$ 500	-0.02	0.00	-0.00	-0.02	-0.00	0.01	0.05	0.02*	0.01
	(0.02)	(0.01)	(0.01)	(0.06)	(0.01)	(0.02)	(0.07)	(0.01)	(0.02)
2 Nearest-Neighbor	-0.01	-0.00	-0.00	0.03	0.00	0.01	0.00	0.02*	0.02
	(0.01)	(0.01)	(0.01)	(0.04)	(0.01)	(0.02)	(0.03)	(0.01)	(0.01)
4 Nearest-Neighbor	-0.01	0.00	0.00	-0.00	-0.01	-0.00	0.00	0.02**	0.01
	(0.01)	(0.00)	(0.01)	(0.03)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
5km Block Fixed-Effects	-0.06	0.00	-0.01	-0.04	0.01	0.07	0.11	0.02	0.02
	(0.12)	(0.01)	(0.01)	(0.15)	(0.01)	(0.04)	(0.11)	(0.02)	(0.04)
10km Block Fixed-Effects	-0.01	0.00	-0.00	0.07	-0.00	0.03	0.05	0.02*	0.02
	(0.04)	(0.01)	(0.01)	(0.07)	(0.01)	(0.02)	(0.06)	(0.01)	(0.02)
Base	0.98	0.97	0.98	0.72	0.87	0.73	0.90	0.92	0.91
N_1	544	7505	3437	541	7248	3365	496	6985	3017
N_0	823	11927	4378	817	11453	4311	814	11748	4145
Panel B. ***Qualified Temporary***									
Kernel($b = 5km$)	0.04	0.03	0.04	0.09***	0.02	0.05*	0.06**	0.04*	0.01
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Kernel($b = 10km$)	0.02	0.02	0.02	0.08***	0.01	0.04**	0.08***	0.05***	0.04*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)
Kernel($b = 10km$) + Popul. $\leq$ 500	0.01	0.04	0.03	0.09***	0.03*	0.05*	0.05*	0.07***	0.04*
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
2 Nearest-Neighbor	0.02	0.00	0.02	0.08***	0.02	0.04**	0.09***	0.05**	0.06***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
4 Nearest-Neighbor	0.02	-0.01	-0.00	0.08***	0.01	0.04***	0.08***	0.04***	0.05***
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
5km Block Fixed-Effects	0.06	0.01	0.03	0.09*	0.02	0.06	0.05	0.01	0.00
	(0.06)	(0.06)	(0.05)	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)
10km Block Fixed-Effects	0.01	-0.01	0.01	0.08**	0.02	0.04	0.04	0.03	-0.01
	(0.03)	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Base	0.25	0.23	0.34	0.15	0.12	0.20	0.23	0.25	0.30
N_1	2358	3024	3645	2574	4165	4340	3739	6036	6195
N_0	2382	2231	2987	2718	3594	3905	3954	4967	5355
Panel C. ***Unqualified Temporary***									
2 Nearest-Neighbor	-0.04	0.03	0.01	0.02	-0.10	-0.08	0.02	0.06	0.07
	(0.08)	(0.06)	(0.14)	(0.06)	(0.06)	(0.08)	(0.05)	(0.03)	(0.04)
4 Nearest-Neighbor	-0.02	0.01	0.08	0.03	-0.18***	-0.09	0.00	0.02	0.04
	(0.07)	(0.07)	(0.13)	(0.05)	(0.05)	(0.07)	(0.05)	(0.03)	(0.03)
10km Block Fixed-Effects	-0.13	-0.22	0.00	-0.02	-0.19*	0.01	-0.02	-0.08	0.03
	(0.13)	(0.17)	(.)	(0.10)	(0.08)	(0.27)	(0.09)	(0.06)	(0.07)
Base	0.30	0.36	0.38	0.34	0.26	0.30	0.27	0.23	0.31
N_1	438	340	148	562	623	391	769	841	768
N_0	112	92	31	184	183	141	290	308	532

Notes: Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. Standard errors are boostraped (200 repetitions) in kernel matching, robust in NN matching, clustered at the region level in the regressions. Each coefficient is a ATE estimator. The dependent variable is an indicator variable equal to 1 if teacher stay at the same school next year. Treatment unit: teacher working at Isolated school, Control unit: teacher works at Rural school. Both matching algorithms use a the Euclidean metric. Kernel matching uses a epanechnikov kernel function. All matching estimates use a bias correction procedure that controls for age, gender, education level of school and town characteristics (water, electricity and sanitation services). The 5 and 10 km^2 Block Fixed-Effects provide ATE estimates from a regression with fixed-effects.

Table 10
Treatment Effect on Retention for Teachers at VRAEM Boundary

	Pre-reform		Post-reform	
	(1) 2013	(2) 2014	(3) 2015	(4) 2016
Permanent	0.02	-0.01*	-0.03	0.02
	(0.02)	(0.00)	(0.03)	(0.03)
Qualified Temporary	0.04	0.14**	0.04	0.08
	(0.05)	(0.05)	(0.04)	(0.05)
Permanent				
Base Retention Rate	0.92	0.99	0.87	0.92
Obs.	2,905	2,458	3,055	2,006
Qualified Temporary				
Base Retention Rate	0.14	0.25	0.15	0.27
Obs.	805	1,614	2,086	2,256

Notes: Bootstrapped standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. Each coefficient is a ATE estimator. The dependent variable is an indicator variable equal to 1 if teacher stayed an extra year at the same school the same school next year. Treatment unit: teacher working at VRAEM school (S/300 bonus), Control unit: teacher works at non-VRAEM school (S/0 bonus). I employ an epanechnikov kernel-based matching algorithm using euclidean distance with a 5 km half-bandwidth. I use a bias-adjustment procedure that controls for age, gender, education level of school and town characteristics (water, electricity and sanitation services).

Table 11
Treatment Effect on Staying at School for Teachers at VRAEM Bonus Boundary

	2014			2015			2016		
	(1) Kinder	(2) Elem.	(3) H-S	(4) Kinder	(5) Elem.	(6) H-S	(7) Kinder	(8) Elem.	(9) H-S
Panel A. ***Permanent***									
Kernel($b = 5km$)	0.00	-0.00	-0.01	-0.13	0.01	0.06	-0.05	-0.00	-0.04
	(.)	(0.01)	(0.01)	(0.18)	(0.04)	(0.05)	(0.69)	(0.03)	(0.07)
Kernel($b = 10km$)	0.00	-0.00	-0.01	-0.04	-0.02	-0.04	-0.04	-0.03	0.02
	(.)	(0.01)	(0.00)	(0.17)	(0.03)	(0.04)	(0.11)	(0.02)	(0.05)
2 Nearest-Neighbor	-0.00	-0.01	-0.01***	-0.01	0.05	-0.12***	0.03	-0.00	0.10**
	(0.00)	(0.01)	(0.00)	(0.08)	(0.04)	(0.03)	(0.07)	(0.03)	(0.04)
4 Nearest-Neighbor	-0.00	-0.01	-0.01**	-0.00	0.02	-0.09**	0.11	0.01	0.08*
	(0.00)	(0.01)	(0.00)	(0.08)	(0.04)	(0.03)	(0.06)	(0.03)	(0.04)
Base	0.99	0.98	0.99	0.90	0.89	0.81	0.93	0.93	0.93
N_1	187	2524	1312	195	2500	1344	211	2436	1317
N_0	441	1965	617	457	1914	619	449	1899	567
Panel B. ***Qualified Temporary***									
Kernel($b = 5km$)	-0.05	0.04	0.24**	0.09*	0.03	0.10	0.13*	-0.07	-0.12
	(0.10)	(0.10)	(0.08)	(0.04)	(0.06)	(0.06)	(0.06)	(0.12)	(0.15)
Kernel($b = 10km$)	-0.03	0.06	0.10	0.09**	0.04	0.05	0.12**	-0.03	-0.06
	(0.08)	(0.06)	(0.07)	(0.03)	(0.05)	(0.05)	(0.05)	(0.09)	(0.09)
2 Nearest-Neighbor	-0.05	0.07	0.10	0.09*	0.10*	0.05	0.09*	-0.03	0.12**
	(0.07)	(0.06)	(0.06)	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.04)
4 Nearest-Neighbor	-0.06	0.06	0.10*	0.06	0.02	0.04	0.08*	-0.04	0.11**
	(0.06)	(0.05)	(0.05)	(0.03)	(0.04)	(0.05)	(0.04)	(0.05)	(0.04)
Base	0.24	0.22	0.30	0.12	0.16	0.20	0.22	0.27	0.31
N_1	742	788	1336	840	1173	1652	1041	1443	2067
N_0	459	381	413	559	549	552	655	780	726

Notes: Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. Standard errors are boostraped (200 repetitions) in kernel matching, robust in NN matching, clustered at the region level in the regressions. Each coefficient is a ATE estimator. The dependent variable is an indicator variable equal to 1 if teacher stay at the same school next year. Treatment unit: teacher working at VRAEM school, Control unit: teacher works at Non-VRAEM school. Both matching algorithms use a the Euclidean metric. Kernel matching uses a epanechnikov kernel function. All matching estimates use a bias correction procedure that controls for age, gender, education level of school and town characteristics (water, electricity and sanitation services). The 5 and 10 km^2 Block Fixed-Effects provide ATE estimates from a regression with fixed-effects.

Table 12
Treatment Effect on Retention for Teachers at Frontier Boundary

	Pre-reform		Post-reform	
	(1) 2013	(2) 2014	(3) 2015	(4) 2016
Permanent	0.01	0.05	-0.00	-0.00
	(0.01)	(0.03)	(0.01)	(0.01)
Qualified Temporary	0.05	0.08*	0.04	0.12***
	(0.06)	(0.03)	(0.02)	(0.03)
Permanent				
Base Retention Rate	0.97	0.99	0.94	0.97
Obs.	10,030	9,667	10,064	9,023
Qualified Temporary				
Base Retention Rate	0.22	0.30	0.21	0.35
Obs.	628	1,571	3,148	3,125

Notes: Bootstrapped standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. Each coefficient is a ATE estimator. The dependent variable is an indicator variable equal to 1 if teacher stayed an extra year at the same school the same school next year. Treatment unit: teacher working at Frontier school (S/100 bonus), Control unit: teacher works at non-Frontier school (S/0 bonus). I employ an epanechnikov kernel-based matching algorithm using euclidean distance with a 5 km half-bandwidth. I use a bias-adjustment procedure that controls for age, gender, education level of school and town characteristics (water, electricity and sanitation services).

Table 13
Treatment Effect on Staying at School for Teachers at Frontier Bonus Boundary

	2014			2015			2016		
	(1) Kinder	(2) Elem.	(3) H-S	(4) Kinder	(5) Elem.	(6) H-S	(7) Kinder	(8) Elem.	(9) H-S
Panel A. ***Permanent***									
Kernel($b = 5km$)	0.00	0.18	-0.00	-0.04*	-0.00	-0.00	-0.01	-0.01	0.04
	(0.01)	(0.09)	(0.01)	(0.02)	(0.01)	(0.01)	(0.03)	(0.01)	(0.04)
Kernel($b = 10km$)	0.00	0.16	-0.00	-0.03	0.01	0.02	-0.01	-0.01	0.02
	(0.01)	(0.09)	(0.01)	(0.02)	(0.03)	(0.01)	(0.02)	(0.01)	(0.03)
2 Nearest-Neighbor	-0.00	0.08***	0.00	-0.04	-0.00	-0.02	-0.01	-0.02	0.01
	(0.01)	(0.02)	(0.02)	(0.05)	(0.02)	(0.02)	(0.03)	(0.01)	(0.03)
4 Nearest-Neighbor	0.01	0.06***	0.00	-0.02	-0.04*	-0.00	0.01	-0.01	0.00
	(0.02)	(0.01)	(0.02)	(0.04)	(0.02)	(0.02)	(0.03)	(0.01)	(0.02)
Base	0.99	0.98	0.98	0.89	0.94	0.91	0.96	0.97	0.96
N_1	652	5401	4354	675	5350	4362	713	5338	4401
N_0	752	4029	1582	768	3956	1591	667	4102	1428
Panel B. ***Qualified Temporary***									
Kernel($b = 5km$)	0.17**	0.12	-0.09	0.12***	-0.01	-0.07	0.07	0.06	0.12**
	(0.06)	(0.08)	(0.09)	(0.04)	(0.06)	(0.05)	(0.06)	(0.05)	(0.04)
Kernel($b = 10km$)	0.13**	0.00	0.01	0.11***	0.02	-0.05	0.06	0.12**	0.09*
	(0.05)	(0.06)	(0.06)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
2 Nearest-Neighbor	0.06	0.10*	-0.03	0.11**	0.04	0.04	0.10*	0.13**	0.10**
	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.02)	(0.05)	(0.05)	(0.04)
4 Nearest-Neighbor	0.05	0.08	-0.02	0.08*	0.01	0.07***	0.09*	0.09*	0.11**
	(0.04)	(0.04)	(0.05)	(0.03)	(0.04)	(0.02)	(0.04)	(0.04)	(0.04)
Base	0.27	0.27	0.41	0.14	0.24	0.25	0.29	0.38	0.37
N_1	772	866	1294	861	1396	1979	1224	1820	2742
N_0	536	506	271	752	917	488	803	1138	659

Notes: Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. Standard errors are boostraped (200 repetitions) in kernel matching, robust in NN matching, clustered at the region level in the regressions. Each coefficient is a ATE estimator. The dependent variable is an indicator variable equal to 1 if teacher stay at the same school next year. Treatment unit: teacher working at Frontier school, Control unit: teacher works at Non-Frontier school. Both matching algorithms use a the Euclidean metric. Kernel matching uses a epanechnikov kernel function. All matching estimates use a bias correction procedure that controls for age, gender, education level of school and town characteristics (water, electricity and sanitation services). The 5 and 10 km^2 Block Fixed-Effects provide ATE estimates from a regression with fixed-effects.

Table 14
Treatment Effect on Filling Opening with Permanent Teacher at Isolated/Rural Boundary

	(1) 2015	(2) 2016	(3) 2017
Kernel ($b = 10km$)	0.00	0.01	0.01
	(0.02)	(0.01)	(0.03)
Kernel ($b = 20km$)	0.00	0.00	0.00
	(0.02)	(0.01)	(0.03)
2 Nearest-Neighbor	-0.02	-0.00	0.03*
	(0.02)	(0.01)	(0.02)
4 Nearest-Neighbor	-0.03	-0.00	0.03*
	(0.02)	(0.01)	(0.02)
6 Nearest-Neighbor	-0.04**	-0.01*	0.04**
	(0.02)	(0.01)	(0.01)
Base	0.19	0.065	0.15
N_1	784	3,156	625
N_0	728	2,907	696

Notes: Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. Standard errors are boostraped (200 repetitions) in kernel matching, robust in NN matching, clustered at the region level in the regressions. Each coefficient is an ATE estimator. The dependent variable is an indicator variable equal to 1 if job opening is filled by a permanent teacher. Treatment unit: job opening at Isolated school, Control unit: job opening at Rural school. Both matching algorithms use the Euclidean metric. Kernel matching uses a epanechnikov kernel function. All matching estimates use a bias correction procedure that controls for education level of school and town characteristics (water, electricity and sanitation services).

Table 15
Treatment Effect on Filling Opening with Permanent Teacher at VRAEM Bonus Boundary

	(1)	(2)	(3)
	2015	2016	2017
Kernel ($b = 10km$)	-0.07	-0.03	0.05
	(0.20)	(0.04)	(0.13)
Kernel ($b = 20km$)	-0.05	-0.03	0.01
	(0.10)	(0.03)	(0.06)
2 Nearest-Neighbor	-0.14	-0.01	0.05
	(0.12)	(0.02)	(0.05)
4 Nearest-Neighbor	-0.22*	-0.01	0.03
	(0.09)	(0.02)	(0.05)
6 Nearest-Neighbor	-0.12	-0.03	0.03
	(0.07)	(0.02)	(0.05)
Base	0.17	0.094	0.090
N_1	429	942	211
N_0	84	301	71

Notes: Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. Standard errors are boostraped (200 repetitions) in kernel matching, robust in NN matching, clustered at the region level in the regressions. Each coefficient is an ATE estimator. The dependent variable is an indicator variable equal to 1 if job opening is filled by a permanent teacher. Treatment unit: job opening at VRAEM school, Control unit: job opening at Non-VRAEM school. Both matching algorithms use the Euclidean metric. Kernel matching uses a epanechnikov kernel function. All matching estimates use a bias correction procedure that controls for education level of school and town characteristics (water, electricity and sanitation services).

Table 16
Treatment Effect on Filling Opening with Permanent Teacher at Frontier Bonus Boundary

	(1)	(2)	(3)
	2015	2016	2017
Kernel ($b = 10km$)	0.04	-0.12*	0.02
	(0.05)	(0.06)	(0.03)
Kernel ($b = 20km$)	-0.01	-0.07	0.02
	(0.05)	(0.04)	(0.01)
2 Nearest-Neighbor	-0.06	-0.15**	0.01
	(0.05)	(0.05)	(0.01)
4 Nearest-Neighbor	-0.03	-0.09*	0.01
	(0.03)	(0.04)	(0.01)
6 Nearest-Neighbor	-0.02	-0.07*	0.01
	(0.03)	(0.03)	(0.01)
Base	0.29	0.090	0.026
N_1	385	865	43
N_0	184	467	36

Notes: Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. Standard errors are boostraped (200 repetitions) in kernel matching, robust in NN matching, clustered at the region level in the regressions. Each coefficient is an ATE estimator. The dependent variable is an indicator variable equal to 1 if job opening is filled by a permanent teacher. Treatment unit: job opening at Frontier school, Control unit: job opening at Non-Frontier school. Both matching algorithms use the Euclidean metric. Kernel matching uses a epanechnikov kernel function. All matching estimates use a bias correction procedure that controls for education level of school and town characteristics (water, electricity and sanitation services).

Table 18
Treatment Effect on Teacher Retention at Isolated/Rural Boundary

	2013			2014			2015			2016		
	(1) Kinder	(2) Elem.	(3) H-S	(4) Kinder	(5) Elem.	(6) H-S	(7) Kinder	(8) Elem.	(9) H-S	(10) Kinder	(11) Elem.	(12) H-S
Panel A. ***Permanent***												
Kernel($b = 5km$)	0.01	-0.01	-0.01	0.01	-0.01	0.02	0.01	0.02	0.03	0.01	0.03	0.02
	(0.05)	(0.01)	(0.03)	(0.04)	(0.01)	(0.02)	(0.09)	(0.01)	(0.04)	(0.06)	(0.01)	(0.03)
Kernel($b = 10km$)	-0.00	-0.01	-0.01	0.02	-0.01	0.02	-0.02	0.01	0.02	0.05	0.01	0.02
	(0.03)	(0.01)	(0.02)	(0.03)	(0.01)	(0.01)	(0.05)	(0.01)	(0.03)	(0.05)	(0.01)	(0.02)
2 Nearest-Neighbor	-0.01	-0.02	-0.02	0.02	-0.01	0.02^{*}	-0.06	0.02	-0.01	0.02	0.02^{*}	-0.01
	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.03)	(0.01)	(0.02)	(0.03)	(0.01)	(0.01)
4 Nearest-Neighbor	-0.02	-0.01	-0.02	0.01	-0.00	0.02^{*}	-0.06	0.01	-0.02	0.01	0.02^{*}	-0.02
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
10km Block Fixed-Effects	-0.01	-0.01	0.01	0.02	-0.00	0.01	0.00	0.00	0.03	0.07	0.02	0.02
	(0.05)	(0.01)	(0.03)	(0.05)	(0.01)	(0.02)	(0.08)	(0.01)	(0.04)	(0.06)	(0.01)	(0.03)
Mean Retention - Control	0.84	0.91	0.87	0.94	0.90	0.93	0.71	0.84	0.69	0.89	0.91	0.90
N_1	674	5621	1064	618	5606	1088	608	5194	1082	515	4935	1103
N_0	900	5510	926	861	5677	964	842	5362	957	834	5547	971
Panel B. ***Qualified Temporary***												
Kernel($b = 5km$)	0.03	-0.01	-0.03	0.01	0.01	-0.01	0.09^{***}	0.03	0.04	0.07^{***}	0.03	-0.01
	(0.03)	(0.03)	(0.04)	(0.03)	(0.02)	(0.04)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)
Kernel($b = 10km$)	0.01	-0.02	-0.05	0.03	0.00	0.01	0.07^{***}	0.02	0.03	0.07^{***}	0.05^{**}	0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
2 Nearest-Neighbor	-0.01	-0.02	0.00	-0.00	0.00	0.00	0.08^{***}	0.02^{*}	0.03^{*}	0.09^{***}	0.06^{***}	0.04^{*}
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
4 Nearest-Neighbor	-0.00	-0.01	-0.01	0.00	-0.00	0.01	0.07^{***}	0.02^{*}	0.03^{*}	0.09^{***}	0.06^{***}	0.04^{**}
	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
10km Block Fixed-Effects	0.02	-0.00	-0.03	0.01	-0.01	-0.02	0.07^{**}	0.02	0.03	0.04	0.04^{*}	-0.01
	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.04)	(0.02)	(0.02)	(0.03)
Mean Retention - Control	0.13	0.083	0.16	0.21	0.13	0.27	0.12	0.078	0.17	0.21	0.19	0.26
N_1	1449	1999	890	2571	3159	1325	2628	3896	1406	3617	4750	1604
N_0	1340	1347	603	2371	2039	972	2463	2820	1065	3343	3331	1131
Panel C. ***Unqualified Temporary***												
Kernel($b = 5km$)	-0.04	0.03	0.00	0.05	-0.00	-0.17	0.04	0.02	-0.01	-0.07	0.04	0.01
	(0.17)	(0.10)	(.)	(0.13)	(0.14)	(0.19)	(0.13)	(0.11)	(0.07)	(0.10)	(0.07)	(0.05)
Kernel($b = 10km$)	-0.03	-0.11	0.25	-0.09	-0.05	0.03	-0.02	-0.01	-0.00	-0.05	0.00	-0.02
	(0.09)	(0.08)	(0.22)	(0.09)	(0.08)	(0.23)	(0.08)	(0.05)	(0.04)	(0.06)	(0.05)	(0.03)
2 Nearest-Neighbor	-0.04	-0.09	-0.06	-0.05	0.11^{**}	-0.06	0.03	-0.01	0.01	-0.03	0.06	0.01
	(0.07)	(0.06)	(0.08)	(0.06)	(0.04)	(0.09)	(0.04)	(0.03)	(0.02)	(0.04)	(0.03)	(0.02)
4 Nearest-Neighbor	0.02	-0.09	-0.04	-0.02	0.09^{*}	0.01	0.04	-0.00	0.01	-0.00	0.08^{**}	0.01
	(0.05)	(0.06)	(0.07)	(0.05)	(0.04)	(0.08)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)
10km Block Fixed-Effects	-0.04	-0.06	0.00	-0.05	-0.10	0.12	0.00	-0.09	0.02	-0.03	-0.04	0.02
	(0.20)	(0.14)	(.)	(0.10)	(0.21)	(0.14)	(0.13)	(0.10)	(0.11)	(0.06)	(0.08)	(0.06)
Mean Retention - Control	0.16	0.20	0.26	0.22	0.20	0.17	0.25	0.13	0.15	0.19	0.11	0.19
N_1	271	315	104	493	414	176	718	890	540	1108	1104	779
N_0	83	77	36	132	99	75	255	305	284	439	434	539

Notes: Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. Standard errors are boostraped (200 repetitions) in kernel matching, robust in NN matching, clustered at the region level in the regressions. Each coefficient is a ATE estimator. The dependent variable is an fraction of teachers sta variable equal to 1 if teacher stay at the same school next year. Treatment unit: teacher working at school, Control unit: teacher works at school. Both matching algorithms use a the Euclidean metric. Kernel matching uses a epanechnikov kernel function. All matching estimates use a bias correction procedure that controls for age, gender, education level of school and town characteristics (water, electricity and sanitation services). The 10 km^2 Block Fixed-Effects provide ATE estimates from a regression with fixed-effects.

Table 19
Treatment Effect on Teacher Retention at VRAEM Boundary

	2013			2014			2015			2016		
	(1) Kinder	(2) Elem.	(3) H-S	(4) Kinder	(5) Elem.	(6) H-S	(7) Kinder	(8) Elem.	(9) H-S	(10) Kinder	(11) Elem.	(12) H-S
Panel A. ***Permanent***												
Kernel($b = 5km$)	0.00	-0.02	-0.01	0.32	-0.03	0.04	-0.17	-0.01	0.03	-0.06	-0.01	-0.06
	(0.16)	(0.02)	(0.08)	(0.42)	(0.02)	(0.04)	(0.18)	(0.04)	(0.15)	(0.21)	(0.03)	(0.15)
Kernel($b = 10km$)	0.00	-0.05***	0.01	0.09	-0.01	0.04	-0.10	-0.01	-0.06	-0.02	-0.03	0.02
	(0.06)	(0.01)	(0.04)	(0.08)	(0.02)	(0.03)	(0.10)	(0.03)	(0.08)	(0.08)	(0.04)	(0.06)
2 Nearest-Neighbor	-0.08	-0.04*	0.01	-0.02	-0.01	-0.00	-0.11	-0.00	-0.12*	0.03	-0.03	-0.01
	(0.05)	(0.02)	(0.03)	(0.05)	(0.02)	(0.02)	(0.06)	(0.04)	(0.06)	(0.07)	(0.02)	(0.05)
4 Nearest-Neighbor	-0.07	-0.03*	0.03	-0.00	-0.01	0.01	-0.08	-0.00	-0.09*	0.02	-0.00	-0.01
	(0.05)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.07)	(0.04)	(0.04)	(0.06)	(0.03)	(0.04)
10km Block Fixed-Effects	0.00	0.01	0.03	0.07	-0.03	0.03	-0.21*	0.00	0.06	-0.08	-0.02	-0.06
	(0.13)	(0.02)	(0.05)	(0.09)	(0.02)	(0.03)	(0.10)	(0.04)	(0.07)	(0.12)	(0.02)	(0.12)
Mean Retention - Control	0.89	0.92	0.87	0.97	0.91	0.94	0.86	0.81	0.72	0.91	0.89	0.88
N_1	151	939	247	138	906	239	141	894	252	152	863	261
N_0	189	596	103	210	602	105	224	572	105	205	568	104
Panel C. ***Qualified Temporary***												
Kernel($b = 5km$)	0.07	-0.01	-0.05	0.02	0.04	0.15*	0.06	0.06	0.02	0.08	-0.02	0.10
	(0.04)	(0.04)	(0.09)	(0.06)	(0.07)	(0.06)	(0.03)	(0.05)	(0.09)	(0.05)	(0.10)	(0.09)
Kernel($b = 10km$)	0.06*	0.00	-0.02	0.01	0.05	0.08	0.04	0.01	0.02	0.08*	0.02	-0.00
	(0.03)	(0.04)	(0.07)	(0.04)	(0.05)	(0.05)	(0.03)	(0.04)	(0.03)	(0.04)	(0.06)	(0.06)
2 Nearest-Neighbor	0.10***	0.05	-0.07	0.00	0.08	0.09*	0.05*	0.06	0.03	0.09	-0.01	-0.00
	(0.03)	(0.05)	(0.04)	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)	(0.05)	(0.08)	(0.06)
4 Nearest-Neighbor	0.09**	0.04	-0.08*	-0.02	0.05	0.09*	0.03	0.02	0.02	0.05	-0.02	-0.03
	(0.03)	(0.03)	(0.04)	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.05)	(0.06)
10km Block Fixed-Effects	0.01	0.00	-0.10**	0.02	-0.00	0.02	0.03	0.05	0.03	0.07	0.03	-0.04
	(0.07)	(0.04)	(0.04)	(0.04)	(0.04)	(0.07)	(0.04)	(0.04)	(0.06)	(0.03)	(0.07)	(0.10)
Mean Retention - Control	0.090	0.084	0.16	0.17	0.15	0.20	0.099	0.10	0.16	0.20	0.21	0.26
N_1	485	498	251	674	572	308	701	703	325	818	744	368
N_0	290	174	67	455	316	123	467	362	141	455	444	145

Notes: Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. Standard errors are boostraped (200 repetitions) in kernel matching, robust in NN matching, clustered at the region level in the regressions. Each coefficient is a ATE estimator. The dependent variable is an fraction of teachers sta variable equal to 1 if teacher stay at the same school next year. Treatment unit: teacher working at school, Control unit: teacher works at school. Both matching algorithms use a the Euclidean metric. Kernel matching uses a epanechnikov kernel function. All matching estimates use a bias correction procedure that controls for age, gender, education level of school and town characteristics (water, electricity and sanitation services). The 10 km^2 Block Fixed-Effects provide ATE estimates from a regression with fixed-effects.

Table 20
Treatment Effect on Teacher Retention at Frontier Bonus Boundary

	2013			2014			2015			2016		
	(1) Kinder	(2) Elem.	(3) H-S	(4) Kinder	(5) Elem.	(6) H-S	(7) Kinder	(8) Elem.	(9) H-S	(10) Kinder	(11) Elem.	(12) H-S
Panel A. ***Permanent***												
Kernel($b = 5km$)	-0.01	0.01	-0.02	0.01	0.02	-0.01	-0.02	0.03	0.00	0.01	0.01	0.00
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.05)	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)
Kernel($b = 10km$)	-0.01	0.00	-0.02	0.03	0.03	0.02	-0.01	0.00	0.03	-0.02	0.01	0.03
	(0.02)	(0.01)	(0.02)	(0.03)	(0.02)	(0.02)	(0.04)	(0.03)	(0.03)	(0.03)	(0.02)	(0.04)
2 Nearest-Neighbor	0.02	0.01	0.01	0.02	-0.01	0.02	-0.01	0.01	0.05	0.01	0.01	0.03
	(0.04)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.06)	(0.03)	(0.03)	(0.04)	(0.02)	(0.04)
4 Nearest-Neighbor	-0.02	0.01	0.01	0.07*	0.00	0.01	0.05	0.00	0.02	0.03	0.01	0.02
	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.05)	(0.03)	(0.03)	(0.04)	(0.02)	(0.03)
10km Block Fixed-Effects	-0.01	0.02	0.00	0.02	0.02	-0.01	-0.04	0.01	-0.03	0.01	0.01	0.04
	(0.02)	(0.02)	(0.03)	(0.02)	(0.01)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.01)	(0.04)
Mean Retention - Control	0.95	0.91	0.94	0.91	0.89	0.93	0.80	0.81	0.82	0.94	0.91	0.93
N_1	272	1351	486	282	1408	497	298	1348	496	305	1298	494
N_0	300	636	151	303	642	156	312	620	158	267	565	134
Panel C. ***Qualified Temporary***												
Kernel($b = 5km$)	-0.04	-0.13	0.06	0.16***	0.06	0.02	0.08**	0.04	0.02	0.06	0.12**	0.09
	(0.07)	(0.13)	(0.16)	(0.04)	(0.05)	(0.07)	(0.03)	(0.04)	(0.07)	(0.05)	(0.04)	(0.07)
Kernel($b = 10km$)	-0.05	-0.05	0.04	0.10*	0.03	0.02	0.09***	0.05*	-0.00	0.06	0.12***	0.08
	(0.05)	(0.06)	(0.08)	(0.04)	(0.03)	(0.05)	(0.02)	(0.02)	(0.05)	(0.04)	(0.03)	(0.04)
2 Nearest-Neighbor	-0.05	-0.02	0.06	0.09**	0.09***	0.05	0.10***	0.07*	0.06	0.03	0.10*	0.04
	(0.04)	(0.04)	(0.05)	(0.03)	(0.02)	(0.06)	(0.02)	(0.03)	(0.04)	(0.04)	(0.04)	(0.03)
4 Nearest-Neighbor	-0.04	0.04	0.05	0.08**	0.08**	0.03	0.09***	0.05	0.05	0.03	0.08*	0.05
	(0.04)	(0.03)	(0.04)	(0.03)	(0.02)	(0.05)	(0.02)	(0.02)	(0.03)	(0.03)	(0.04)	(0.03)
10km Block Fixed-Effects	-0.06	-0.07	0.06	0.11**	0.02	-0.05	0.06	0.01	-0.00	0.04	0.07	0.07
	(0.06)	(0.11)	(0.07)	(0.04)	(0.03)	(0.10)	(0.04)	(0.03)	(0.07)	(0.04)	(0.05)	(0.10)
Mean Retention - Control	0.16	0.082	0.27	0.20	0.12	0.26	0.12	0.11	0.18	0.27	0.24	0.30
N_1	466	475	321	717	675	452	676	927	522	891	1029	564
N_0	241	180	83	426	301	135	533	371	160	501	433	149

Notes: Standard errors in parentheses. * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Standard errors are boostraped (200 repetitions) in kernel matching, robust in NN matching, clustered at the region level in the regressions. Each coefficient is a ATE estimator. The dependent variable is an fraction of teachers sta variable equal to 1 if teacher stay at the same school next year. Treatment unit: teacher working at school, Control unit: teacher works at school. Both matching algorithms use a the Euclidean metric. Kernel matching uses a epanechnikov kernel function. All matching estimates use a bias correction procedure that controls for age, gender, education level of school and town characteristics (water, electricity and sanitation services). The 10 km^2 Block Fixed-Effects provide ATE estimates from a regression with fixed-effects.

Table 21
Treatment Effect on Teacher Composition at Isolated/Rural Boundary

	2013			2014			2015			2016		
	(1) Kinder	(2) Elem.	(3) H-S	(4) Kinder	(5) Elem.	(6) H-S	(7) Kinder	(8) Elem.	(9) H-S	(10) Kinder	(11) Elem.	(12) H-S
Panel A. ***Frac. of Permanent***												
Kernel($b = 5km$)	-0.04	-0.03**	-0.02	-0.04*	-0.05***	-0.08*	-0.06***	-0.06***	-0.04	-0.06***	-0.05**	-0.05*
	(0.03)	(0.01)	(0.03)	(0.02)	(0.01)	(0.03)	(0.02)	(0.01)	(0.03)	(0.02)	(0.02)	(0.02)
Kernel($b = 10km$)	-0.07***	-0.02**	-0.06***	-0.05***	-0.06***	-0.09***	-0.06***	-0.07***	-0.06***	-0.06***	-0.07***	-0.06***
	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
2 Nearest-Neighbor	-0.06***	-0.03***	-0.09***	-0.06***	-0.06***	-0.11***	-0.07***	-0.08***	-0.08***	-0.07***	-0.09***	-0.08***
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
4 Nearest-Neighbor	-0.06***	-0.04***	-0.08***	-0.07***	-0.07***	-0.11***	-0.08***	-0.08***	-0.08***	-0.08***	-0.10***	-0.08***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
10km Block Fixed-Effects	-0.07*	-0.03**	-0.06*	-0.05**	-0.05***	-0.09**	-0.05*	-0.06***	-0.05	-0.05***	-0.05***	-0.06**
	(0.03)	(0.01)	(0.03)	(0.02)	(0.01)	(0.03)	(0.02)	(0.01)	(0.03)	(0.01)	(0.01)	(0.02)
Mean Fraction - Control	0.38	0.88	0.72	0.23	0.83	0.53	0.21	0.73	0.45	0.16	0.68	0.37
N_1	2265	6454	1248	3510	6999	1409	3661	7282	1474	4802	7927	1668
N_0	2120	5715	989	3049	6020	1103	3106	5974	1120	4080	6448	1173
Panel B. ***Frac. of Qualified Temp.***												
Kernel($b = 5km$)	0.01	0.03**	0.02	0.04*	0.04***	0.08**	0.05*	0.05***	0.05	0.04*	0.04**	0.08**
	(0.03)	(0.01)	(0.03)	(0.02)	(0.01)	(0.03)	(0.02)	(0.01)	(0.03)	(0.02)	(0.02)	(0.03)
Kernel($b = 10km$)	0.06**	0.02*	0.06**	0.06***	0.06***	0.09***	0.05**	0.06***	0.06***	0.04*	0.07***	0.09***
	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)
2 Nearest-Neighbor	0.06***	0.02**	0.08***	0.05**	0.04***	0.10***	0.04**	0.04***	0.06***	0.03*	0.07***	0.10***
	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
4 Nearest-Neighbor	0.06***	0.03***	0.07***	0.05***	0.05***	0.10***	0.05***	0.04***	0.06***	0.03**	0.07***	0.09***
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
10km Block Fixed-Effects	0.04	0.03**	0.05	0.04	0.04***	0.09**	0.03	0.06***	0.04	0.04*	0.05***	0.08***
	(0.03)	(0.01)	(0.03)	(0.02)	(0.01)	(0.03)	(0.02)	(0.01)	(0.03)	(0.02)	(0.01)	(0.02)
Mean Fraction - Control	0.59	0.11	0.28	0.73	0.16	0.45	0.73	0.25	0.50	0.77	0.29	0.52
N_1	2265	6454	1248	3510	6999	1409	3661	7282	1474	4802	7927	1668
N_0	2120	5715	989	3049	6020	1103	3106	5974	1120	4080	6448	1173

Notes: Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. Standard errors are boostraped (200 repetitions) in kernel matching, robust in NN matching, clustered at the region level in the regressions. Each coefficient is a ATE estimator. The dependent variable is an fraction of teachers sta variable equal to 1 if teacher stay at the same school next year. Treatment unit: teacher working at school, Control unit: teacher works at school. Both matching algorithms use a the Euclidean metric. Kernel matching uses a epanechnikov kernel function. All matching estimates use a bias correction procedure that controls for age, gender, education level of school and town characteristics (water, electricity and sanitation services). The 10 km^2 Block Fixed-Effects provide ATE estimates from a regression with fixed-effects.

Table 22
Treatment Effect on Teacher Composition at VRAEM Boundary

	2013			2014			2015			2016		
	(1) Kinder	(2) Elem.	(3) H-S	(4) Kinder	(5) Elem.	(6) H-S	(7) Kinder	(8) Elem.	(9) H-S	(10) Kinder	(11) Elem.	(12) H-S
Panel A. ***Frac. of Permanent***												
Kernel($b = 5km$)	0.00	-0.01	-0.19	0.04	0.04	0.03	-0.13*	-0.04	0.13	0.06	0.02	-0.00
	(0.06)	(0.03)	(0.12)	(0.05)	(0.04)	(0.08)	(0.05)	(0.03)	(0.14)	(0.04)	(0.05)	(0.07)
Kernel($b = 10km$)	-0.00	-0.01	-0.13*	-0.01	0.06	0.06	-0.11**	0.01	0.11**	0.05	0.03	0.01
	(0.04)	(0.02)	(0.06)	(0.04)	(0.03)	(0.05)	(0.04)	(0.03)	(0.04)	(0.02)	(0.04)	(0.05)
2 Nearest-Neighbor	-0.02	-0.04	-0.21***	-0.00	0.07*	0.01	-0.12***	0.03	0.09**	0.03	0.04	-0.02
	(0.06)	(0.03)	(0.05)	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)	(0.05)	(0.04)
4 Nearest-Neighbor	-0.03	-0.05*	-0.20***	0.00	0.05	-0.01	-0.10**	0.01	0.08**	0.02	0.03	-0.03
	(0.05)	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.02)	(0.04)	(0.03)
10km Block Fixed-Effects	-0.02	-0.02	-0.17	0.03	0.02	-0.04	-0.01	-0.00	0.06	0.03	0.07	0.05
	(0.05)	(0.02)	(0.13)	(0.05)	(0.03)	(0.09)	(0.05)	(0.02)	(0.07)	(0.04)	(0.04)	(0.05)
Mean Fraction - Control	0.37	0.86	0.66	0.28	0.74	0.44	0.26	0.67	0.35	0.24	0.58	0.29
N_1	648	1085	304	822	1057	314	880	1131	339	1005	1161	376
N_0	434	633	112	603	691	130	641	697	152	611	741	152
Panel B. ***Frac. of Qualified Temp.***												
Kernel($b = 5km$)	0.00	0.01	0.17	-0.04	-0.05	-0.06	0.17**	0.03	-0.03	0.04	0.01	0.14
	(0.06)	(0.03)	(0.12)	(0.05)	(0.04)	(0.07)	(0.06)	(0.04)	(0.20)	(0.06)	(0.05)	(0.08)
Kernel($b = 10km$)	0.00	0.01	0.12	0.00	-0.06	-0.08	0.19***	-0.00	-0.01	0.05	0.03	0.06
	(0.05)	(0.03)	(0.07)	(0.04)	(0.03)	(0.05)	(0.04)	(0.03)	(0.05)	(0.04)	(0.04)	(0.05)
2 Nearest-Neighbor	0.04	0.04	0.19***	-0.01	-0.07	-0.04	0.18***	-0.03	0.01	0.06	0.05	0.10*
	(0.06)	(0.03)	(0.05)	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)	(0.03)	(0.05)	(0.06)	(0.05)
4 Nearest-Neighbor	0.04	0.06*	0.17***	-0.03	-0.05	-0.02	0.17***	-0.03	-0.00	0.05	0.05	0.10*
	(0.05)	(0.02)	(0.04)	(0.03)	(0.03)	(0.04)	(0.05)	(0.04)	(0.03)	(0.04)	(0.05)	(0.04)
10km Block Fixed-Effects	0.02	0.02	0.14	-0.03	-0.02	0.01	0.05	-0.00	-0.00	0.03	0.00	-0.00
	(0.04)	(0.02)	(0.11)	(0.06)	(0.03)	(0.09)	(0.07)	(0.03)	(0.07)	(0.06)	(0.06)	(0.06)
Mean Fraction - Control	0.61	0.14	0.33	0.68	0.25	0.55	0.64	0.30	0.49	0.65	0.36	0.50
N_1	648	1085	304	822	1057	314	880	1131	339	1005	1161	376
N_0	434	633	112	603	691	130	641	697	152	611	741	152

Notes: Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. Standard errors are boostraped (200 repetitions) in kernel matching, robust in NN matching, clustered at the region level in the regressions. Each coefficient is a ATE estimator. The dependent variable is an fraction of teachers sta variable equal to 1 if teacher stay at the same school next year. Treatment unit: teacher working at school, Control unit: teacher works at school. Both matching algorithms use a the Euclidean metric. Kernel matching uses a epanechnikov kernel function. All matching estimates use a bias correction procedure that controls for age, gender, education level of school and town characteristics (water, electricity and sanitation services). The 10 km^2 Block Fixed-Effects provide ATE estimates from a regression with fixed-effects.

Table 23

Treatment Effect on Teacher Composition at Frontier Bonus Boundary

	2013			2014			2015			2016		
	(1) Kinder	(2) Elem.	(3) H-S	(4) Kinder	(5) Elem.	(6) H-S	(7) Kinder	(8) Elem.	(9) H-S	(10) Kinder	(11) Elem.	(12) H-S
Panel A. ***Frac. of Permanent***												
Kernel($b = 5km$)	-0.10**	0.05*	-0.00	-0.05	0.03	0.03	-0.06	0.05	0.03	-0.07	0.01	0.06
	(0.04)	(0.02)	(0.03)	(0.04)	(0.02)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)
Kernel($b = 10km$)	-0.07*	0.04	-0.02	-0.04	0.02	-0.00	-0.09*	0.02	-0.01	-0.06	-0.00	-0.01
	(0.03)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.04)	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)
2 Nearest-Neighbor	-0.05	0.02	-0.06	-0.02	-0.01	-0.05	-0.06*	-0.02	-0.00	-0.06*	0.01	0.02
	(0.03)	(0.03)	(0.04)	(0.03)	(0.02)	(0.06)	(0.03)	(0.03)	(0.04)	(0.02)	(0.03)	(0.04)
4 Nearest-Neighbor	-0.06*	0.01	-0.07*	-0.03	-0.01	-0.02	-0.06*	-0.02	0.02	-0.05*	0.00	0.01
	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)	(0.05)	(0.02)	(0.02)	(0.04)	(0.02)	(0.02)	(0.04)
10km Block Fixed-Effects	-0.05	0.02	0.00	-0.01	0.03	0.03	0.01	0.06*	0.06	-0.00	0.03	0.06
	(0.04)	(0.02)	(0.03)	(0.04)	(0.02)	(0.04)	(0.04)	(0.03)	(0.06)	(0.04)	(0.02)	(0.05)
Mean Fraction - Control	0.56	0.82	0.83	0.40	0.76	0.74	0.31	0.66	0.62	0.28	0.57	0.52
N_1	787	1571	529	1061	1688	545	1062	1790	560	1332	1878	600
N_0	486	709	158	620	741	168	757	769	182	663	771	158
Panel B. ***Frac. of Qualified Temp.***												
Kernel($b = 5km$)	0.08	-0.04	0.00	0.05	-0.02	-0.04	0.05	-0.03	-0.00	0.11**	0.01	-0.02
	(0.04)	(0.03)	(0.03)	(0.04)	(0.02)	(0.04)	(0.05)	(0.03)	(0.03)	(0.04)	(0.03)	(0.05)
Kernel($b = 10km$)	0.06	-0.04	0.02	0.04	-0.01	-0.00	0.08	-0.00	0.02	0.10**	0.02	0.03
	(0.03)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.04)	(0.02)	(0.03)	(0.04)	(0.03)	(0.03)
2 Nearest-Neighbor	0.04	-0.02	0.04	-0.01	0.01	0.01	0.02	0.01	-0.00	0.09**	0.01	-0.07
	(0.03)	(0.02)	(0.04)	(0.04)	(0.02)	(0.06)	(0.03)	(0.03)	(0.05)	(0.03)	(0.03)	(0.04)
4 Nearest-Neighbor	0.03	-0.02	0.04	-0.01	0.02	-0.02	0.03	0.02	-0.02	0.06*	0.02	-0.06
	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)	(0.05)	(0.03)	(0.02)	(0.04)	(0.03)	(0.02)	(0.03)
10km Block Fixed-Effects	0.02	-0.01	-0.01	0.02	-0.02	-0.03	-0.01	-0.02	-0.03	0.04	0.00	-0.03
	(0.04)	(0.01)	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.05)	(0.02)	(0.05)
Mean Fraction - Control	0.42	0.14	0.16	0.55	0.19	0.25	0.59	0.25	0.31	0.61	0.32	0.38
N_1	787	1571	529	1061	1688	545	1062	1790	560	1332	1878	600
N_0	486	709	158	620	741	168	757	769	182	663	771	158

Notes: Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001. Standard errors are boostraped (200 repetitions) in kernel matching, robust in NN matching, clustered at the region level in the regressions. Each coefficient is a ATE estimator. The dependent variable is an fraction of teachers sta variable equal to 1 if teacher stay at the same school next year. Treatment unit: teacher working at school, Control unit: teacher works at school. Both matching algorithms use a the Euclidean metric. Kernel matching uses a epanechnikov kernel function. All matching estimates use a bias correction procedure that controls for age, gender, education level of school and town characteristics (water, electricity and sanitation services). The 10 km^2 Block Fixed-Effects provide ATE estimates from a regression with fixed-effects.

Table 24
Monthly Base Salary Schedule and Teacher Distribution as of 2017

(1) Step	(2) Monthly Base Salary (S/)	(3) % of First Step	(4) Distribution of Permanent Teachers	(5) Distribution of Temporary Teachers
First	1,781	100%	34%	100%
Second	1,959	110%	29%	0%
Third	2,137	120%	20%	0%
Fourth	2,315	130%	11%	0%
Fifth	2,671	150%	4%	0%
Sixth	3,116	175%	1%	0%
Seventh	3,383	190%	0%	0%
Eight	3,739	210%	0%	0%

Source: Ministry of Education, administrative teacher data combined with School Census (Censo Escolar).

Table 25
Regression of Share of Qualified Temporary Teachers at the School Level

	(1) All	(2) 2013	(3) 2014	(4) 2015	(5) 2016	(6) 2017	(7) All
Rural	0.021**	0.036***	0.034***	0.032***	0.015*	0.00076	
	0.0092	0.012	0.011	0.010	0.0086	0.0082	
Suburban	0.027**	0.041***	0.039***	0.038***	0.021**	0.0090	
	0.011	0.014	0.012	0.011	0.010	0.010	
Urban	0.044***	0.053***	0.054***	0.053***	0.037**	0.033**	
	0.016	0.018	0.017	0.017	0.015	0.015	
Constant	0.91***	0.92***	0.92***	0.91***	0.91***	0.90***	0.93***
	0.0081	0.0098	0.0089	0.0088	0.0076	0.0076	0.00033
Obs.	171455	19255	29876	37103	42355	42866	171455
R^2	0.40	0.47	0.42	0.38	0.39	0.39	0.74
School Dist. $\times$ Year FE	yes	yes	yes	yes	yes	yes	yes
School FE							yes

Source: Ministry of Education, administrative teacher dataset combined with School Census (Censo Escolar) data.
Notes: The dependent variable is the share of qualified temporary teachers. Each observation is a year-school pair in schools with temporary teachers. Base category : Isolated schools.
Clustered standard errors at the school district in parenthesis.

Table 26
Share of Temporary Teacher by Year and Geographical Classification

	Isolated		Rural		Suburban		Urban		VRAEM		Frontier		Nu
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	share	N	share	N	share	N	share	N	share	N	share	N	Temporary
2013	0.35	29,696	0.23	39,777	0.14	32,097	0.10	156,933	0.40	7,708	0.20	16,522	40,383
2014	0.48	36,388	0.32	45,646	0.21	35,558	0.16	169,565	0.46	8,693	0.27	18,462	65,898
2015	0.55	41,030	0.42	50,381	0.31	38,248	0.24	174,688	0.54	9,497	0.37	19,741	98,372
2016	0.64	45,846	0.49	55,445	0.39	42,535	0.29	186,408	0.59	10,364	0.44	21,262	127,611
2017	0.67	48,317	0.53	57,874	0.44	44,625	0.32	191,931	0.61	10,846	0.46	21,824	144,364

Source: Ministry of Education, administrative teacher dataset combined with School Census (Censo Escolar) data.
Notes: The "share" column represents the share of temporary teacher with respect to the total number of teachers, it is not be con "N" column represents the total number of teachers (permanent and temporary). Urban/Non-urban classification is exhausti VRAEM and Frontier classifications can overlap with Urban/Non-urban classification.

Table 27
Validity Test for Observable Variables Pre-Reform (2013)

	(1) Student to Teacher Ratio	(2) Reading Score	(3) Math Score	(4) Town with Electric service	(5) Town with Water service	Tow
Isolated vs Rural	-0.382	-2.634	-0.729	-0.0502***	-0.0564***	
s.e.	(0.29)	(3.28)	(4.37)	(0.01)	(0.01)	
N	16,485.0	6,069.0	6,070.0	15,675.0	15,675.0	
VRAEM	1.13	14.69	19.14	(0.02)	(0.05)	
s.e.	(0.944)	(11.2)	(12.99)	(0.028)	(0.0435)	
N	2,693	844	844	2,638	2,638	
Frontier	-0.317	5.742	14.27	0.0299	0.129***	
s.e.	(0.57)	(7.16)	(8.20)	(0.02)	(0.02)	
N	3,956	1,166	1,166	3,672	3,672	

Notes: Standard errors in parentheses * p<0.05, ** p<0.01, *** p<0.001. Each coefficient is an ATET estimator. I employ an epanechnikov kernel-based matching distance with a 5 km half-bandwidth.