Teacher Compensation and Structural Inequality: Evidence from Centralized Teacher School Choice in Perú*

Matteo Bobba† Tim Ederer‡ Gianmarco León-Ciliotta§
Christopher A. Neilson¶ Marco Nieddu‖

March 31, 2024

Abstract

We exploit data on the universe of public-school teachers and students in Perú to establish that wage rigidity makes teachers choose schools based on non-pecuniary factors, magnifying the existing urban-rural gap in student achievement. Leveraging a reform in the teacher compensation structure, we provide causal evidence that increasing salaries in less desirable locations is effective at improving student learning by attracting higher-quality teachers. We then build and estimate a model of teacher sorting across schools and student achievement production, whereby teachers are heterogeneous in their preferences over non-wage attributes and their comparative advantages in teaching different student types. Counterfactual compensation policies that leverage information about teachers' preferences and value-added can result in a substantially more efficient and equitable allocation by inducing teachers to sort based on their comparative advantage.

*We thank the Ministry of Education of Perú, and in particular Camila Alva, for supporting rigorous policy evaluation efforts and continued collaboration. Mohammad Akbarpour, Ernesto Dal Bo, Michael Dinerstein, Pascaline Dupas, Frederico Finan, Yinghua He, Imran Rasul, Jesse Silbert, Alessandro Tarozzi, and Takuro Yamashita generously provided us with insightful feedback on earlier versions of this draft. We also thank Mariel Bedoya, Manuela Cardona Badillo, Maria Elena Guerrero, Isabel Jacas, Cecilia Moreira, and Eduardo Vargas for outstanding research assistance. Matteo Bobba acknowledges financial support from the AFD, the H2020-MSCA-RISE project GEMCLIME-2020 GA No 681228, and the ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program). Gianmarco León-Ciliotta recognizes financial support from the Spanish Ministry of Economy and Competitiveness (through the Severo Ochoa Programme for Centres of Excellence in R&D (CEX2019-000915-S), grant AEI/10.13039/501100011033 and RyC-2017-23172 financed by MICIU/AEI /10.13039/501100011033 and FSE invierte en tu futuro), and the BSE Seed Grant.

†Toulouse School of Economics & CEPR & IZA, matteo.bobba@tse-fr.eu
‡University of Chicago, tederer@uchicago.edu
§Universitat Pompeu Fabra & BSE & IPEG & BREAD & CEPR, gianmarco.leon@upf.edu
¶Yale University & NBER, christopher.neilson@yale.edu
‖University of Cagliari & CRENoS, mgnieddu@unica.it

1 Introduction

Children born in remote or rural communities achieve significantly lower levels of academic achievement than their urban counterparts (World Bank, 2018). These disparities are particularly salient in developing countries, where distressed and remote areas suffer from a structural and historically persistent underdevelopment (Sokoloff and Engerman, 2000; Banerjee and Iyer, 2005; Dell, 2010). Designing effective policies that provide equal opportunities despite these barriers is thus a first-order concern. In this paper, we argue that structural inequalities in schooling outcomes can be mitigated or reinforced by the inherent mobility of a critical factor in the production of human capital: teachers.

Teachers largely contribute to student learning and long-run outcomes (Rivkin et al., 2005; Chetty et al., 2014b; Araujo et al., 2016; Jackson, 2018). Recent evidence further documents that teachers hold comparative advantages in teaching different types of students (Gershenson et al., 2022; Ahn et al., 2023; Graham et al., 2023), implying that the allocation of teachers is not only relevant for equity, but may also have crucial implications for efficiency. Yet, little is known about the fundamental drivers of teacher sorting across schools. Wage rigidity – a common feature of most public schooling systems around the world – has been highlighted as an important friction that would make teachers sort on non-pecuniary aspects of employment (Rosen, 1986), potentially resulting in low-quality teachers disproportionately working in disadvantaged schools (Clotfelter et al., 2005; Mansfield, 2015). However, allowing school districts to flexibly set wages has been shown to yield limited equity and efficiency gains (Biasi et al., 2021), suggesting that direct public intervention might be more effective.

In this paper, we conduct a comprehensive analysis of how the design of teacher compensation schemes jointly shapes teacher sorting across locations and the spatial distribution of student achievement. Our study takes place in Perú – a large developing country characterized by a diverse geography and population, as well as profound spatial inequalities in schooling outcomes. Using a unique dataset on the universe of public-school teachers and their students, we show that wage rigidity prompts teachers to sort on non-pecuniary aspects of employment, leading high-quality teachers to concentrate in urban areas. We leverage a reform in teacher compensation structure to show that raising wages in disadvantaged schools is effective at attracting high-quality teachers and improving student learning. We then build and estimate a model of teacher sorting across schools and student achievement in which teachers are heterogeneous in their preferences over jobs and in both their absolute and comparative advantages in teaching different student types. Finally, we provide a framework for teacher compensation design and find substantial equity and efficiency gains from alternative compensation policies that would induce teachers to sort based on their comparative advantages.

Our analysis draws on very rich administrative panel data linking the universe of applicants and jobs posted within two consecutive rounds of the nationwide centralized recruitment drive for public teachers in Perú with an array of additional data sources on public primary schools and students. All applicants take a standardized national competency test and choose their preferred position sequentially according to their test score rank. We first document that students in rural areas lack access to basic amenities and attend schools with notably inferior infrastructure. Teachers working in rural schools are significantly less qualified, as evidenced by their average lower performance (0.43σ) in the national teacher competency test. Survey data eliciting teachers' preferences over various non-wage job amenities suggest that wage rigidity induces teachers to prefer urban areas that offer better amenities. As a result, highly competent teachers disproportionately concentrate in urban areas. These inequities are associated with a staggering difference of 0.58 of a standard deviation (σ) in students' scores in standardized tests between urban and rural areas.

Against this backdrop, the Ministry of Education of Peru (MoE) introduced a policy that attributes wage bonuses to teachers working in rural schools. These bonuses get increasingly large with remoteness through discrete jumps determined by arbitrary cutoffs on the school locality's population and its distance to the provincial capital. Leveraging a 13% wage increase embedded in this bonus scheme in a regression discontinuity design, we find that teachers' labor supply largely increased in schools offering higher wages. This translated into a 0.39σ increase in newly recruited teachers' scores at the national competency test as well as a 0.32σ and 0.23σ increase in students' test scores in both math and language, respectively. We provide evidence that these gains did not come at the expense of lowering the quality of teachers in lower-paying schools located near the threshold. Even though incumbent teachers were also eligible to receive the wage bonus, the effects are entirely driven by newly recruited teachers. The policy was thus effective at increasing student achievement in targeted areas by attracting higher quality teachers, but not through increased productivity of incumbent teachers.

Despite our positive assessment of the efficacy of the wage reform at improving learning outcomes in rural schools, it remains unclear the extent to which this policy instrument can be leveraged to reduce geographical disparities in student achievement. We thus build an empirical model of teacher sorting across schools and student achievement to quantify and decompose the aggregate effects of the policy, as well as to characterize the potential equity and efficiency gains from alternative compensation policies. We allow teachers to have heterogeneous preferences over wages and non-wage job attributes. Specifically, factors such as local amenities, geographical proximity, ethnolinguistic alignment, and teaching conditions induce both vertical and horizontal differentiation across jobs. In line with the institutional framework, wages are fixed, and we assume that the equilibrium teacher-school match is stable with respect to teachers' preferences over schools and schools' priorities. Teacher sorting maps into the distribution of student achievement through a potential outcomes framework where the value-added of a given teacher is allowed to be heterogeneous and vary with students' prior achievement measures and demographics. Teachers' absolute and comparative advantages flexibly correlate with latent teacher attributes governing their willingness to pay for non-wage amenities as well as their valuation of the outside option. This potentially captures intrinsic motivation that cannot be explained by observable teacher characteristics and allows for selection on unobserved teaching quality in response to counterfactual teacher compensation policies.

Stability implies that each teacher is matched to their preferred school among their feasible choice set, i.e. the set of schools that would be willing to rematch with them. As school priorities are observed, we can construct the set of feasible schools of each teacher directly from the data. This unique feature of the data at our disposal allows us to express the equilibrium teacher-school allocation as the outcome of a discrete choice problem. We leverage this insight to characterize the mapping between teachers' preferences and the observed teacher-school match, as well as between teacher effectiveness and the realized distribution of student achievement. We show that these relationships are invertible, such that teachers' preferences and value-added coefficients are identified.

The estimated model replicates the main features of the data, namely teacher sorting across locations, threshold-crossing effects on teacher quality and student achievement induced by the wage bonus policy, as well as moments of the distributions of matched teacher and school characteristics. We find that the average willingness to pay for non-wage amenities is substantial and greatly varies across teachers. Consistent with the descriptive survey evidence, teachers have a high willingness to pay for proximity to home and better teaching conditions, implying that they strongly prefer urban areas. Very remote schools would need to offer two to four times the wage increase implied by the current bonus structure to compensate teachers for differences in non-pecuniary benefits with respect to urban areas.

We report substantial heterogeneity in teachers' absolute and comparative advantage. Teachers in the top 5% of the value-added distribution generate an average test score gain of 0.93σ in math and 0.82σ in Spanish, compared to the median teacher. Importantly, 11-18% of the overall variance in teacher effectiveness can be explained by differences in their comparative advantage. The variance in teacher value-added is larger for students already lagging behind, implying that rural schools would highly benefit from making teachers sort based on their comparative advantage. Observed teacher characteristics explain little of the overall variance in teachers' absolute and comparative advantage. Instead, teachers' latent types driving their preferences over job postings are highly predictive of their effectiveness. This shows that combining data on teachers' school choices with data on student achievement can significantly help us get a better understanding of what makes a good teacher.

We use the estimated parameters to simulate the equilibrium teacher-school allocation and the country-wide distribution of student achievement that would have occurred in the absence of the wage bonus reform. Despite the large local effects estimated at the population threshold, the policy only decreased the overall urban-rural gap in student test scores by 0.08σ. Teacher sorting across locations remains unequal and favors urban areas, as we observe a residual 0.08-0.11σ urban-rural gap in teacher value added. We also find that the existing wage reform fails to incentivize teachers to sort based on their comparative advantage, implying large potential efficiency and equity gains from alternative teacher-school allocations.

We quantify these potential efficiency and equity gains by characterizing a set of counterfactual allocations that maximize total student achievement given relative weights put on rural areas against urban areas. We show that it would be possible to close the urban-rural gap in teacher value-added at no cost for urban areas by making teachers sort based on their comparative advantage. We also highlight the benefits of using Bayesian shrinkage in such assignment problems with treatment effect heterogeneity. By attenuating the estimates of teachers' comparative advantage proportionally to how imprecise they are, this approach limits the potential for efficiency losses by prioritizing teacher-student matches that are high-value with high certainty. We provide evidence from Monte Carlo simulations supporting this claim.

Finally, we provide a framework to design teacher compensation policies that internalize information on teachers' preferences and effectiveness. We consider the problem of a policy-maker choosing how to set wages in each school with the objective of inducing teachers to sort more equitably and efficiently at a minimal cost. Specifically, we set the objective as ensuring that teacher value-added in each rural school uniformly does not fall below a given threshold. We then consider a counterfactual economy where schools would be allowed to bid for higher quality teachers by increasing wages until they succeed at meeting this objective. We leverage results from Hatfield and Milgrom (2005) to show that the solution to the policymaker's problem is equivalent to the outcome of the school-proposing generalized Deferred Acceptance algorithm within this counterfactual economy. Using this procedure, we find that it would be feasible to fully close the urban-rural gap in teacher value added at a slightly higher cost than the status quo policy. By inducing teachers to sort on their comparative advantages, the increase in teacher value-added in rural areas is achieved at no cost for urban areas. Accounting for teachers' preferences over non-wage attributes as well as teachers' absolute *and* comparative advantage is thus crucial to design cost-effective compensation policies aimed at alleviating structural inequalities.

Our findings speak to several strands of literature. A large body of work studies how teachers contribute to student achievement (Rivkin et al., 2005; Chetty et al., 2014a; Bau and Das, 2020; Ahn et al., 2023). We first highlight the benefits of informing teacher value-added models with data on teachers' choices over schools. We find that such data allow to identify teachers' latent types by revealing their preferences over job attributes, which in turn predict a large share of the variance of teacher value-added. In contrast, observable teacher characteristics have been shown to be poor predictors of teacher effectiveness (Chetty et al., 2014a; Bau and Das, 2020). This approach has the additional advantage of capturing selection on unobserved quality that would occur as a response to teacher compensation policies (Rothstein, 2015; Brown and Andrabi, 2020). We further show that allowing for teacher value-added to be heterogeneous and vary with students' types is crucial to assess the implications of teacher sorting. Specifically, we find large equity and efficiency gains from alternative teacher assignments due to the presence of match effects in the student achievement production function. Finally, we provide a *countrywide* assessment of the impact of teacher sorting on the distribution of student achievement. We find that the within-country variance in teacher value-added is almost twice as large as similar within-district estimates from developed and developing countries (Chetty et al., 2014a; Bau and Das, 2020). This suggests that teachers contribute significantly more to the overall within-country variance in student achievement than previously expected.

This paper also contributes to a recent literature studying the link between teacher sorting and student achievement through equilibrium models of the labor market for teachers (Boyd et al., 2013; Bonhomme et al., 2016; Tincani, 2021; Biasi et al., 2021; Bates et al., 2022). In contrast to the literature, our approach utilizes panel data on a nation-wide centralized allocation mechanism for public-sector teachers, providing detailed information on each applicant's choices and choice sets over several years. This exceptional dataset, combined with quasi-experimental variation in wages, allows us to identify and estimate, under minimal assumptions, a rich empirical model of teacher sorting and student achievement accounting for widely heterogeneous preferences over non-wage attributes and teachers' comparative advantages.

Finally, our work is broadly related to a growing literature studying personnel and organizational policies in the public sector (Finan et al., 2017; Khan et al., 2019). While there is large body of work studying the effectiveness of pay-for-performance schemes for teachers (Muralidharan and Sundararaman, 2011; Barrera-Osorio and Raju, 2017; Biasi, 2021; Leaver et al., 2021; Gilligan et al., 2022; Brown and Andrabi, 2020), there is relatively little work on the effects of unconditional pay increases on teacher sorting and student outcomes (Clotfelter et al., 2008; de Ree et al., 2018; Pugatch and Schroeder, 2018; Cabrera and Webbink, 2020). In line with this literature, we find that such interventions do not prompt increased effort from incumbent teachers. However, we document that they can largely increase student achievement by attracting higher-quality teachers. We borrow tools from the empirical market design literature (Agarwal, 2017; Agarwal and Budish, 2021) to show how to leverage information on teachers' preferences and effectiveness to inform the design of teacher compensation policies. This approach may be relevant in a variety of other settings that typically feature rigid wage profiles, whereby (re-)allocating public employees is likely consequential for equity and efficiency considerations in the provision of public services.

2 Context, Data, and Descriptive Evidence

Our analysis focuses on public primary education for two reasons. First, secondary schools are notably less prevalent in rural regions. Public schools constitute 75% of nationwide primary school enrollment in Peru. In rural areas, more than 26,000 public primary schools cater to 99% of school-aged children. We thus choose to focus on the education level with the broadest coverage across the country. Second, students in primary schools are typically exposed to a single teacher per grade, allowing us to precisely isolate the relative impact of each teacher on students' academic achievement.

2.1 Institutional Background

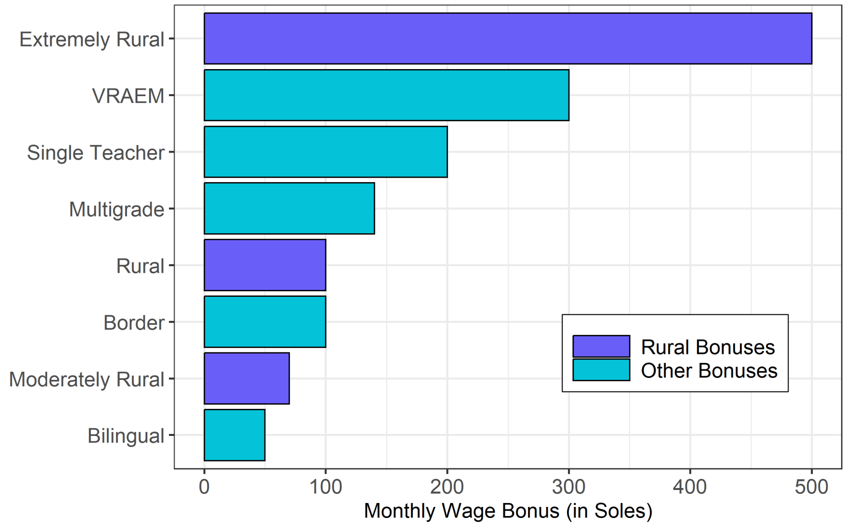
Public school teachers in Perú are hired under two distinct types of contracts. Permanent teachers (*docentes nombrados*) are civil servants with secure employment conditions. Contract teachers (*docentes contratados*) are hired on a fixed one-year contract by a specific school, renewable for an additional year upon approval from the school's principal.

Primary school teachers' earnings in Perú rank second to last among liberal professions, trailing only behind translators and interpreters (INEI, 2016). In 2016, all contract teachers were receiving a fixed base monthly wage of S/ 1,396 (US$ 402) while permanent teachers were receiving S/ 1,550 (US$ 447), irrespective of where they worked.[1] Regardless of contract type, all public-sector teachers receive wage bonuses linked to specific school appointments. Figure 1 illustrates the various wage bonuses that were in place during our analysis period, which range between 4% (for bilingual schools) and 36% (for extremely rural locations, as detailed in Section 3.1) of the monthly base wage. These bonuses are additive such that teachers working in schools meeting multiple criteria (e.g., being both multi-grade and rural) accumulate bonuses.

Historically, the recruitment of public teachers followed a decentralized approach, granting regional and local officials substantial discretion in hiring and resource allocation (Bertoni et al., 2019; Estrada, 2019). To enhance transparency and fairness, the government introduced a nationwide recruitment process in 2015 centralizing all job postings and teacher applications in a unified platform. This process takes place every two years since then. The MoE starts

[1]Both baseline salaries increased to S/ 2,000 in 2018. The average wage for a primary-school teacher in the private sector was S/ 950 (US$ 274) per month (MINEDU, 2014).

Figure 1: Wage Bonuses for Teachers in Disadvantaged Schools



Notes. This figure shows the monetary amount in Peruvian Soles for the different wage bonuses implemented by the Government during the period 2015-2018. VRAEM corresponds to schools located in the region of *Valle de los Rios Apurimac, Ene y Mantaro*, which is extremely poor and presents security concerns due to the activity of drug cartels. Border categorizes schools that are close to the country's international borders.

by making the list of new permanent and contract positions publicly available. All applicants for positions in public schools are required to have a teaching accreditation (i.e. a teacher degree) and to undergo a national competency test. The test comprises three modules, each contributing different weights to the total score: logical reasoning (25 percent), reading comprehension (25 percent), and curricular and pedagogical knowledge (50 percent).[2]

Both types of positions are allocated through a sequential process (for more details, see Appendix A). Permanent positions are only accessible to teachers scoring above selective thresholds and are assigned through a two-step matching process in which schools have a fair amount of discretionary power over hires.[3] Temporary positions are allocated through a serial dictatorship algorithm where schools are required to vertically rank teachers according to their competency score. Applicants first select a school district (corresponding to an administrative unit in Perú) and then choose from the available vacancies within that district by order of their rank. Once a vacancy is filled, it is removed from the list and the subsequent lower-ranked applicant selects their preferred option. This procedure continues until all positions are filled or the lowest-ranked applicant makes their choice. In instances where vacancies

[2]Figure B.1 provides relevant individual-level correlates of teacher performance in each module of the national competency test.

[3]In our data, only about 11% of the applicants are eligible for a permanent teaching position. Among this 11%, half of them end up applying for short-term teaching positions.

persist at the end of this process, unassigned applicants are granted an opportunity to select from remaining openings in other school districts. Positions that remain unfilled nonetheless move on to a decentralized market which potentially involves non-certified teachers.

We restrict the analysis to the recruitment of contract teachers for several reasons. First, the majority of vacancies that are open in rural areas end up being filled by teachers on a temporary contract (85%), thus making the market for short-term teaching positions particularly relevant for the distribution of student achievement in remote locations. Second, as previously indicated, the recruitment of contract teachers is merit-based, as schools are forced to use teacher competency scores as priorities. This unique feature enables us to directly observe teachers' choice sets. Furthermore, it eliminates discretionary hiring practices by school principals, which have been shown to undermine the intended impact of merit-based recruitment reforms for teachers (Biasi et al., 2021; Bates et al., 2022; Ederer, 2023).

2.2 Data Description and Sample Selection

Our analysis draws upon a comprehensive collection of administrative data encompassing all schools, teachers, and students within the Peruvian public education system from 2015 to 2018 (for further details, refer to Appendix B).

The *centralized assignment data* collects information on all job postings and applicants, including the posted wage, applicants' scores in the national competency test and basic demographics such as education, age, gender, and native language. Approximately 25,000 short-term primary school vacancies are posted each year (2016 and 2018) for 60,000 applicants. We observe the realized teacher-school match resulting from the centralized assignment mechanism. Roughly one-third of the applicants are assigned to 75% of the available vacancies. Among the remaining unassigned applicants, half of them find alternative teaching positions within the public sector in later decentralized rounds. The remaining half chooses options outside of the public sector such that we do not observe whether they work in the private sector or in another occupation (if any). Despite the large excess supply of certified applicants, about 15% of the vacancies are eventually filled by non-certified teachers. Importantly, 75% of applicants participating in the 2016 recruitment drive reapply in 2018. This allows us to construct a panel and track teachers' re-matching decisions.

We complement this dataset with several auxiliary sources providing a wide range of additional schools' and applicants' characteristics. The *school census*, collecting information on the number of students and teachers in each school, school infrastructure (libraries, computers, classrooms, sports facilities, etc.), as well as access to basic local amenities like electricity, water, internet, banking, and public libraries. The *MoE's administrative records*, tracking all employed public-sector teachers from 2012 to 2019 and carrying information on their workplace and contract type. Finally, a government-collected *household-level dataset* primarily utilized for targeting social programs and which we can link to approximately two-thirds of applicants. This dataset includes important household characteristics such as their residential location.

Student academic performance is obtained from the national standardized test evaluating proficiency in math and Spanish language (ECE, for its acronym in Spanish). This assessment provides individual test scores for fourth-grade students attending public primary schools in 2016 and 2018. In total, 97% of eligible students take the exam.[4] For approximately 44% of the assigned applicants, we are able to uniquely link teachers to classrooms – and thus to students' academic performance– through an administrative teacher-classroom dataset held by the MoE (SIAGIE, for its acronym in Spanish).

Finally, to elicit teachers' preferences over various non-wage attributes, we conducted an *online survey among applicants* for permanent positions within the 2016 centralized job application process, in which we obtained a response rate of just under 20% (5,553 teachers). As shown in Table B.1, observable teacher characteristics of survey respondents align closely with the overall pool of applicants. Our survey module includes questions on teachers' application decisions, why applicants opted to apply through the centralized mechanism, and their prioritization of school characteristics (see Tables B.2 and B.3).

We make two major sample restrictions throughout the rest of the paper. First, to analyze teachers' choices over job postings, we only consider applicants for whom we have data on the location of their primary residence, as geographical proximity has been shown to be an important predictor of preferences. Second, to identify and estimate teacher effectiveness, we only consider teachers who can be directly linked to the specific classrooms they are assigned to within schools. These sub-samples remain largely representative of the country-wide distribution of applicants to the assignment system across schools and localities (for more details, see Tables B.4 and B.5).

[4]The 2017 examination was cancelled due to nationwide floods.

2.3 Wage Rigidity, Sorting, and Inequalities

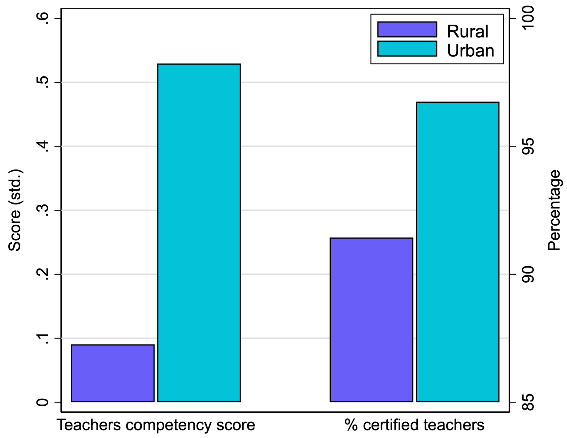
Despite concerted efforts to overcome a historically persistent lack of investment directed toward distressed areas (Bertoni et al., 2020), profound inequalities in education outcomes persist in Perú. These disparities become evident in the contrasting urban-rural gaps in schooling inputs and student performance. A large share of rural schools are situated in villages that lack basic facilities: 20% of these schools do not have electricity, and 40% lack access to drinking water. In contrast, such amenities are almost always available in urban areas. Rural schools are also less than half as likely to have libraries and sports facilities compared to their urban counterparts, and up to two-thirds less likely to have internet access. These disparities, along with other differences in school, teacher, and student characteristics, are detailed in Table B.6.

Figure 2 visually portrays the striking disparities in teaching quality and student achievement between urban and rural primary schools. As shown in Panel A, teachers working in rural schools score, on average, 0.43σ lower at the national competency test than teachers working in urban schools. Urban schools are twice more likely than rural schools to employ a certified teacher. These spatial disparities in teacher quality strongly correlate with inequalities in student achievement. Panel B displays the distribution of students' academic performance at the national standardized evaluation for Spanish and mathematics. We observe a wide urban-rural gap in student achievement ranging between 0.58 and 0.69σ.

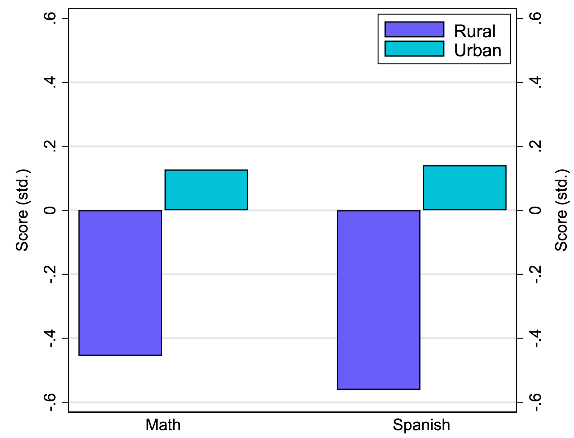
The bottom panels in Figure 2 provide a geographical visualization of these disparities by mapping the distribution of teachers' competency scores and students' test scores across provinces. Panel C shows a significant concentration of highly competent teachers in affluent coastal cities, while their presence is scarce in the highlands and Amazonian regions. The spatial distribution of student achievement displayed in Panel D closely mirrors the distribution of teachers' competency score.

Several factors might explain the observed teacher sorting patterns. Teachers working in impoverished rural areas contend with several challenges, including the scarcity of fundamental school resources, inadequate services, limited access to public goods, and, for many, geographical distance from their home. Our survey of applicants to the 2016 centralized assignment process reveals that non-monetary factors significantly influence teachers' choices over job postings. 44% of teachers highlight 'proximity to home' as a crucial factor guiding their preference ranking (see Table 1). Additionally, attributes such as prestige, safety, and

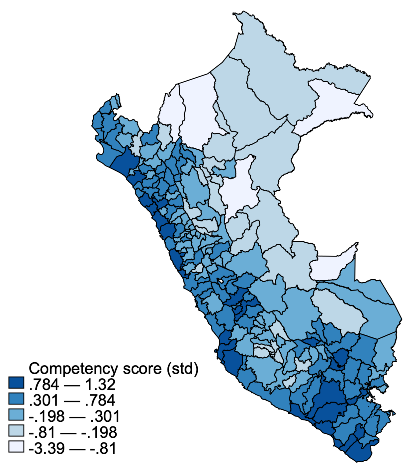
Figure 2: Geographic Distribution of Teacher Competency and Student Achievement



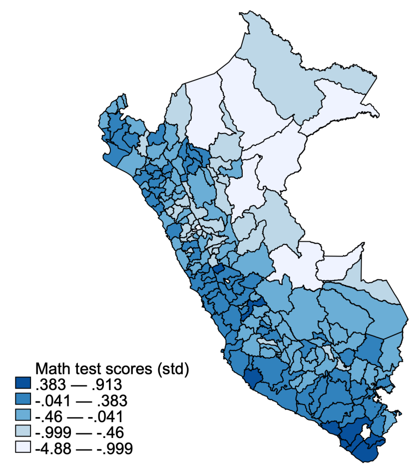
a) Teacher Competency by Urban/Rural



b) Student Achievement by Urban/Rural



c) Teacher Competency by Province



d) Student Achievement (Math) by Province

Notes: These charts show different summary statistics about teachers and students across different regions. Table B.6 in the Appendix presents a broader set of indicators for school and community-level characteristics across urban and rural areas. Rural and urban schools are classified based on whether their locality population is below or above 2,000 inhabitants (respectively). The left part of Panel A shows, separately for rural and urban schools, the average score that teachers obtained in the 2016 and 2018 centralized assignment process, which include both assigned and non-assigned applicants in each assignment round. The right part of Panel A shows, for the universe of teaching positions in primary schools, the share of teachers with teaching certifications. Panel B shows the average student's score in the Spanish and math modules of the national standardized evaluation. The bottom panels of the figure depict the geographical variation in the teachers' competency scores (Panel C) and student's test scores in Math (Panel D) within each province of Perú. In both panels, darker colors indicate higher average scores, with class intervals defined based on the quintiles of the overall score distributions. Both figures are obtained by pooling the data across two school years (2016 and 2018).

cultural considerations are frequently cited as relevant when assessing teaching positions.

As such non-wage amenities tend to be worse in rural areas and rigid wages merely compensate for these differences, teachers' preferences would tend to be skewed toward urban schools. Descriptive evidence from the assignment mechanism confirms this hypothesis as teachers with higher priority (i.e. those with a higher score in the centralized test) dispropor-

Table 1: Applicant Survey (Choice Attributes)

	Rank			
	1^{st}	2^{nd}	3^{rd}	In Top 3
Close to House	44.17	11.66	8.00	63.83
Safe	10.66	24.19	19.25	54.1
Well Connected	9.69	20.62	20.20	50.51
Prestige	17.92	14.12	12.29	44.33
Cultural Reasons	10.61	9.67	12.31	32.59
Good Infrastructure	2.02	8.40	12.86	23.28
Good Students	1.24	4.52	6.08	11.84
Possibility other Jobs	1.93	3.72	4.90	10.55
Career	1.76	3.10	4.09	8.95

NOTES. This table summarizes the answers of 5,553 survey respondents to the question "What are the most important characteristics for your ranked choices?". The first three columns show the share of respondents that ranked the corresponding answer first, second or third. Column (4) shows the share of respondents that listed the corresponding choice in their top 3 reasons. For other determinants of participation into the assignment mechanism and more results on heterogeneity in responses by competency score, see Table B.3.

tionately choose schools in urban areas. As a result, over half of urban postings (compared to a quarter in rural areas) are occupied by teachers ranked in the top 20% of the applicant pool within their school districts. In this context, wage rigidity seems to be an important contributor to unequal teacher sorting and spatial disparities in student achievement.

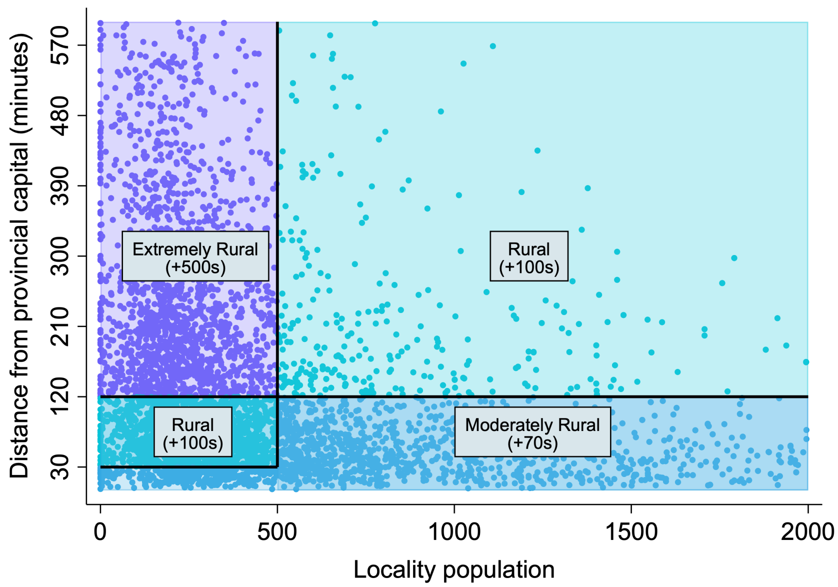
3 Teacher Compensation Reform

In the previous section, we provided descriptive evidence suggesting that wage rigidity induces teachers to sort on non-wage amenities contributing to poorer staffing and worse academic outcomes in rural schools. In this section, we study the role of teacher compensation in addressing these disparities by leveraging a policy that provided wage bonuses to teachers working in rural schools. Appendix C presents a series of additional results and robustness checks related to the empirical analysis discussed in the main text.

3.1 The Rural Wage Bonus Policy

Rural bonuses for contract teachers were introduced in August 2015, i.e., during the school year prior to the first wave of the centralized teacher recruitment drive. Importantly, the reform was only announced briefly before being actually implemented and both incumbent and

Figure 3: The Distribution of Rural Schools and the Wage Bonuses



NOTES. This figure shows the spatial distribution of rural primary schools along the two dimensions that determine the assignment of the rural wage bonus. *Extremely Rural* schools are the dark blue dots, *Rural* are light blue and *Moderately Rural* schools are green.

newly assigned teachers benefited from these bonuses. The policy established three categories of "rurality", which are defined based on specific thresholds of the school locality's population and travel time to the provincial capital. The population of the locality is measured by population counts in the latest available census (2007), while travel time to the provincial capital is computed based on the school's GPS coordinates (taken on-site by government inspectors), the type of roads available, and the most frequent modes of transport.

Figure 3 shows how wage bonuses are distributed across schools based on the two measures defining the rural categories. *Extremely Rural* schools are in localities with less than 500 inhabitants and situated more than 120 minutes away from the province capital. Teachers in these schools receive a bonus of S/ 500 (US$ 144), representing between 25% to 36% of their base wage (depending on the year of the assignment).[5] *Rural* schools are either in localities with less than 500 inhabitants and situated between 30 and 120 minutes away from the province capital, or in localities with 500 to 2,000 inhabitants that are farther than 120 minutes from the province capital. The bonus received by teachers in these schools is S/

[5]The base monthly wage of contract teachers increased from S/ 1,396 in 2016 to S/ 2,000 (US$ 576) by the end of 2017.

100 (US$ 29). Finally, *Moderately Rural* schools are either in localities with 500 to 2,000 inhabitants that are within 120 minutes of the province capital or in localities with less than 500 inhabitants that are within 30 minutes of the province capital. In these schools, teachers receive a bonus of S/ 70 (US$ 20). All other schools are classified as Urban and are therefore not entitled to the rural wage bonus.

There is a large mass of schools around both the time-to-travel (30 minutes and 120 minutes from the provincial capital) and the population cutoffs (500 inhabitants). As localities become more remote, they are more likely to have few inhabitants and predominantly fall into the *Extremely Rural* category. Likewise, as localities become more populated, they are less likely to be remote and fall into the *Moderately Rural* category.

3.2 Regression Discontinuity Design

To study the effects of increasing compensation in rural schools on teacher sorting and the distribution of student achievement, we exploit the sharp thresholds that determine the allocation of the rural wage bonuses in a regression discontinuity (RD) design. The validity of this research design relies on two assumptions: (i) continuity of potential outcomes around the cutoffs, and (ii) independence between the potential outcomes of each unit and the treatment status of other units in a neighborhood of the cutoffs (SUTVA).

Continuity may be violated if the introduction of the rural wage bonus generated incentives for school administrators to manipulate the information used to determine eligibility for the bonus. The population cutoff of 500 inhabitants is based on census data collected before the policy was announced, and as such, is impossible to manipulate. The time-to-travel cutoffs at 30 minutes and 120 minutes are based on GPS measures gathered periodically by government inspectors to account for possible changes in the transportation network and could be subject to manipulation.

Figure C.1 shows a large a significant jump in the density of schools located just above the time-to-travel threshold at 120 minutes in 2018 (but not for 2016). Instead, there are no significant jumps in the density of schools at the population threshold for either years. Table C.1 further shows that pre-determined school and locality-level covariates are smooth around the population cutoff (including the determinants of the other wage bonuses reported in Figure 1), with point estimates that are very small and not statistically different from zero in all but five out of 29 cases for 2016, and in all cases for 2018.[6] Given the possible manipulation of the time-to-travel threshold, we only exploit the variation in wages generated by the population threshold in this part of the analysis.

SUTVA may be violated if the policy triggered spillovers through teacher sorting *around* the population cutoff – e.g. if teachers who chose a position in a high-bonus school just below the threshold would have otherwise chosen a position just above the threshold in the absence of the wage bonus policy. We provide suggestive evidence that this is unlikely to have happened. First, we document that localities around the population threshold are not necessarily geographically close to one another, limiting the possibility of spillovers. In fact, for any given school close to the threshold, the median (geodesic) distance to its first, second, and third closest school across the cutoff is approximately 10km, 20km, and 30km, respectively. Second, we leverage data on teachers' previous job to directly test for whether teachers choosing a school below the threshold are more likely to have been working in a school above the threshold the year before. Figure C.2 shows that high-paying schools did not disproportionately attract teachers from low-paying schools close to the threshold but rather drew them evenly from across the entire country, and particularly from urban areas.

We consider the following RD specification:

$$y_{ijt} = \gamma_0 + \gamma_1 \mathbf{1}(pop_{jt} < pop_c) + g(pop_{jt}, pop_c) + \delta_t + u_{ijt}, \tag{1}$$

where y_{ijt} is an outcome variable for teacher (or teaching position) i in school j at time t, $g(\cdot)$ is a flexible polynomial in the population of the locality of the school on both sides of the population cutoff, δ_t denotes time indicators for the specific year of the recruitment drive, and u_{ijt} is an error term clustered at the school-year level. Equation (1) can be further used to study the effects of the wage bonuses on student outcomes, such as a standardized test score for student l taught by teacher i. The parameter of interest is γ_1, which represents the average outcome difference between teachers or students in localities that are just below or above the population cutoff. We estimate γ_1 non-parametrically using the robust estimator proposed by Calonico et al. (2014) through bias-corrected local linear regressions that are defined within the mean squared error optimal bandwidths. The RD estimates reported in

[6] Table C.1 also shows that manipulation of the 120-minutes threshold seems to be smooth at the population threshold. This addresses concerns that schools might have manipulated the 120-minutes threshold to a greater extent below the 500 inhabitants threshold (where the benefits from manipulating are larger), which would threaten the validity of our research design.

this section are robust to alternative specifications and estimation choices, which we show in Figure C.3.

Teachers recruited in localities with slightly less than 500 inhabitants earn, on average, about S/ 225 (US$ 65) more than teachers hired in localities that are just above the cutoff (corresponding to an increase of about 11%)[7] In what follows, we refer to the observations in rural areas that are above the population cutoff of 500 inhabitants as 'Low-Bonus' and to those below the cutoff as 'High-Bonus'.

3.3 Teacher Sorting Patterns

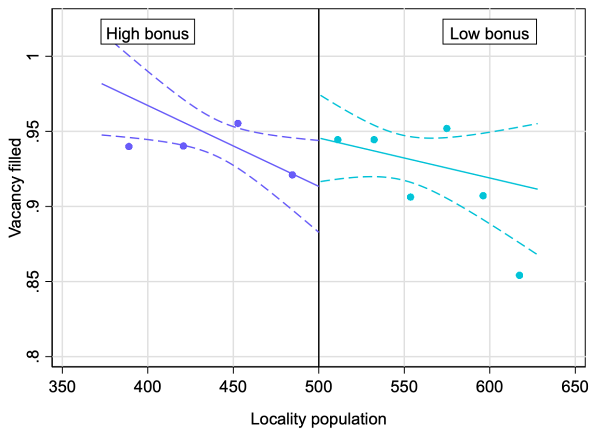
We start by investigating how teachers' school choices responded to the wage increase at the threshold. A first-order objective of the centralized assignment system is to fill as many positions as possible. The graphical evidence displayed in Panel A in Figure 4, along with the corresponding RD estimates reported in Column (1) of Table 2, show that the wage increase at the threshold had a small and statistically insignificant effect on the probability that a vacancy is filled by a certified teacher. While it may be the case that the wage bonus policy induced teachers to choose positions that would have otherwise been unfilled, this margin does not seem to be relevant for vacancies located close to the population threshold where the share of filled vacancies is already very high (95% in the Low-Bonus RD sample).

We next restrict our attention to filled vacancies on either side of the threshold and consider as outcome variables a measure of teachers' preference intensity and the competency score of the assigned teachers.[8] The preference index takes the value of zero if the position is filled last and the value of one if the position is filled first in a school district. We find that high-paying vacancies were filled at a significantly faster rate than low-paying vacancies. Column (2) of Table 2 shows that the average vacancy in localities just above the population cutoff is filled by a teacher ranked in the 33^{rd} percentile ($= 1 - 0.67$) of the score distribution of applicants in that school district, while schools that offer a higher wage bonus fill the

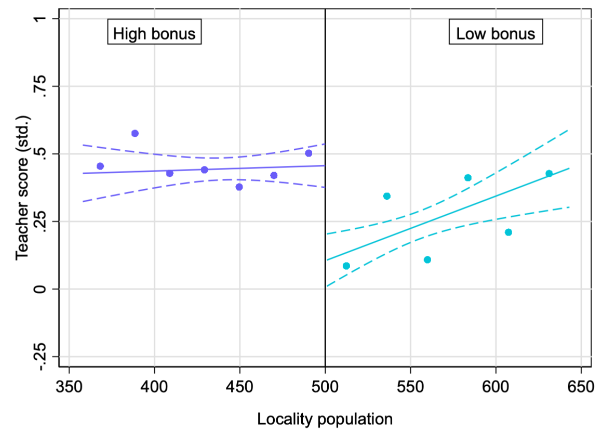
[7] We exclude all urban schools and rural schools in localities within 30 minutes of the province capital since, for them, crossing the population cutoff does not lead to an increase in the rural wage bonus. Note that the wage increase at the threshold would be larger if we restrict our estimation sample to schools above the time-to-travel cutoff of 120 minutes (see Table C.2). However, this alternative approach would imply conditioning on a partially manipulated variable and decrease the sample size.

[8] To deal with the potential issue of endogenous selection into the sample, in Table C.3 we report RD bounds for both of these outcome variables using the approach outlined in Gerard et al. (2020). The bounds are in general quite tight, thereby suggesting that the censorship in the density of the observations due to the fact that some vacancies remain unfilled is inconsequential for the RD estimates.

Figure 4: Teacher Choices and Sorting



a) Vacancy Filled



b) Competency Score

NOTES. This figure shows how applicants' preferences and quality vary based on the difference between the 500-inhabitants cutoff and the population of the community where the school is located. In Panel A the outcome is an indicator variable that is equal to one if a vacancy was filled by a (certified) contract teacher during the centralized assignment, while Panel B uses the standardized score obtained in the centralized test by the newly-assigned contract teacher. Each marker indicates the median of the outcome variable within each bin, defined following the IMSE-optimal evenly spaced method by Calonico et al. (2015). Solid lines represent the predictions from linear regressions estimated separately for observations to the left and to the right of the cutoff, assuming a triangular kernel function. Dashed lines are 95% asymptotic confidence intervals.

Table 2: Teacher Choices and Sorting

	(1) Vacancy filled	(2) Preferences	(3) Teacher Score (Std.)
High Bonus	-0.043	0.103	0.386
	(0.040)	(0.035)	(0.137)
Bandwidth	127.521	157.452	141.447
Schools	715	850	764
Observations	1851	2080	1870

NOTES. This table reports the effect of crossing the population threshold on different outcomes. In Column (1) the outcome variable is an indicator for whether the vacancy was filled by a certified teacher in the assignment process for contract teachers. Column (2) shows the effect on the rank in which a vacancy was chosen within a school district (normalized so that it takes values from zero to one). Column (3) uses as outcome variable the standardized competency score obtained by the teachers in the centralized test. Cells report the bias-corrected regression-discontinuity estimates obtained using the robust estimator proposed in Calonico et al. (2014). Regressions are defined within a mean-square error optimal bandwidth (BW), reported at the bottom part of the table. Standard errors are clustered at the school×year level.

position with an applicant ranked in the 23^{rd} percentile on average $(= 1 - 0.67 - 0.10)$.

Given that priority for choosing vacancies in the recruitment drive for short-term positions is determined by teachers' competency scores only, we expect that the increase in preferences for high-bonus schools documented in Column (2) translates into an increase in quality of recruited teachers. Both the graphical evidence in the left panel of Figure 4 and the RD estimates in Column (3) of Table 2 document that teachers who select into high-bonus schools have higher competency scores than those who choose a position in low-bonus rural schools. The effect is sizeable, at 0.39σ of the overall distribution of the competency score (p-value=0.005).[9] As a benchmark, the average gap in teacher competency between *Extremely Rural* schools and other rural schools is approximately 0.30σ, whereas the overall urban-rural gap is about 0.50σ.

Additional evidence seems to support the limited role of spillovers around the 500 population threshold. In particular, Table C.4 reports the results of a difference-in-discontinuity analysis, where we compare threshold-crossing estimates on teacher competency before and after the wage bonus policy was introduced. The magnitude of the estimates are very close to what is reported in Table 2. Interestingly, the estimated coefficients for the post-implementation dummy are positive throughout and statistically insignificant. This evidence suggests that the increase in teacher competency in high-bonus schools did not come at the expense of lowering the quality of teachers in lower-paying schools located near the threshold.

In sum, we find that increasing wages in disadvantaged locations effectively steered teachers' labor supply toward the targeted job postings. The observed change in teachers' behavior does not seem to have significantly affected the probability of creating new matches, but instead lead to an inflow of higher-quality teachers in high-paying schools.[10]

3.4 Student Achievement

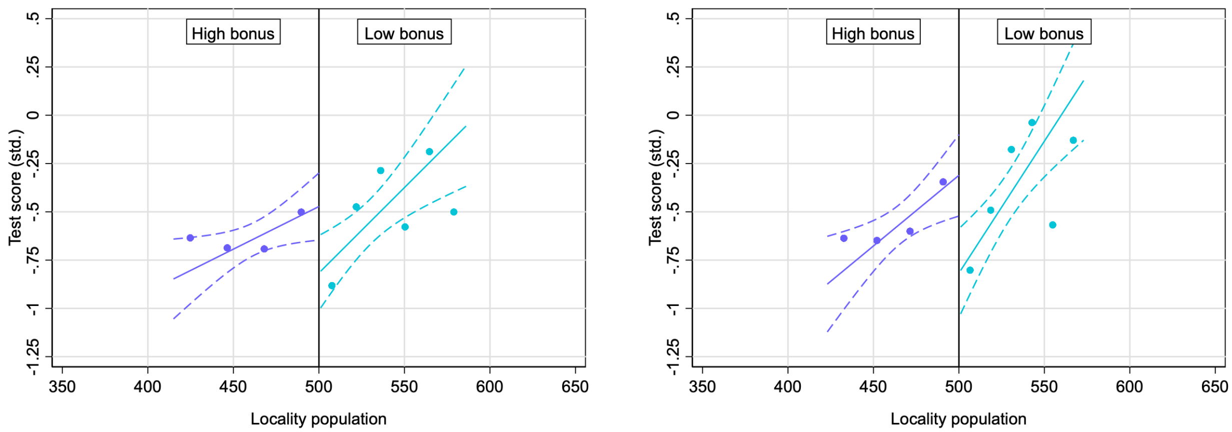
Offering higher wages for positions in rural locations could potentially improve student outcomes mainly through two mechanisms: (i) the recruitment of higher-quality teachers and (ii) an increased effort from incumbent teachers. To parse out the effort margin from the recruitment margin, we leverage matched teacher-classroom data and run the analysis separately for classrooms taught by newly recruited teachers and for those taught by incumbent teachers.[11] We use fourth-grade student test scores in math and Spanish in 2016 and 2018 as measures of academic achievement.

[9] Despite finding similar effects on applicants' preference intensity in the permanent teacher recruitment drive, we do not find evidence of increased competency for hired permanent teachers (see Figures C.4 and C.5). This evidence supports the hypothesis that efforts aimed at increasing teachers' labor supply in distressed areas might not translate into upward pressure on teacher quality in the presence of discretionary hiring practices (Biasi et al., 2021; Bates et al., 2022; Ederer, 2023).

[10] This result is consistent with recent evidence reported in Agarwal (2017), which documents that the primary effect of financial incentives was to increase the quality, not numbers, of medical residents in rural America.

[11] As shown in Table C.5, both the number of vacancies per school and the probability that a vacancy is open in the centralized assignment are smooth around the RD cutoff.

Figure 5: Wage Bonus and Student Achievement



a. Spanish **b. Math**

NOTES. This figure shows how student achievement – as measured by the score on the ECE standardized test – varies based on the difference between the 500-inhabitants cutoff and the population of the community where the school is located. We consider the sample of schools with at least one opening for a contract teacher in the 2015 or 2017 recruitment drives. Panel A considers the (standardized) test score in Spanish, while Panel B that in math. Each marker indicates the median of the outcome variable within each bin, defined following the IMSE-optimal evenly spaced method by Calonico et al. (2015). Solid lines represent the predictions from linear regressions estimated separately for observations to the left and to the right of the cutoff, assuming a triangular kernel function. Dashed lines are 95% asymptotic confidence intervals.

Figure 5 plots the threshold crossing effects on students' academic achievement in classrooms taught by newly recruited teachers. Students exposed to higher-quality teachers attracted by the wage increase at the threshold perform significantly better in standardized achievement tests. Column (1) of Table 3 shows sizable effects ranging from 0.40σ in Spanish to 0.58σ in math. We find no evidence of an increased effort from incumbent teachers in response to the wage increase at the threshold. Column (2) of Table 3 shows very small effect sizes of the bonuses when restricting the sample to schools with no open vacancies. Finally, Column (3) displays the aggregate effect on the overall sample. Students in high-paying schools perform overall better in Spanish and math, with effect sizes of 0.2-0.3σ.

Taken together, this evidence strongly suggests that the wage bonus policy triggered an inflow of higher quality teachers translating into large improvements in students' learning outcomes. While there may also be an effort response to the wage incentives for newly recruited teachers, the evidence reported in Table C.6 documents little if no composition effects at the population cutoff along teachers' observable characteristics like gender, age, experience, native mother tongue, or having a university degree. We also do not find supportive evidence for alternative mechanisms through which the rural wage bonus may affect student

Table 3: Wage Bonus and Student Achievement

Panel A: Dependent Variable is Spanish Test (z-score)			
	(1) Vacancy	(2) No vacancy	(3) All
High Bonus	0.395	-0.004	0.232
	(0.152)	(0.127)	(0.088)
Bandwidth	107.818	148.920	105.822
Schools	264	451	832
Observations	4635	6773	16681
Panel B: Dependent Variable is Math Test (z-score)			
	(1) Vacancy	(2) No vacancy	(3) All
High Bonus	0.579	0.067	0.317
	(0.193)	(0.143)	(0.105)
Bandwidth	85.848	155.174	95.638
Schools	220	470	764
Observations	3939	7039	15363

Notes. This table reports the effect of crossing the population threshold on student achievement in Math and Spanish. The outcome variables are standardized test scores in Spanish (Panel A) and Math (Panel B) for students in fourth grade. The sample in Columns (2) and (3) is split based on whether the school had an open vacancy (of any type) in the 2015 and/or 2017 centralized recruitment drives. Each cell reports the bias-corrected regression-discontinuity estimates obtained using the robust estimator proposed in Calonico et al. (2014). Regressions are defined within a mean-square error optimal bandwidth (BW), reported at the bottom of the table. Standard errors are clustered at the school level.

outcomes. For example, the wage increase at the threshold did not affect student achievement by changing the size and composition of the teaching staff (see Table C.7). Additionally, we do not find any effect on teacher retention (see Table C.8) that might have independently influenced student achievement.

4 An Empirical Model of Teacher Sorting and Student Achievement

Despite our positive assessment of the efficacy of the wage bonus policy in shaping teacher sorting and diminishing spatial disparities in student achievement, the extent to which teacher compensation reform can effectively address the prevailing spatial inequalities in education outcomes remains uncertain. To answer this question, we build a model of teacher sorting across schools and student achievement in which teachers are heterogeneous in their preferences over jobs and in both their absolute and comparative advantages in teaching different student types.

4.1 Wages, Preferences, and Equilibrium

Consider an economy with a set of teachers indexed by $i = 1, ..., n_w$, a set of schools indexed by $j = 1, ..., n_m$, and a set of students indexed by $l = 1, ..., n_s$. Time is indexed by t. We define a teacher-school match as a mapping μ_w, where $\mu_w(i,t) = j$ if teacher i is matched to school j in period t. Similarly, we define a student-teacher match as a mapping μ_s, where $\mu_s(l,t) = i$ if student l is matched to teacher i in period t.

Wages are fixed by the government through a known deterministic rule. They are posted ex-ante for each available vacancy, observed by applicants before making their choices, and cannot be renegotiated. Teachers receive the same fixed baseline wage, irrespective of the school they work in, and a set of wage bonuses, which vary with pre-determined locality and school characteristics (see Figure 1).

We define the indirect utility that teacher i gets from being matched with school j in year t as:

$$U_{ijt} = w_{jt} + \alpha_i^{-1}(u(a_{jt}, x_{it}) + \epsilon_{ijt}), \tag{2}$$

where w_{jt} is the posted wage in school j in year t. The parameter α_i controls teacher i's taste for wages relative to non-pecuniary amenities captured by $u(a_{jt}, x_{it}) + \epsilon_{ijt}$. $u(a_{jt}, x_{it})$ is an arbitrarily flexible function of observed school and locality characteristics (a_{jt}) as well as observed teacher characteristics (x_{it}) giving rise to vertical and horizontal differentiation between jobs offering the same wage. ϵ_{ijt} is an unobserved taste shock introducing further horizontal differentiation in teachers' preferences.

A significant share of applicants remain unassigned at the end of the centralized assignment procedure (see Section 2.2). A first group of teachers chooses to remain in the public sector by either filling a school vacancy that becomes available later in the academic year or by taking a teaching position outside of the pedagogical area in which they originally applied. We denote this alternative as $j = p$ and normalize its utility as follows:

$$U_{ipt} = \alpha_i^{-1}(x'_{it}\beta_p + \epsilon_{ipt}), \tag{3}$$

where $x'_{it}\beta_p$ captures heterogeneity across teachers in the value of participating in this secondary decentralized market.

A second group of teachers leaves the public sector for another occupation, which might include teaching in the private sector or staying unemployed. We denote this outside option as $j = 0$ and normalize its utility as follows:

$$U_{i0t} = \alpha_i^{-1}(\beta_i + \epsilon_{i0t}), \tag{4}$$

where β_i is a teacher-specific coefficient capturing observed and unobserved heterogeneity in the pecuniary and non-pecuniary benefits of choosing to leave the public sector. We assume that both outside options $j = p$ and $j = 0$ are in the choice sets of all teachers.

As described in Section 2.1, applicants are ranked based on their competency score within their school district and choose, by order of priority, their preferred school among those that still have open vacancies. This precludes the existence of blocking pairs, i.e. teachers would not be accepted by a school they strictly prefer to the one they choose. This directly implies that the resulting match is stable within each district (Roth and Sotomayor, 1992). As teachers choose their district before knowing which options are available to them, they might realize ex-post that they would have preferred a feasible school in another district. However, additional data coming from the aftermarket show that less than one percent of applicants choose to re-match in another school district after the first round of the assignment. We thus assume that the resulting equilibrium μ_w is stable both within and across school districts. As a result, we can characterize the equilibrium teacher-school match μ_w^* as follows:

$$\mu_w^*(i,t) = \underset{j \in \Omega(s_{it})}{\arg\max}\, U_{ijt}, \tag{5}$$

where $\Omega(s_{it})$ is teacher i's feasible choice set: the set of schools that still have remaining vacancies after all applicants with a score larger than s_{it} made their choice. As $\Omega(s_{it})$ is directly observed in the data, the matching equilibrium can be rewritten as the outcome of a discrete choice problem with personalized choice sets (Fack et al., 2019).

4.2 Student Achievement

We consider a potential outcomes framework that maps any potential teacher-student match into the distribution of student achievement. Recent work has highlighted the presence of substantial heterogeneity in school value-added across students with different demographics

(Walters, 2018; Abdulkadiroğlu et al., 2020). We follow a similar approach and allow teacher effectiveness to be heterogeneous and vary with students' observed types as in Ahn et al. (2023). Accounting for potential match effects in teacher value-added is crucial to quantify the potential equity and efficiency gains from teacher reallocation and design more effective teacher compensation policies. In particular, we posit that student l's test score in fourth grade when being enrolled in school j and taught by teacher i in year t can be written as follows:

$$Y_{lijt} = c_{jt}'\beta + z_{lt}'\delta_i + \nu_{lijt}, \tag{6}$$

such that, by construction, $\mathbb{E}[\nu_{lijt}|c_{jt}, z_{lt}] = 0$. In this framework, $c_{jt}'\beta$ captures observed determinants of school and classroom effects that are not explained by teacher effectiveness while δ_i measures the average effectiveness of teacher i and the returns of student characteristics z_{lt} when being exposed to teacher i. Equation (6) can be rewritten as follows:

$$Y_{lijt} = c_{jt}'\beta + z_{lt}'\bar{\delta} + z_{lt}'(\delta_i - \bar{\delta}) + \nu_{lijt}, \tag{7}$$

where $z_{lt}'\bar{\delta}$ corresponds to student l's baseline ability and $z_{lt}'(\delta_i - \bar{\delta})$ corresponds to teacher i's treatment effect on student l. We include in c_{jt} a set of observed school and classroom characteristics such as classroom- and school-level averages of students' lagged test scores (i.e., in second grade), age, gender, and ethnicity. We include in z_{lt} an intercept, students' lagged scores, gender, and ethnicity. We normalize z_{lt} to be mean zero such that the coefficient associated with the intercept corresponds to the average treatment effect of teacher i. We do not allow for complementarities between teacher effectiveness and other schooling inputs, which have limited within-teacher variation in a short panel. As noted by Ahn et al. (2023), this framework nests other approaches used in the teacher value-added literature, which either assume constant effects (Chetty et al., 2014b) or constant effects within student subpopulations (Biasi et al., 2021; Bates et al., 2022).

4.3 Assumptions

We define $\theta_i = (\log \alpha_i, \beta_i)$ and impose the following assumptions on the payoff functions as well as the student achievement production function:

Assumption 1 . *(i) The function u is bounded and differentiable.*

(ii) ϵ_{ijt} *are iid across* i*,* j *and* t *and are Extreme Value Type I distributed.*

Assumption 2 *.* $\mathbb{E}[\nu_{lijt}|c_{jt}, z_{lt}, i = \mu_s(l,t), j = \mu_w(i,t)] = 0$

Assumption 3 *.* $(\theta_i, \delta_i)|x_{it} \sim \mathcal{N}(\gamma(x_{it}), \Sigma)$ *where* $\Sigma = \begin{pmatrix} \Sigma_{\theta,\theta} & \Sigma_{\theta,\delta} \\ \Sigma_{\delta,\theta} & \Sigma_{\delta,\delta} \end{pmatrix}$ *and* $\gamma(x_{it}) = \begin{pmatrix} x'_{1it}\gamma^{\theta} \\ x'_{2it}\gamma^{\delta} \end{pmatrix}$,

where x_{1it}, x_{2it} are sub-vectors of x_{it}.

Assumption 1 (i) is a mild regularity condition satisfied by most functional forms used in practice. Assumption 1 (ii) might seem restrictive as it often implies that conditional choice probabilities exhibit the Independence of Irrelevant Alternative (IIA) property. This is precluded under Assumption 3, which introduces unobserved preference heterogeneity through the correlated random coefficients θ_i. These capture time-invariant idiosyncratic factors that affect how teachers substitute non-pecuniary benefits against wages, as well as how they value the outside option. Allowing for these coefficients to be correlated is crucial to correctly pin down substitution patterns from the outside option in response to changes in wages across job postings. For example, if the correlation between α_i and β_i is positive, drawing teachers from the outside option might be easier than inducing currently employed teachers to relocate across schools.

Assumption 2 posits that the allocation of teachers to students is as good as random after conditioning on students' and schools' characteristics (including students' lagged test scores).[12] The school and teacher value-added literature argues that controlling for students' lagged test scores is sufficient to ensure the validity of this assumption (Chetty et al., 2014b; Angrist et al., 2017). We evaluate the validity of this design in our setting in Section 5.3 by leveraging the rural bonus policy and show that the estimated threshold crossing effects on predicted teacher value-added closely match their counterparts on student achievement.

Finally, getting precise estimates of teacher effectiveness may be challenging when teachers are not exposed to many students throughout the available data. Recent work has highlighted the benefits of using Empirical Bayes (EB) methods when making decisions involving ranking and selection in the presence of such statistical uncertainty (Gu and Koenker, 2023; Kline et al., 2023). We argue that EB shrinkage is particularly relevant for assignment problems with heterogeneous treatment effects. By attenuating the estimates of teachers' comparative

[12] Individual effects in linear models are typically poorly identified from data on sparse networks (Jochmans and Weidner, 2019). A more flexible two-way fixed effects empirical strategy would thus not give reliable estimates of both student and teacher effects.

advantage proportionally to how imprecise they are, this approach limits the potential for efficiency losses by prioritizing teacher-student matches that are high-value with high certainty. We provide evidence from Monte Carlo simulations supporting this claim in Section 6.2. We assume that δ_i are *iid* draws from a common Gaussian prior through Assumption 3. The mean of the prior is anchored on a sub-vector of teacher characteristics x_{2it} including their certification status, competency score, gender, and ethnicity. Importantly, we allow the teacher value-added coefficients δ_i to correlate with the random coefficients θ_i. This has the advantage of reducing the variance of the posterior by conditioning on teachers' observed characteristics and choice behavior. It also allows us to study the link between teacher effectiveness and latent factors captured by their choice behavior that could reveal their intrinsic motivation. Finally, it captures any potential selection on unobserved teacher quality that would occur as a result of counterfactual compensation policies.

4.4 Identification

We first characterize the mapping between preferences and equilibrium sorting. From Equation (5) and Assumptions 1 and 3, we can write the conditional probability of observing the realized matching history of a given teacher i $(\mu_w^*(i,t))_{t=1}^T$ as:

$$\mathbb{P}(\{\mu_w^*(i,t)\}_{t=1}^T | x_{it}, \boldsymbol{w}, \boldsymbol{a}, \Omega(s_{it})) = \tag{8}$$
$$\int \prod_{t=1}^{T} \frac{\exp\{\alpha_i w_{\mu_w^*(i,t)t} + u(a_{\mu_w^*(i,t)t}, x_{it})\}}{\exp\{\beta_i\} + \exp\{x_{it}'\beta_p\} + \sum_{k\in\Omega(s_{it})} \exp\{\alpha_i w_{kt} + u(a_{kt}, x_{it})\}} \phi(\theta_i | \gamma^\theta, \Sigma_{\theta,\theta}) d\theta_i.$$

We directly identify the left-hand side of Equation (8) from panel data on realized matches, observed amenities, wages, teacher characteristics, and choice sets. This is a standard expression for mixed-logit models that has been shown to be invertible in Fox et al. (2012), such that the mapping between primitives $(u, \gamma^\theta, \Sigma_{\theta,\theta})$ and our data is one-to-one. More specifically, observing two choices per teacher and having variation in choice sets within and across teachers is crucial to identify the distribution of random coefficients (Berry et al., 2004). For example, teachers choosing the outside option (4) in one year and a public school in another year provide essential variation to identify the off-diagonal elements of $\Sigma_{\theta,\theta}$.

We use the panel matched teacher-classroom data described in Section 2.2 to identify the parameters of the student achievement production function (β, γ^δ and $\Sigma_{\delta,\delta}$). Under

Assumption 2, we can show that:

$$\delta_i = \mathbb{E}\left[z_{\mu_s(i,t)t} z'_{\mu_s(i,t)t}\right]^{-1} \mathbb{E}\left[z_{\mu_s(i,t)t}(Y_{\mu_s(i,t)ijt} - c'_{jt}\beta)\right] \tag{9}$$

$$\beta = \mathbb{E}\left[c_{jt}c'_{jt}\right]^{-1} \mathbb{E}\left[c_{jt}(Y_{l\mu_s(l,t)jt} - z'_{lt}\delta_{\mu_s(l,t)})\right], \tag{10}$$

which can be solved to recover β and δ_i for each teacher i, provided that there is sufficient within-teacher variation in student characteristics for $\mathbb{E}\left[z_{\mu_s(i,t)t} z'_{\mu_s(i,t)t}\right]$ to be invertible. Under Assumption 3, we can then write the conditional probability density function of δ_i as follows:

$$f(\delta_i | x_{2it}) = \phi(\delta_i | x'_{2it}\gamma^\delta, \Sigma_{\delta,\delta}). \tag{11}$$

With knowledge of δ_i from Equation (9), we can identify the left-hand side of Equation (11) and use variation in x_{2it} to recover $(\gamma^\delta, \Sigma_{\delta,\delta})$.

Finally, to identify $\Sigma_{\theta,\delta}$, we link the centralized assignment data with the matched teacher-classroom data. Under Assumption 3, we can rewrite the probability density function (11) conditional on θ_i as follows:

$$f(\delta_i | x_{1it}, x_{2it}, \theta_i) = \phi(\delta_i | x'_{2it}\gamma^\delta + \Sigma_{\delta,\theta}\Sigma^{-1}_{\theta,\theta}(\theta_i - x'_{1it}\gamma^\theta), \Sigma_{\delta,\delta} - \Sigma_{\delta,\theta}\Sigma^{-1}_{\theta,\theta}\Sigma_{\theta,\delta}) \tag{12}$$

Using (12) we can write the joint probability of observing teacher i's matching history $(\mu^*(i,t))_{t=1}^T$ along with her teaching effectiveness δ_i as follows:

$$\mathbb{P}(\{\mu^*(i,t)\}_{t=1}^T, \delta_i | x_{it}, \boldsymbol{w}, \boldsymbol{a}, \Omega(s_{it})) = \tag{13}$$
$$\int \prod_{t=1}^{T} \frac{\exp\{\alpha_i w_{\mu^*(i,t)t} + u(a_{\mu^*(i,t)t}, x_{it})\}}{\exp\{\beta_i\} + \exp\{x'_{it}\beta_p\} + \sum_{k\in\Omega(s_{it})} \exp\{\alpha_i w_{kt} + u(a_{kt}, x_{it})\}} f(\delta_i | x_{1it}, x_{2it}, \theta_i)\phi(\theta_i | \gamma^\theta, \Sigma_{\theta,\theta}) d\theta_i,$$

which we can invert to recover $\Sigma_{\theta,\delta}$ conditional on already knowing $u, \Sigma_{\theta,\theta}, \Sigma_{\delta,\delta}, \gamma^\theta, \gamma^\delta$. See Appendix D.1 for a proof.

4.5 Parameterization and Estimation

We parameterize $u(a_{jt}, x_{it})$ as a flexible function of a wide range of schools' and teachers' characteristics:

$$u(a_{jt}, x_{it}, \boldsymbol{\Gamma}) = x'_{1it}\Gamma_1 q_{jt} + x'_{1it}\Gamma_2 d_{ijt} + x'_{1it}\Gamma_3 m_{ij} + \kappa_j, \tag{14}$$

where the vector q_{jt} contains a summary index of different indicators measuring the quality of infrastructure in the locality (see Panel D of Table B.6) as well as an asset-based measure of poverty at the individual level that we aggregate at the locality level (which enters with a negative sign in the overall index). It further includes a set of indicator variables for whether a given school belongs to specific regimes that determine eligibility for other wage bonuses such as multi-grade, single-teacher, bilingual, and/or to the specific geographic areas (see Figure 1). By accounting for those characteristics as well as flexible polynomials in the locality's population and the time-to-travel (in hours) to the province capital, the residual identifying variation in wages stems from the discrete changes at the thresholds induced by the rural wage policy (see Section 3.1.)

Moving costs and other costs associated with switching jobs are captured by d_{ijt}, a vector containing linear splines of the geodesic distance between the location of school j and teacher i's home location, as well as an indicator for whether teacher i was working in school j in the previous year. The vector m_{ij} contains ethnolinguistic match effects, indicating whether teacher i's indigenous native language (if any) and school j's secondary language of instruction (if any) coincide (see Section 2.1). These capture language barriers that teachers might face when working in a school where the prevalent language is different from theirs and, more broadly, any specific taste for living in a community with shared linguistic or cultural traits. To avoid sparseness in the data, we only consider the two most prominent linguistic groups (*Quechua* and *Aymara*).

Preferences over all these non-pecuniary aspects of the job are allowed to vary with x_{1it}, a sub-vector of teacher characteristics, including gender, experience in the public and private sector, and competency score. Finally, we include a set of time-invariant province fixed effects κ_j that capture vertical differentiation on unobserved amenities across these geographic areas.

Estimation is done in two steps. We first estimate (β, δ_i) by taking the empirical counterparts of (9)-(10) and solving for the unknown parameters. We then estimate $(\mathbf{\Gamma}, \gamma, \Sigma)$ by maximizing the following log-likelihood function:

$$L(\mathbf{\Gamma}, \gamma, \Sigma) = \sum_{i=1}^{n_w} \sum_{t:\{\mu^*(i,t) \neq \emptyset\}} \log \mathbb{P}\left((\mu^*(i,t))_{t=1}^{T}, \hat{\delta}_i | x_i, \boldsymbol{w}, \boldsymbol{a}, \Omega(s_{it}) \right), \tag{15}$$

where we sum over all teachers and over the years in which they participated in the assignment mechanism (see Appendix D.2 for more details).

5 Estimation Results

This section discusses the implications of the model estimates and assesses their validity through various measures of fit. The full set of model estimates and additional validity checks on the predictions of the model are displayed in Appendix D.5.

5.1 Teachers Preferences

We report the implied willingness-to-pay for several non-pecuniary job and locality characteristics in Table 4. Consistently with the survey evidence displayed in Table 1, teachers attach a high value to local amenities, geographical and cultural proximity, as well as good teaching conditions (Columns 1 and 2). The average teacher would be willing to give up 10% of their wage for a one standard deviation increase in the quality of school infrastructure and local amenities. We estimate very high moving costs, as teachers would be on average willing to pay between 1 to 10% of their base wage to avoid traveling one kilometer further away from home.[13] Interestingly, moving costs are highly non-linear as the cost of traveling an additional kilometer decreases by 43% after 20km. Teachers have a high willingness to pay to teach in a school where the language of instruction is the same as theirs. The average willingness to pay to avoid teaching in remote schools close to the country's borders, in multi-grade, or single-teacher schools ranges between 20% and 88% of the base wage. More importantly, Columns (3) to (6) of Table 4 show that these willingness-to-pay estimates are highly heterogeneous across teachers. Non-pecuniary attributes thus induce substantial vertical and horizontal differentiation across schools and locations.

Jobs located in rural areas tend to have worse teaching conditions, worse local amenities, and can potentially be very remote, compared to jobs in urban areas. This translates into large spatial differences in the utility associated to a given job. Figure 6 plots the distribution of non-pecuniary differences in utility along the population and distance to provincial capital margins, between rural schools and the average urban school. Utility here is measured in monthly wages (in Peruvian Soles). We find that, on average, utility differences are merely compensated by the current wage bonus scheme. The vast majority of the schools categorized as *Extremely Rural* (i.e. population<500 and time-to-travel>120min) would need to offer an

[13] One kilometer measured in geodesic distance may entail substantial travel time in some regions of Perú due to poor road infrastructure quality. This may explain the large magnitudes of the estimated moving costs.

Table 4: Monthly Willingness to Pay for Non-Pecuniary Job Characteristics

	Mean		10% Quantile		90% Quantile	
	Soles	% Wage	Soles	% Wage	Soles	% Wage
	(1)	(2)	(3)	(4)	(5)	(6)
Amenities, Infrastructure and Remoteness						
Amenity/Infrastructures	200	10	30	2	440	22
Closer to Home by 1km						
$0 \leq$ Distance < 20	200	10	33	2	443	22
$20 \leq$ Distance < 100	113	6	23	1	243	12
Distance ≥ 100	20	1	3	0	43	2
Ethnolinguistic Proximity						
Same Language: Spanish	2,777	139	393	20	6,180	309
Same Language: Quechua	986	49	303	15	1,929	96
Same Language: Aymara	3,264	163	656	33	6,976	349
Teaching Conditions						
No Border	406	20	-97	-5	1,122	56
No Multigrade	962	48	147	7	2,121	106
No Single Teacher	1,758	88	120	6	4,123	206

NOTES. This Table displays the mean and the bottom and top deciles of the distribution of the estimated willingness to pay for several non-wage characteristics. Each number is displayed in monthly wage equivalent in Peruvian soles, and in percentage of the base wage in 2018 (2000 soles).

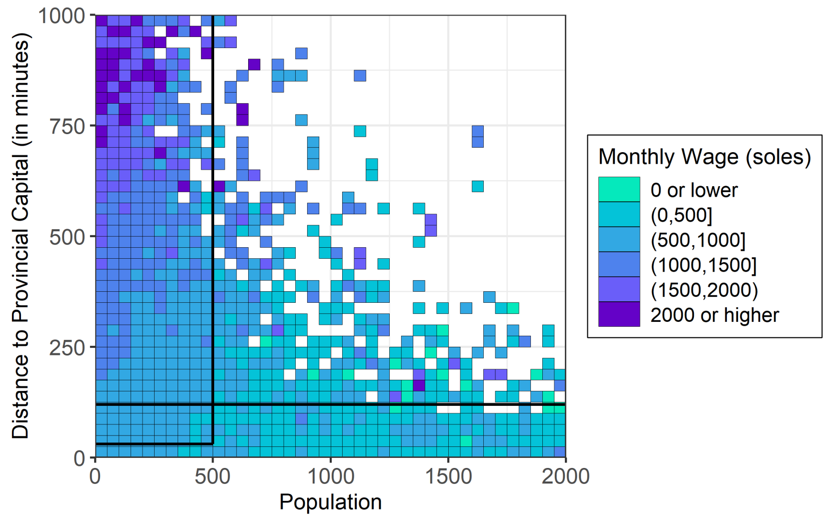
amount that ranges from two to four times the existing S/ 500 monthly wage bonus to fully compensate for the differences in non-pecuniary benefits with respect to schools in urban areas.

Overall, preference estimates suggest that teachers' labor supply is likely to be disproportionately concentrated in urban areas, even in the presence of the current wage bonus policy. As a result, the current teacher allocation is likely to remain highly unequal and favor urban schools. We find evidence supporting this hypothesis in Section 6.1, where we explore the aggregate effects of the rural wage bonus.

5.2 Teacher Value Added

We find that teachers are major inputs to student achievement production. A 1σ increase in teacher value added ($z_{lt}'\delta_i$) corresponds to a 0.50σ increase in students' math test scores and a 0.44σ increase in Spanish scores. This finding alone indicates that teacher sorting has the potential to largely reduce or amplify the urban-rural gap in student achievement.

Figure 6: Rural vs. Urban Non-Pecuniary Utility Differences



Notes. This figure plots the distribution of non-pecuniary utility differences between rural and urban schools measured in monthly wages. To construct this figure, we simulate the non-pecuniary part of utility $U_{ijt} - w_{ij}$ for each teacher-school pair using the estimated parameters from equation (2) and fixed draws of ϵ_{ijt} and θ_i. We construct the top decile of $U_{ijt} - w_{ij}$ for each school j and compute its difference for each rural school with the average urban school. This figure plots this difference averaged at the level of equally spaced cells of dimension 50×25 in the population-distance to provincial capital space.

Teachers also widely differ in their comparative advantage in teaching students with different characteristics. Table 5 shows the existence of significant match effects on lagged measures of student achievement. For example, a student with second-grade test scores 1σ below average can experience fourth-grade test score gains of 0.145σ by being matched with a teacher with similar average effectiveness but with a 1σ higher match effect. More generally, depending on the subject taught (math or Spanish), we find that 82-88% of the total variance in teacher value-added can be explained by differences in average effectiveness, while the remaining 12-18% is explained by differences in teachers' comparative advantage (see Panel A of Table 6).

The variance of value added explained by match effects can be mostly attributed to comparative advantage on students' lagged achievement (approx. 90%). This evidence is consistent with findings in Ahn et al. (2023); Graham et al. (2023), which document that students lagging behind have the largest potential gains from teachers' reallocation based on comparative advantage. It also points to large potential gains for rural areas from attracting better teachers along *both* absolute and comparative advantages. The estimates of the co-variance matrix of teacher value added ($\Sigma_{\delta,\delta}$) show that being an effective teacher on average is correlated with having a comparative advantage on teaching relatively older students, who tend to have lower grades (see Table D.5). This piece of evidence hints at the possibility

Table 5: Standard Deviation Value Added Coefficients

	Math (1)	Spanish (2)
ATE	0.465 (0.006)	0.408 (0.006)
Lagged Score	0.145 (0.005)	0.150 (0.005)
Lagged Score^2	0.049 (0.004)	0.061 (0.003)
Female	0.098 (0.010)	0.083 (0.013)
Quechua - Aymara	0.040 (0.030)	0.067 (0.019)
Age	0.115 (0.007)	0.110 (0.008)

Notes. This Table displays the estimates of the standard deviations of the teacher value-added coefficients from equation (6). Estimates of the full variance-covariance matrix $\Sigma_{\delta,\delta}$ can be found in Table D.2. Standard errors are in parentheses.

Table 6: Variance Decomposition of Teacher Value Added

	Math		Spanish	
	Variance	%	Variance	%
Panel A: Total Variance Explained by Match Effects				
ATE	0.2196	87.6	0.1606	81.1
Lagged Score + Age	0.0284	11.3	0.0347	17.5
Ethnicity + Gender	0.0027	1.1	0.0026	1.3
Total	0.2507	100	0.1979	100
Panel B: Variance Coefficients Explained by Teacher Characteristics				
ATE	0.0156	6.7	0.0143	7.9
Lagged Score	0.0008	3.7	0.0011	4.8
Lagged Score^2	0.0002	6.0	0.0002	4.1
Female	0.0001	0.6	0.0000	0.6
Quechua-Aymara	0.0002	9.5	0.0003	6.0
Age	0.0005	3.9	0.0003	2.6

Notes. Panel A of this Table decomposes the total variance of teacher value-added $z_l'\delta_i$ from equation (6) in the components related to match effects with specific sets of student covariates. Panel B displays the share of the total variance of each teacher value-added coefficient explained by observable teacher characteristics.

that reallocating teachers to raise efficiency by leveraging match effects might also improve equity. We quantify the magnitude of these potential equity and efficiency gains in Section 6.2.

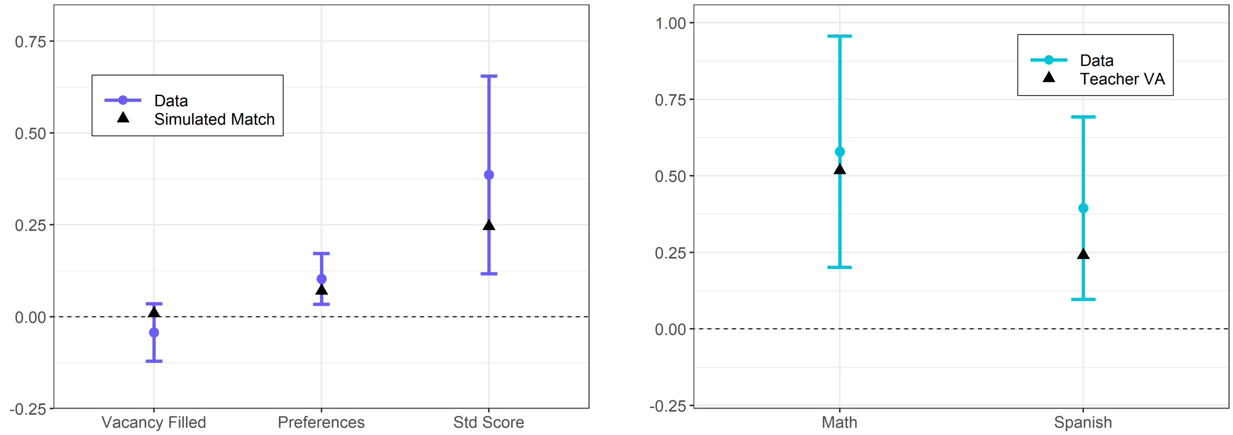
We also find that the coefficients of teacher value-added correlate with a range of observed teacher characteristics (see Table D.6). Teacher competency scores strongly correlate with teacher value-added. Increasing test scores results in logical reasoning or curricular knowledge by 1σ is associated with a 0.07σ increase in teacher value-added. Female teachers tend to have a higher average value-added than male teachers. The value-added of non-certified teachers is on average 0.15σ lower than for certified teachers. Surprisingly, we do not find strong evidence that teacher observables explain their potential comparative advantage. Female teachers do not seem to be more effective when teaching to female students. *Quechua* or *Aymara* teachers do not seem to be more effective when teaching to students from the same ethnicity.

Consistently with a large body of work on teacher effectiveness across both developed and developing countries (Rivkin et al., 2005; Araujo et al., 2016; Bau and Das, 2020), Panel B of Table 6 shows that observed measures of teaching quality only explain between 1% and 8% of the overall variance in the teacher value-added coefficients. Instead, we find that random coefficients governing teachers' preferences strongly correlate with the teacher value-added coefficients. The estimated covariance matrix between the random coefficients (α_i, β_i) and the value-added coefficients (δ_i) is displayed in Table D.7. Teachers who are less responsive to wage differences are more effective on average and have a higher comparative advantage with students with lower measures of prior achievement. Teachers with a higher outside option are more effective on average. We find that, if we would directly observe the random coefficients θ_i and condition on them, the standard deviation of teacher math value-added would shrink from 0.50σ to 0.22σ. This suggests that teachers' unobserved types driving their preferences over job postings likely reveal information about latent factors, such as their intrinsic motivation, that cannot be captured by observable characteristics. Combining data on teachers' school choices with data on student achievement could thus significantly help us get a better understanding of what makes a good teacher.

5.3 Model Fit

We now turn to assess whether the estimated model replicates the main features of the data through several exercises. We first test whether the model can replicate the threshold-crossing effects on teacher sorting discussed in Section 3. This provides a direct assessment of the validity of the estimated wage elasticities. We simulate the equilibrium teacher-school match predicted by the model using the teacher-proposing DA algorithm given the estimated preference parameters and a random draw of ϵ and θ to construct u_{ij} for all teacher-school

Figure 7: Model Validation



a. Preference Estimates

b. Teacher Value Added Estimates

NOTES. Panel A of this figure compares estimates of the threshold crossing effects displayed in Table 2 with the same estimates derived from a simulated teacher-school match. This simulated teacher-school match is constructed using the teacher-proposing DA algorithm given the estimated preference parameters from equation (2) and a random draw of ϵ and θ to construct U_{ijt} for all teacher-school pairs (i, j). Panel B of this figure shows the threshold-crossing effects on students' test scores in math and Spanish displayed in Column (1) of Table 3 along with threshold-crossing effects on shrunken teacher value-added estimates (see Appendix D.2 for details on constructing the posterior distribution).

pairs (i, j). We then replicate the RD analysis on this simulated match and compare the resulting estimates with the RD estimates obtained with the actual data (see Table 2). Panel A of Figure 7 shows that the estimated choice model predicts well the observed changes in teacher sorting patterns induced by the rural wage bonus.[14] However, it slightly under-predicts the threshold-crossing effects on teachers' competency score indicating that our estimates of the wage elasticity may be a lower bound.

The RD variation can be further leveraged to validate the student achievement production model (7). Panel B of Figure 7 shows the threshold-crossing effects, within the teacher-school match observed in the data, on students' test scores in math and Spanish displayed in Column (1) of Table 3 along with threshold-crossing effects on predicted teacher value-added corresponding to shrunken estimates of $z'_{lt}\delta_i$ (see Appendix D.3 for details on constructing the posterior distribution of δ_i). The predictions from the teacher-value added model match closely the threshold-crossing effects on standardized student test scores found in the data (see Table 3). As teacher effectiveness is the only input in the student achievement production

[14]Notice that, due to the nature of the matching equilibrium, an out-of-sample model validation—whereby the model would be estimated on the subset of the data with a Low Bonus, then simulated to predict the effect of the High Bonus, and finally assessed based on its performance in replicating the RD estimates in the data—is not feasible in our context.

function that varies non-smoothly at the population cutoff, this provides credibility to the value-added estimates discussed in the previous sub-section.

Finally, we test whether the predictions of the model match a wide range of moments of the distribution of matched characteristics and student achievement. Figure D.1 shows that the estimated model can replicate nearly exactly the urban-rural gaps in teacher competency scores (Panel A) and student achievement (Panel B). Similarly, Table D.8 shows that the model predicts well averages of matched teacher and school characteristics as well as the share of teachers choosing the two outside options (3)-(4). Overall, we find that the estimated model replicates quite closely a diverse set of features of our data.

6 Counterfactual Analysis

In this section, we first use the estimated model to decompose and quantify the aggregate effects of the rural wage bonus policy studied in Section 3. We then characterize the potential equity and efficiency gains from alternative teacher assignments. We finally build a framework for designing compensation policies aimed at attaining those gains at a minimal cost by leveraging information on teachers' preferences and effectiveness.

6.1 Aggregate Effects of the Rural Wage Bonus

We start by simulating the counterfactual teacher-school match and the implied distribution of student achievement resulting from removing the wage bonuses attributed to *Extremely Rural*, *Moderately Rural*, and *Rural* schools. We do this by predicting teachers' preferences over schools u_{ij} both in the absence and in the presence of the rural wage bonus for the overall sample of school vacancies from the estimated parameters and a random draw of ϵ_{ijt} and θ_i. We then compute the stable matching equilibria by running the DA algorithm. We finally predict the distribution of teacher value-added without and with rural wage bonuses by using the shrunken estimates of δ_i (i.e. the mean of the posterior distribution of δ_i).

Table 7 summarizes the results of this exercise by reporting the distribution of average teacher effectiveness in math under each scenario (see Table D.9 for the corresponding results on teacher effectiveness in Spanish). We find that despite the large local effects around the population threshold documented in the RD analysis of Section 3, the aggregate effects of the rural wage bonus are more modest. The policy decreased the overall urban-rural

Table 7: Evaluation of the Rural Wage Bonus - Math

	Status Quo (1)	No Rural Bonus (2)	Policy Effect (3)
Panel A: Total Value Added			
Urban-Rural Gap	0.077	0.164	-0.087
Urban	0.024	0.059	-0.036
Rural	-0.053	-0.105	0.052
Moderately Rural	-0.033	-0.055	0.022
Rural	-0.111	-0.049	-0.063
Extremely Rural	0.067	-0.099	0.166
Panel B: Match Effects			
Urban	-0.007	0.002	-0.009
Rural	0.008	0.001	0.007

Notes. Column (1) and Column (2) summarize the spatial distribution of teacher value added in the simulated teacher-school match in the status quo and in the absence of the rural wage bonus policy, respectively. Each simulated teacher-school match is constructed using the teacher-proposing DA algorithm given the estimated preference parameters from equation (2) and a random draw of ϵ and θ to construct U_{ijt} for all teacher-school pairs (i, j). Teacher value-added corresponds to shrunken estimates of $z'_{lt}\delta_i$ from equation (6)—see Appendix D.3 for details on how to construct the posterior distribution of δ_i. Column (3) is the difference between Column (1) and (2).

gap in student achievement by 0.09σ in math and 0.07σ in Spanish. Most of the effect is concentrated in *Extremely Rural* locations, which receive the largest bonus. Other areas either lose or benefit very little. Teacher sorting remains highly unequal under the status quo allocation (Column 1 of Table 7) to the extent that average teacher value-added in rural areas is still 0.08σ lower than in urban areas for math and 0.11σ for Spanish.

It is worth noticing that the increase in student achievement in rural areas triggered by the policy does not translate into an equivalent loss for public schools in urban areas. Teacher value-added increases by 0.05σ in rural areas while decreasing by only 0.04σ in urban areas. We find that these net gains are in a large part fueled by higher value-added teachers substituting away from the outside option (see Table D.10). Note that these teachers might be substituting away from urban private schools implying that we may be underestimating the negative impact on urban areas.

Finally, we find that the wage bonus policy fails to make teachers sort based on their comparative advantages (see Panel B of Table 7). These results suggest the existence of large potential equity and efficiency gains from counterfactual teacher assignments, which we characterize in the next sub-section.

6.2 Equity and Efficiency Gains from Reallocation

To characterize the potential gains from teacher reallocation across schools we consider the following problem:

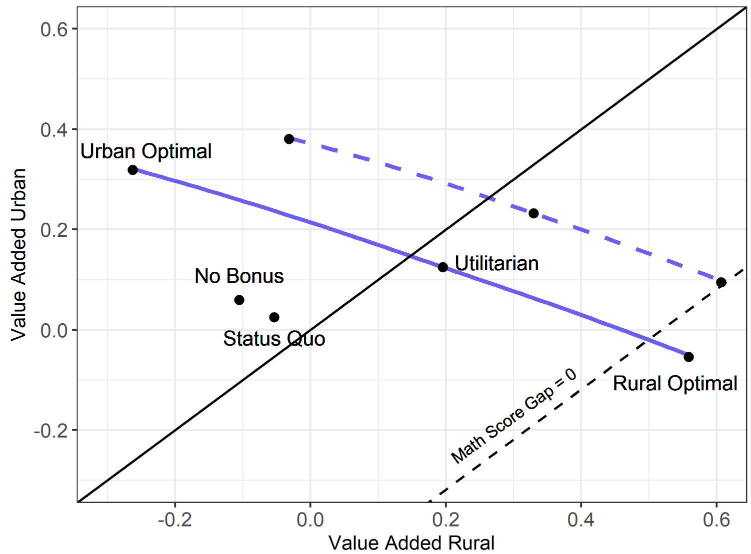
$$\max_{\mu} \sum_{i=1}^{n_w} \sum_{j=1}^{n_m} \pi_j \bar{z}_j' \delta_i \mathbb{1}\{\mu(i) = j\} \tag{16}$$

where $\bar{z}_j'\delta_i$ is the average teacher value-added in school j when assigned teacher i, and π_j determines the weight put on students in school j. We aim to find the solution $\mu^*(\pi)$ of this linear program for a given set of weights π. We plug in the shrunken estimates of δ_i to approximate the objective function. We fix $\pi_j = 1$ for all schools j located in urban areas. We then specify $\pi_k = x$ for all schools k located in rural areas. By varying x we can then characterize a set of counterfactual teacher-school allocations that maximize total output while giving more or less priority to rural areas. Putting all weight on rural schools gives us the *Rural Optimal* allocation. Conversely, putting zero weight on rural schools gives us the *Urban Optimal* allocation. We label the allocation reached with $x = 1$ as the *Utilitarian* allocation.

Figure 8 displays the results of this exercise. The solid purple line represents the solution to the linear program (16) for different values of x. The dashed purple line is the result of the same exercise when augmenting the pool of teachers to draw from with teachers that chose the outside option in the status quo match. Therefore, the difference between the solid and dashed lines quantifies the gains stemming from the extensive margin, as we draw teachers from a larger set of applicants. Increasing the pool of teachers by considering applicants that chose the outside option further pushes upward the efficiency frontier by 0.2σ suggesting large potential gains from drawing more applicants into the public sector.

We find that there are large attainable efficiency *and* equity gains, both with respect to the teacher allocation that would occur in the absence of the rural wage bonus as well as the status quo teacher allocation (No Bonus and Status Quo points in Figure 8). Leveraging match effects by reallocating the pool of existing teachers could increase student outcomes in rural areas by up to 0.45σ while keeping student outcomes in urban areas unchanged. Reaching the Rural Optimal allocation would increase outcomes in rural areas by 0.61σ at the cost of decreasing student achievement in urban areas by 0.08σ. Teacher sorting can largely exacerbate or close the urban-rural gap in student outcomes. The Urban Optimal

Figure 8: Potential Efficiency and Equity Gains



Notes. This figure displays the solution to the problem described in Equation (16) for different values of x ranging from 0.1 (Urban Optimal) to 10 (Rural Optimal). The solid purple line corresponds to the case where the pool of available teachers is restricted to the ones already assigned in the status quo. The dashed purple line corresponds to the case where the pool of available teachers includes those who chose the outside option in the status quo as well. The dashed black line corresponds to all points such that the urban-rural gap in students' math test scores is exactly equal to 0. All points above correspond to teacher-school matches where student achievement is higher in urban areas than in rural areas.

allocation would entail a 1.08σ urban-rural gap in student achievement, while the Rural Optimal allocation would tilt the balance in favor of rural schools creating an urban-rural output gap of -0.11σ. Lastly, we find that both the Urban Optimal and Rural Optimal allocations would lead to the same total output gain as the Utilitarian allocation, suggesting that striving for equity does not harm efficiency.

Finally, we assess the validity of this exercise by comparing the realized efficiency gains when solving for the Utilitarian optimal allocation using the true teacher value-added coefficients δ_i (Oracle), the mean of its estimated posterior distribution (Shrunken), or the unshrunken OLS estimates (Naive) through Monte Carlo simulations (see Appendix D.4 for additional details). Table 8 shows that the decision rule based on the shrunken estimates performs well and achieves 74% of the total efficiency gains that could be achieved if we knew the true value of δ_i. In contrast, the OLS estimates perform poorly and only achieve 38% of these efficiency gains. The OLS estimates perform even worse in rural areas where the potential efficiency gains are larger. This highlights the benefits of using Bayesian shrinkage in assignment problems with treatment effect heterogeneity. As noisy estimates are shrunk towards the mean, match effects coefficients shrink to zero, attenuating teachers' comparative

Table 8: Monte Carlo Simulations - Efficiency Gains

	Oracle (1)	Shrunken (2)	Naive (3)
Total Output Gain	0.135	0.099 (74%)	0.051 (38%)
Urban	0.103	0.074 (72%)	0.047 (46%)
Rural	0.224	0.170 (76%)	0.061 (27%)

NOTES. This table displays the results of Monte Carlo simulations assessing the efficiency gains realized when solving problem (16) using different estimators for δ_i from equation (6)—see Appendix D.4 for additional details. Column (1) corresponds to the case where we use the true value-added coefficients δ_i. Column (2) and (3) correspond to the cases where we use instead the shrunken estimates or the OLS estimates of δ_i, respectively. Efficiency gains are measured in math test scores gains with respect to the status quo allocation.

advantage and largely reducing the potential for efficiency gains. This allows to put priority on matches that are high quality with high certainty when solving Equation (16) and thus limits the potential for making costly mistakes that would harm efficiency.

6.3 Teacher Compensation Design

The results discussed in the previous section document that naive teacher compensation policies, such as the rural wage bonus, fail to achieve large potential efficiency and equity gains. In this section, we study the extent to which those gains can be achieved by developing a framework for designing cost-effective teacher compensation policies that fully leverage information on teachers' preferences and effectiveness.

6.3.1 Framework

We consider the point of view of a policymaker who has two instruments available: (i) setting the priorities used to rank teachers in the centralized assignment mechanism, and (ii) setting the wages offered in each school, w_j. The objective of the policymaker is to use these two instruments such that each school $j \in \mathcal{S}$ is assigned at least one teacher with value-added above a predefined threshold c_j under the match resulting from the centralized assignment mechanism. We formalize the problem as follows:

$$\min_{w} \sum_{j} w_j, \text{ s.t.} \begin{cases} \max_{i \in \mu(j)} z_{lt}' \delta_i \geq c_j, \ \forall j \in \mathcal{S} & \text{(C1)} \\ \mu \text{ is stable given } w \text{ and using } z_{lt}'\delta_i \text{ as priorities} & \text{(C2)} \end{cases} \tag{17}$$

Condition (C1) imposes that the policy maker's objective is attained under the resulting teacher-school match μ. Condition (C2) makes sure that the solution is implementable by imposing that μ would result from the current assignment mechanism if we set wages to w and priorities to $z'_{lt}\delta_i$. To be consistent with the institutional setting, we do not allow for teachers working in the same school to be paid differently.

We show that the solution of this problem is the decentralized equilibrium outcome of a counterfactual economy where schools are willing to bid for teachers by increasing wages until (C1) is satisfied. This environment is called the matching with contracts framework (Kelso and Crawford, 1982; Hatfield and Milgrom, 2005) and augments the definition of a match as being a teacher-school-wage contract. We show that if we specify schools' preferences over contracts as follows:

(P1) For a fixed wage, schools follow strict priorities and rank teachers according to $z'_{lt}\delta_i$.

(P2) Any contract satisfying (C1) is strictly preferred to a contract not satisfying (C1) for any given wage. Within contracts satisfying (C1), or within contracts not satisfying (C1), the allocation with the lower wage is always strictly preferred.

and specify teachers' preferences over contracts according to the model described in Equation (2), we can establish the following result.

Proposition 1 *Under (P1)-(P2), the school-optimal stable set of contracts is the solution to Equation* (17)*.*

See Appendix E.2 for a proof. This result stems from Theorem 3 in Hatfield and Milgrom (2005) showing that a stable set of contracts always exists in this counterfactual economy. More importantly, there exists a school-optimal stable allocation, which is unanimously preferred by schools, and a teacher-optimal stable allocation, which is unanimously preferred by teachers. Intuitively, the proof shows that, since (P2) implies that schools are willing to increase wages until (C1) is satisfied, stability implies that (C1) is satisfied. However, (P2) also implies that schools strictly prefer allocations with lower wages meaning that the school-optimal stable allocation will satisfy (C1) while minimizing wages. (P1) implies that for fixed wages the allocation is stable with respect to school priorities which satisfies the implementability constraint (C2). Note that for this result to hold, the preference ordering described by (P1)-(P2) needs to imply that contracts are substitutes (see Appendix E.1 for a formal definition and proof).

The school-optimal stable set of contracts can be reached through the school-proposing generalized deferred acceptance algorithm—see Theorem 3 in Hatfield and Milgrom (2005). Under (P1) and (P2) the school-proposing generalized DA algorithm essentially entails schools bidding for teachers by increasing wages until (C1) is satisfied. We set the step size for wage increases to 50 soles (2.5% of base wage).

6.3.2 Results

We leverage Proposition 1 to study whether the wage setting protocol outlined above can achieve the large equity and efficiency gains documented in Section 6.2. Teachers' preferences over schools are constructed from the estimated preference parameters and a fixed random draw of ϵ and θ to construct U_{ijt} for all teacher-school pairs (i, j). Teacher value-added corresponds to shrunken estimates of $z_{lt}'\delta_i$ (see Appendix D.3 for details on how to construct the posterior distribution of δ_i).[15] We consider different thresholds for value-added by setting $c \in \{-0.4, -0.3, -0.2, -0.1, 0\}$ in each rural school. Table 9 shows the result of this exercise. We find that large gains in teacher value-added are reachable in rural areas at a moderate cost. Imposing $c = -0.3$ allows to fully close the urban-rural gap in teacher value-added and entails an average wage increase of 14% which is only slightly higher than the cost of the current policy. Most of the gains achieved through the optimal wage policy are accrued by inducing teachers to match on their comparative advantage, as shown in Panel B of Table 9. This shows that teacher compensation policies incorporating knowledge about teachers' preferences and effectiveness in their design can achieve substantial equity and efficiency gains at a moderate cost. Columns (4)-(6) show that achieving further gains for rural areas is attainable but at an increasingly larger cost.

In practice, it might be unfeasible to change schools' priorities such that they reward teacher value-added instead of other observed measures of teacher quality. We investigate what would be the distribution of teacher value-added resulting from the wage schedules derived for Table 9 when fixing schools' priorities to be the same as in the status quo (i.e. competency scores). Table E.1 shows that the counterfactual wage policies would still achieve larger equity gains than the status quo policy for the same cost. However, they achieve significantly lower efficiency gains, as schools do not give priority to teachers with a higher

[15] For teachers that are not in our teacher-classroom data (see Section 2.2), we predict their value-added from the mean of the prior conditional on their observables and a draw of the random coefficients.

Table 9: Counterfactual Teacher Compensation Policies

	Status Quo	Teacher Value Added Threshold				
		$c = -0.4$	$c = -0.3$	$c = -0.2$	$c = -0.1$	$c = 0$
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Teacher Value Added						
Urban	0.055	0.036	0.035	0.019	-0.009	-0.058
Rural	-0.048	0.015	0.076	0.133	0.197	0.258
Moderately Rural	0.025	0.007	0.058	0.040	0.127	0.203
Rural	-0.154	-0.060	0.034	0.094	0.117	0.199
Extremely Rural	-0.022	0.080	0.131	0.225	0.296	0.357
Panel B: Match Effects						
Urban	0.019	0.017	0.018	0.018	0.013	0.022
Rural	0.040	0.063	0.111	0.137	0.180	0.191
Moderately Rural	0.008	0.002	0.031	0.022	0.065	0.089
Rural	0.039	0.085	0.141	0.107	0.154	0.161
Extremely Rural	0.070	0.106	0.168	0.218	0.247	0.300
Panel C: Monthly Total Cost (in Soles)						
% Base Wage	0.111	0.086	0.140	0.234	0.379	0.621
Mean Bonus per School	223	171	279	467	759	1,242
SD Bonus per School	220	407	576	839	1,184	1,698

NOTES. This table summarizes the spatial distribution of teacher value added (Panel A) and its match effect component (Panel B) under the simulated equilibrium teacher-school match resulting from different wage bonus policies. Column (1) corresponds to the absence of the rural wage bonus, Columns (2) to (6) correspond to the wage policy solving Equation (17) when condition (C1) is to have at least one teacher with value-added above c in every rural school where c takes different values. Teachers' preferences over schools are constructed from the estimated preference parameters and a fixed random draw of ϵ and θ to construct U_{ijt} for all teacher-school pairs (i, j). Teacher value-added corresponds to shrunken estimates of $z_{lt}'\delta_i$ (see Appendix D.3 for details on how to construct the posterior distribution of δ_i). Panel C summarizes the cost of each policy.

comparative advantage with their students.

Finally, we consider an alternative objective that requires no information on teacher value-added: filling at least one vacancy per school with a certified teacher—i.e., a teacher with a competency score. If a vacancy remains unfilled after the matching algorithm, schools usually recruit teachers without credentials who are substantially less effective. As we showed in Section 5.2, the value added of non-certified teachers is 0.15σ lower than for certified teachers. Column (3) of Table E.2 shows that this objective can be reached at a lower cost in terms of the total wage bill (almost half) than the current rural wage policy. However, the resulting allocation implies slightly lower value-added gains in rural areas than the status quo.

7 Conclusion

This paper establishes that the design of teacher compensation schemes can largely exacerbate or alleviate structural inequalities in schooling outcomes. We assemble rich administrative panel data on the universe of applicants and positions posted within a nationwide recruitment drive for public teachers. This unique dataset allows us to comprehensively analyze teachers' choices over jobs and study how these map into student outcomes. We document that wage rigidity induces teachers to sort on non-pecuniary aspects of employment, resulting in school choices that are skewed towards urban areas. This leads to large spatial disparities in teacher quality that are strongly associated with inequities in other schooling inputs and student achievement. We then leverage a policy reform that largely increased teacher compensation in remote and rural schools to study its impact on teacher sorting across locations and student outcomes. We find that the policy successfully attracted higher quality teachers in the targeted remote areas, which substantially improved student learning.

To go beyond the local estimated effects of the policy and understand the potential equity and efficiency gains that can be achieved through alternative teacher compensation schemes, we build and estimate a model of teacher sorting across schools and student achievement. Teachers have heterogeneous preferences over wage and non-wage attributes that induce vertical and horizontal differentiation across jobs. Teacher sorting maps into student achievement through a teacher value-added model where teacher effectiveness is allowed to be heterogeneous according to a wide range of student characteristics. We then use the estimated model to show the presence of large equity and efficiency gains from teacher reallocation. Importantly, we find that the current wage bonus policy does not realize those gains as it fails to incentivize teachers to sort on their comparative advantage. In fact, despite the current wage bonuses in place, teacher sorting across locations remains highly unequal and accounts for one-quarter of the observed urban-rural gaps in student achievement. We then build a framework for designing cost-effective compensation policies that incorporate information on teachers' preferences and effectiveness. We find that such information is highly valuable as inequalities in teacher effectiveness could be alleviated without negatively affecting urban areas and at a slightly higher cost than the status-quo policy of wage bonuses.

Stretching beyond the specific setting of our analysis, our findings suggest that incorporating measures of preferences and productivity in the design of workers' compensation is a promising alternative to rigid wage schedules and market-based wage setting in a variety of other labor markets. Aligning the private and social returns of worker-firm matches through informed compensation design can result in more equitable and efficient allocations. We believe that this approach is increasingly relevant from a policy perspective, given the widespread availability of administrative data on centralized labor markets as well as recent developments in the tools that enable researchers and practitioners to leverage such data to infer the preferences of participating agents (Roth, 2018; Agarwal and Budish, 2021).

References

Abdulkadiroğlu, Atila, Parag A Pathak, Jonathan Schellenberg, and Christopher R Walters, "Do Parents Value School Effectiveness?," *American Economic Review*, 2020, *110* (5), 1502–1539.

Agarwal, Nikhil, "Policy Analysis in Matching Markets," *American Economic Association P&P*, 2017, *107* (5), 246–50.

_ **and Eric Budish**, "Market Design," in "Handbook of Industrial Organization," Vol. 5, Elsevier, 2021, pp. 1–79.

Ahn, Tom, Esteban Aucejo, and Jonathan James, "The Importance of Matching Effects for Labor Productivity: Evidence from Teacher-Student Interactions," Technical Report, California Polytechnic State University, Department of Economics 2023.

Angrist, Joshua D., Peter D. Hull, Parag A. Pathak, and Christopher R. Walters, "Leveraging Lotteries for School Value-Added: Testing and Estimation," *The Quarterly Journal of Economics*, 2017, *132* (2), 871–919.

Araujo, M Caridad, Pedro Carneiro, Yyannú Cruz-Aguayo, and Norbert Schady, "Teacher quality and learning outcomes in kindergarten," *The Quarterly Journal of Economics*, 2016, *131* (3), 1415–1453.

Banerjee, Abhijit and Lakshmi Iyer, "History, institutions, and economic performance: The legacy of colonial land tenure systems in India," *American economic review*, 2005, *95* (4), 1190–1213.

Barrera-Osorio, Felipe and Dhushyanth Raju, "Teacher performance pay: Experimental evidence from Pakistan," *Journal of Public Economics*, 2017, *148* (C), 75–91.

Bates, Michael D, Michael Dinerstein, Andrew C Johnston, and Isaac Sorkin, "Teacher Labor Market Equilibrium and Student Achievement," *NBER Working Paper 29728*, 2022.

Bau, Natalie and Jishnu Das, "Teacher value added in a low-income country," *American Economic Journal: Economic Policy*, 2020, *12* (1), 62–96.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, 2004, *112* (1), 68–105.

Bertoni, Eleonora, Gregory Elacqua, Diana Hincapie, Carolina Mendez, and Diana Paredes, "Teachers' Preferences for Proximity and the Implications for Staffing Schools: Evidence from Peru," IDB Publications (Working Papers) 01073, Inter-American Development Bank October 2019.

_ , _ , Luana Marotta, Matías Martinez, Humberto Santos, and Sammara Soares, "Is School Funding Unequal in Latin America? A Cross-Country Analysis. CEPA Working Paper No. 20-11.," *Stanford Center for Education Policy Analysis*, 2020.

Biasi, Barbara, "The Labor Market for Teachers under Different Pay Schemes," *American Economic Journal: Economic Policy*, August 2021, *13* (3), 63–102.

_ , Chao Fu, and John Stromme, "Equilibrium in the Market for Public School Teachers: District Wage Strategies and Teacher Comparative Advantage," *NBER Working Paper 28530*, 2021.

Bonhomme, Stéphane, Grégory Jolivet, and Edwin Leuven, "School characteristics and teacher turnover: Assessing the role of preferences and opportunities," *The Economic Journal*, 2016, *126* (594), 1342–1371.

Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff, "Analyzing the Determinants of the Matching of Public School Teachers to Jobs: Disentangling the Preferences of Teachers and Employers," *Journal of Labor Economics*, 2013, *31* (1), 83–117.

Brown, Christina and Tahir Andrabi, "Inducing positive sorting through performance pay: Experimental evidence from Pakistani schools," *University of California at Berkeley Working Paper*, 2020.

Cabrera, Jose Maria and Dinand Webbink, "Do Higher Salaries Yield Better Teachers and Better Student Outcomes?," *Journal of Human Resources*, 2020, *55* (4), 1194–1221.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression?Discontinuity Designs," *Econometrica*, November 2014, *82*, 2295–2326.

— , — , and — , "Optimal Data-Driven Regression Discontinuity Plots," *Journal of the American Statistical Association*, December 2015, *110* (512), 1753–1769.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma, "Simple Local Polynomial Density Estimators," *Journal of the American Statistical Association*, 2020, *115* (531), 1449–1455.

Chetty, Raj, John Friedman, and Jonah Rockoff, "Measuring the Impact of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood," *American Economic Review*, 2014, *104(9)*, 2633–2679.

— , — , and — , "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates," *American Economic Review*, 2014, *104(9)*, 2593–2632.

Clotfelter, Charles, Elizabeth Glennie, Helen Ladd, and Jacob Vigdor, "Would Higher Salaries Keep Teachers in High-Poverty Schools? Evidence from a Policy Intervention in North Carolina," *Journal of Public Economics*, 2008, *92* (5), 1352–1370.

Clotfelter, Charles T, Helen F Ladd, and Jacob Vigdor, "Who teaches whom? Race and the distribution of novice teachers," *Economics of Education review*, 2005, *24* (4), 377–392.

de Ree, Joppe, Karthik Muralidharan, Menno Pradhan, and Halsey Rogers, "Double for Nothing? Experimental Evidence on an Unconditional Teacher Salary Increase in Indonesia," *The Quarterly Journal of Economics*, 2018, *133* (2), 993–1039.

Dell, Melissa, "The persistent effects of Peru's mining mita," *Econometrica*, 2010, *78* (6), 1863–1903.

Ederer, Tim, "Labor Market Dynamics and Teacher Spatial Sorting," *Job Market Paper*, 2023.

Estrada, Ricardo, "Rules versus Discretion in Public Service: Teacher Hiring in Mexico," *Journal of Labor Economics*, 2019, *37* (2), 545–579.

Fack, Gabrielle, Julien Grenet, and Yinghua He, "Beyond Truth-Telling: Preference Estimation with Centralized School Choice and College Admissions," *American Economic Review*, 2019, *109* (4), 1486–1529.

Finan, Frederico, Benjamin A. Olken, and Rohini Pande, "The Personnel Economics of the State," in Abhijit Banerjee and Esther Duflo, eds., *Handbook of Field Experiments, Volume II*, North Holland, December 2017.

Fox, Jeremy T, Kyoo il Kim, Stephen P Ryan, and Patrick Bajari, "The random coefficients logit model is identified," *Journal of Econometrics*, 2012, *166* (2), 204–212.

Gerard, François, Miikka Rokkanen, and Christoph Rothe, "Bounds on treatment effects in regression discontinuity designs with a manipulated running variable," *Quantitative Economics*, 2020, *11* (3), 839–870.

Gershenson, Seth, Cassandra MD Hart, Joshua Hyman, Constance A Lindsay, and Nicholas W Papageorge, "The long-run impacts of same-race teachers," *American Economic Journal: Economic Policy*, 2022, *14* (4), 300–342.

Gilligan, Daniel O., Naureen Karachiwalla, Ibrahim Kasirye, Adrienne M. Lucas, and Derek Neal, "Educator Incentives and Educational Triage in Rural Primary Schools," *Journal of Human Resources*, 2022, *57* (1), 79–111.

Graham, Bryan S, Geert Ridder, Petra Thiemann, and Gema Zamarro, "Teacher-to-classroom assignment and student achievement," *Journal of Business & Economic Statistics*, 2023, *41* (4), 1328–1340.

Gu, Jiaying and Roger Koenker, "Invidious comparisons: Ranking and selection as compound decisions," *Econometrica*, 2023, *91* (1), 1–41.

Hatfield, John William and Paul R. Milgrom, "Matching with Contracts," *American Economic Review*, September 2005, *95* (4), 913–935.

INEI, "Encuesta Nacional de Hogares (ENAHO)," https://www.ilo.org/surveyLib/index.php/catalog/6803/study-description Instituto Nacional de Estadística e Informática 2016.

Jackson, C. Kirabo, "What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes," *Journal of Political Economy*, 2018, *126* (5), 2072–2107.

Jochmans, Koen and Martin Weidner, "Fixed-effect regressions on network data," *Econometrica*, 2019, *87* (5), 1543–1560.

Kelso, Alexander S. and Vincent P. Crawford, "Job Matching, Coalition Formation, and Gross Substitutes," *Econometrica*, 1982, *50* (6), 1483–1504.

Khan, Adnan Q., Asim Ijaz Khwaja, and Benjamin A. Olken, "Making Moves Matter: Experimental Evidence on Incentivizing Bureaucrats through Performance-Based Postings," *American Economic Review*, January 2019, *109* (1), 237–270.

Kline, Patrick, Evan K Rose, and Christopher R Walters, "A Discrimination Report Card," *arXiv preprint arXiv:2306.13005*, 2023.

Leaver, Clare, Owen Ozier, Pieter Serneels, and Andrew Zeitlin, "Recruitment, Effort, and Retention Effects of Performance Contracts for Civil Servants: Experimental Evidence from Rwandan Primary Schools," *American Economic Review*, July 2021, *111* (7), 2213–46.

Mansfield, Richard K, "Teacher quality and student inequality," *Journal of Labor Economics*, 2015, *33* (3), 751–788.

MINEDU, "Encuesta Nacional a Docentes de Instituciones Educativas Estatales y No Estatales (ENDO))," https://hdl.handle.net/20.500.12799/4872 Ministerio de Educación 2014.

Muralidharan, Karthik and Venkatesh Sundararaman, "Teacher Performance Pay: Experimental Evidence from India," *Journal of Political Economy*, 2011, *119* (1), 39–77.

Pugatch, Todd and Elizabeth Schroeder, "Teacher pay and student performance: evidence from the Gambian hardship allowance," *Journal of Development Effectiveness*, April 2018, *10* (2), 249–276.

Rivkin, Steven G, Eric A Hanushek, and John F Kain, "Teachers, Schools, and Academic Achievement," *Econometrica*, 2005, *73* (2), 417–458.

Rosen, Sherwin, "The theory of equalizing differences," *Handbook of labor economics*, 1986, *1*, 641–692.

Roth, Alvin E, "Marketplaces, Markets, and Market Design," *American Economic Review*, 2018, *108* (7), 1609–58.

— **and Marilda Sotomayor**, "Two-Sided Matching," *Handbook of Game Theory with Economic Applications*, 1992, *1*, 485–541.

Rothstein, Jesse, "Teacher quality policy when supply matters," *American Economic Review*, 2015, *105* (1), 100–130.

Sokoloff, Kenneth L and Stanley L Engerman, "History lessons: institutions, factor endowments, and paths of development in the new world," *Journal of Economic perspectives*, 2000, *14* (3), 217–232.

Tincani, Michela M, "Teacher Labor Markets, School Vouchers, and Student Cognitive Achievement: Evidence from Chile," *Quantitative Economics*, 2021, *12* (1), 173–216.

Walters, Christopher R, "The demand for effective charter schools," *Journal of Political Economy*, 2018, *126* (6), 2179–2223.

World Bank, "World Development Report 2018 : Learning to Realize Education's Promise," Technical Report, Washington, DC: World Bank 2018.

Appendices

This supplementary section contains five appendices. Appendix A describes the assignment mechanism for permanent and temporary vacancies. Appendix B contains additional descriptive evidence and summary statistics on the main sample of applicants that we use throughout the analysis, as well as on the on-line survey that we have collected during the 2016 recruitment drive. Appendix C presents a series of additional results and robustness checks related to the RD analysis. Appendix D provides further details on the identification and estimation of the model of teacher sorting and student achievement, as well as the full set of estimated parameters of the sorting model, some moments and statistics assessing the overall fit of the model, and additional results on the evaluation of the rural wage bonus. Finally, Appendix E contains the proofs of the substitute condition and of the main proposition outlined in the matching framework of Section 6.3.1, as well as additional results on the optimal wage policy.

A Institutional Details

A.1 The Assignment Mechanism for Permanent Vacancies

Every permanent position across all education levels is posted on a centralized platform. The opening of each position depends on previous retirements and transfers, as well as the ability of local governments to secure permanent funding for the position. Applicants are required to have a teaching accreditation (i.e., a teaching degree). They must also correctly answer at least 60 percent of the questions in each of the three parts (reading comprehension, logic reasoning, curricular knowledge) of the national competency evaluation.

Eligible applicants can indicate their preferred school district and submit a rank order list of schools within that district. Once preferences are submitted, teachers move on to a decentralized stage of evaluation and enter a shortlist for their top three choices (top two in 2016). This shortlist has a maximum length of 10 (20 in 2016). For schools that are oversubscribed, test scores are used to prioritize candidates. In this second evaluation round, teachers are given another score based on a direct evaluation of their performance in teaching a typical class and an in-person interview with the principal and other school stakeholders.

Points can also be assigned based on their CV. Finally, schools make offers sequentially to the applicants ranked according to the overall score that comprises the competency test and the decentralized evaluation. Unassigned applicants can then participate in an exceptional stage that allocates the remaining unfilled slots. At the end of this round, unassigned teachers can decide to participate in the allocation of temporary positions, which takes place shortly after.

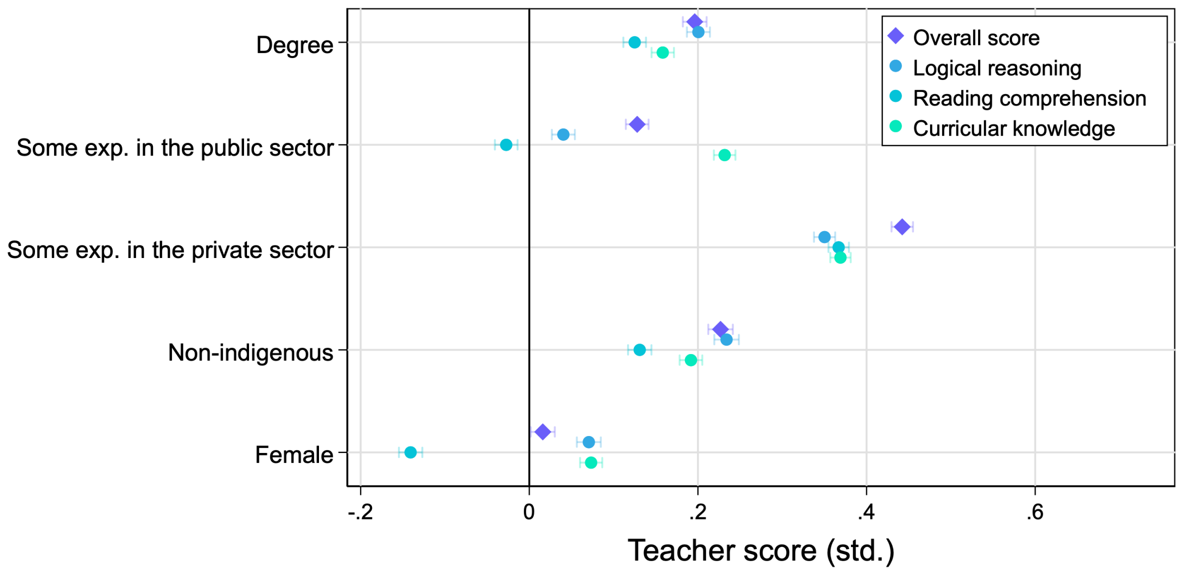
A.2 The Assignment Mechanism for Temporary Vacancies

Contract teacher positions are posted on the website of each school district. The list of vacancies includes both standard positions – that are based on a two-year contract, with the possibility of renewal for a second year upon approval from the school's director – and occasional positions catering to short-term extra needs, such as covering maternity leaves and lasting up to one year. The list of contract teacher vacancies also includes positions that are not filled through the assignment mechanism for permanent teachers, and are later posted as contract teacher vacancies.

Participants in the assignment mechanism for contract teachers are asked to indicate a preferred school district when applying. School districts are administrative units corresponding roughly to Peruvian provinces. As of 2016, there are 226 school districts in Peru. Vacancies are assigned based on a serial dictatorship mechanism. All applicants to the assignment mechanism in a given school district and specialization are ranked based on the score they got on the national competency test with bonus points awarded to those with recognized disabilities or who served in the Peruvian army. The assignment procedure works as follows: the highest-scoring teacher chooses their preferred position, which is thus removed from the list and thus is not available to the subsequent lower-ranked applicant. This procedure continues until all positions in the list are filled or the lowest-ranked applicant makes her choice. Vacancies that remain unfilled are made available to other groups of (unmatched) applicants who initially indicated a different school district or specialization. Specifically, they are first made available to those who initially indicated a different school district within the same region. Second, if vacancies still remain, they are made available to applicants who initially indicated a different region or stage/subject specialization. Any positions not filled through this procedure are then offered to non-certified teachers – who did not participate in the competency test – based on a committee evaluation of their curricula.

B Descriptive Evidence

Figure B.1: Teacher Characteristics and Standardized Competency Scores



Notes: This figure depicts OLS estimates along with their associated 95 percent confidence intervals from a multivariate regression analysis of various individual teacher characteristics on teacher competency scores. These characteristics include an indicator for university (vs technical institute) education, two indicators for experience (at least one year) in the public and private sector, a dummy variable equal to 1 if the teacher does not speak a Peruvian indigenous language, and a female dummy variable. The sample includes all applicants to the assignment mechanism for contract teachers, irrespective of whether they were eventually assigned, in both the 2016 and 2018 rounds.

Table B.1: Sample Selection—Survey Respondents

	All applicants (2016) (1)	Survey respondents (2)
Female	0.725	0.723
	(0.447)	(0.448)
Age	34.546	35.201
	(6.184)	(6.546)
Indigenous language:	0.122	0.093
	(0.328)	(0.291)
- Quechua	0.109	0.084
	(0.312)	(0.277)
- Aimara	0.016	0.012
	(0.127)	(0.107)
Some experience in the public sector	0.812	0.801
	(0.390)	(0.399)
Experience in the public sector (≥3 yrs)	0.621	0.614
	(0.485)	(0.487)
Some experience in the private sector	0.674	0.709
	(0.469)	(0.454)
Experience in the private sector (≥3 yrs)	0.447	0.475
	(0.497)	(0.499)
Overall competecy score (std)	1.593	1.608
	(0.390)	(0.457)
Curricular knowledge (std)	1.158	1.174
	(0.472)	(0.492)
N. of teachers	22,784	5,550

Notes. This table reports the summary statistics for the sample of survey respondents (Column 2) vis-à-vis the sample of applicants to the 2016 assignment system for permanent teaching positions (Column 1). Columns report, for each group, the mean, and the standard deviation for the variables considered.

Table B.2: Applicant Survey (Strategy and Information)

	All	Score in Top Quartile
Panel A: Strategic behavior (% of respondents)		
Preferred school in concurso	63.36	61.37
If preferred school in concurso, which rank?		
Ranked 1^{st}	84.26	88.93
Ranked 2^{nd}	6.28	3.51
Ranked 3^{rd}	2.31	1.32
Ranked 4^{th}	0.71	0.66
Ranked 5^{th}	0.95	0.66
Not Ranked	5.48	4.93
If not ranked first, why?		
High demand and score too low	64.91	41.82
Remuneration not attractive	3.51	5.45
Other	31.58	52.73
Panel B: Information about first choice (% of respondents)		
Had prior information about first choice	50.97	54.01
Does your first choice benefit from wage bonus?		
Yes	16.42	15.08
No	54.53	62.69
Do not know	29.04	22.23
Expected wage - actual wage (in %)	-9.48	-7.87

Notes. This table displays the answers of the 5,553 survey respondents to the corresponding questions. The last columns displays the same results for respondents that scored above the top quartile of the test score distribution for tenured teachers.

Table B.3: Applicant Survey (Participation and Choice Attributes)

	All Teachers Rank				Score in Top Quartile Rank			
	1^{st}	2^{nd}	3^{rd}	In Top 3	1^{st}	2^{nd}	3^{rd}	In Top 3
Panel A: Why did you apply to the centralized assignment mechanism? (% of respondents)								
Career	33.77	30.35	20.57	84.69	33.73	29.97	21.35	85.05
Stability	51.08	17.04	14.76	82.88	50.66	18.26	13.92	82.84
Formation Opportunities	9.63	29.15	21.81	60.59	9.57	26.73	20.32	56.62
Better Wage Opportunities	2.08	9.51	23.84	35.43	2.14	11.41	22.75	36.3
Social Benefits	1.04	7.78	7.96	16.78	1.10	7.00	7.58	15.68
Prestige	1.71	4.28	7.19	13.18	1.62	3.24	7.73	12.59
18 mil Soles Incentive	0.69	1.89	3.87	6.45	1.18	3.39	6.33	10.9
Panel B: What are the most important characteristic for your ranked choices? (% of respondents)								
Close to House	44.17	11.66	8.00	63.83	49.77	13.22	8.76	71.75
Safe	10.66	24.19	19.25	54.1	7.65	24.50	19.35	51.5
Well Connected	9.69	20.62	20.20	50.51	8.23	18.70	19.67	46.6
Prestige	17.92	14.12	12.29	44.33	21.13	15.77	12.68	49.58
Cultural Reasons	10.61	9.67	12.31	32.59	7.58	9.45	12.61	29.64
Good Infrastructure	2.02	8.40	12.86	23.28	1.81	7.23	11.83	20.87
Good Students	1.24	4.52	6.08	11.84	0.84	4.36	5.95	11.15
Possibility other Jobs	1.93	3.72	4.90	10.55	1.62	4.10	4.71	10.43
Career	1.76	3.10	4.09	8.95	1.36	2.67	4.44	8.47

Notes. This table displays the share of the 5,553 survey respondents that chose the corresponding answers to Question A and B. The first three columns show which answer they chose and how they ranked them (by order of importance) while column (4) shows the share of respondents that listed the corresponding choice in their top 3 reasons. The last four columns display the same results for respondents that scored above the top quartile of the test score distribution for tenured teachers.

Table B.4: Applicants for Temporary Teaching Positions

	2016		2018	
	All	w/ location	All	w/ location
	(1)	(2)	(3)	(4)
Female	0.708	0.723	0.706	0.722
	(0.455)	(0.447)	(0.456)	(0.448)
Age	36.703	36.959	38.278	38.537
	(6.801)	(6.889)	(6.930)	(6.987)
Indigenous language:	0.290	0.262	0.241	0.223
	(0.454)	(0.440)	(0.428)	(0.416)
- Quechua	0.246	0.225	0.205	0.192
	(0.431)	(0.417)	(0.404)	(0.394)
- Aimara	0.034	0.029	0.026	0.022
	(0.180)	(0.167)	(0.159)	(0.147)
Some experience in the public sector	0.772	0.766	0.775	0.771
	(0.419)	(0.423)	(0.418)	(0.420)
Experience in the public sector ($\geq$3 yrs)	0.550	0.541	0.568	0.564
	(0.498)	(0.498)	(0.495)	(0.496)
Some experience in the private sector	0.430	0.433	0.497	0.502
	(0.495)	(0.495)	(0.500)	(0.500)
Experience in the private sector ($\geq$3 yrs)	0.261	0.268	0.304	0.313
	(0.439)	(0.443)	(0.460)	(0.464)
Applied to both recruitment drives	0.918	0.918	0.834	0.836
	(0.274)	(0.275)	(0.372)	(0.370)
Overall competecy score (std)	-0.224	-0.221	-0.043	-0.037
	(0.944)	(0.950)	(1.011)	(1.016)
Curricular knowledge (std)	-0.089	-0.085	-0.094	-0.088
	(0.974)	(0.979)	(0.995)	(0.997)
N. of teachers	60,840	40,774	66,280	44,348

Notes. This table reports summary statistics for a set of characteristics for different samples of applicants to the assignment mechanism for temporary positions. These are: the universe of applicants to the centralized assignment system for contract positions (Column 1 and 3); the subset of these applicants for whom the information on their residential location (from the SISFOH dataset) is available (Column 2 and 4). The first two columns refer to the 2016 assignment system, while the remaining two to the 2018 one.

Table B.5: Characteristics of Temporary Teaching Positions

	2016			2018		
	All	w/ student score	w/ class-teach link	All	w/ student score	w/ class-teach link
	(1)	(2)	(3)	(4)	(5)	(6)
Vacancy characteristics						
Rural	0.694	0.565	0.408	0.665	0.552	0.425
	(0.461)	(0.496)	(0.492)	(0.472)	(0.497)	(0.494)
Bilingual	0.294	0.158	0.134	0.239	0.116	0.062
	(0.456)	(0.364)	(0.341)	(0.427)	(0.320)	(0.242)
Multigrade	0.362	0.277	0.134	0.335	0.255	0.113
	(0.481)	(0.448)	(0.341)	(0.472)	(0.436)	(0.317)
Single-teacher	0.100	0.015	0.019	0.082	0.016	0.008
	(0.300)	(0.123)	(0.135)	(0.275)	(0.126)	(0.087)
N. of teachers (school)	6.579	9.614	13.029	6.616	9.108	11.789
	(8.750)	(10.173)	(11.458)	(7.899)	(8.797)	(9.257)
School infrastructure						
Library	0.390	0.478	0.595	0.395	0.468	0.550
	(0.488)	(0.500)	(0.491)	(0.489)	(0.499)	(0.498)
Computer	0.759	0.845	0.914	0.755	0.827	0.888
	(0.428)	(0.362)	(0.281)	(0.430)	(0.378)	(0.315)
Internet	0.504	0.609	0.746	0.533	0.620	0.716
	(0.500)	(0.488)	(0.435)	(0.499)	(0.485)	(0.451)
Sport facility	0.394	0.493	0.599	0.400	0.484	0.576
	(0.489)	(0.500)	(0.490)	(0.490)	(0.500)	(0.494)
Locality infrastructure						
Electricity	0.889	0.946	0.977	0.885	0.941	0.973
	(0.314)	(0.225)	(0.150)	(0.319)	(0.236)	(0.162)
Drinking water	0.715	0.792	0.872	0.711	0.773	0.837
	(0.451)	(0.406)	(0.335)	(0.453)	(0.419)	(0.369)
Sewage	0.513	0.638	0.766	0.519	0.622	0.725
	(0.500)	(0.481)	(0.424)	(0.500)	(0.485)	(0.447)
Medical clinic	0.618	0.746	0.845	0.636	0.743	0.830
	(0.486)	(0.435)	(0.362)	(0.481)	(0.437)	(0.375)
N. of teaching positions	19,743	13,587	6,547	21,299	15,813	11,216

Notes. This table reports summary statistics for a set of vacancy- and school-level characteristics for different samples of vacancies for temporary teaching positions that are made available through the centralized assignment mechanism. These are: the universe of vacancies (Column 1 and 4); the set of teaching positions in schools that took part in the standardized test evaluating student achievement (Column 2 and 5); the set of the former where, in addition, it also is possible to link students to teachers through a classroom-teacher matching (Column 3 and 6). The first three columns refer to the 2016 assignment system, while the remaining three to the 2018 one.

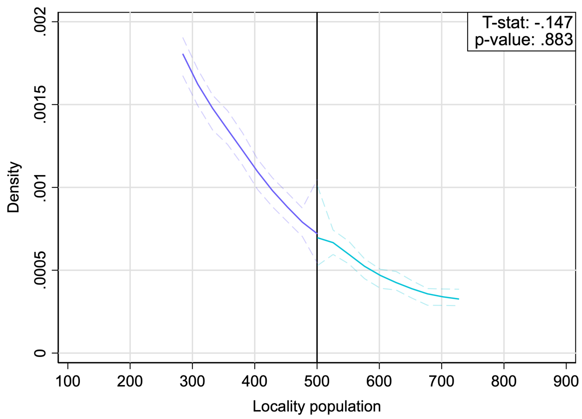
Table B.6: School and Locality Characteristics

	Rural schools		Urban Schools	
	Mean	Std. Dev.	Mean	Std. Dev.
Panel A: School characteristics				
N. of students	40.79	(48.04)	362.6	(285.4)
Bilingual school	0.346	(0.476)	0.0592	(0.236)
Single-teacher school	0.357	(0.479)	0.00764	(0.0871)
Multigrade school	0.496	(0.500)	0.0130	(0.113)
Number of teachers	2.519	(2.391)	15.71	(11.34)
Share of permanent teachers	0.614	(0.487)	0.775	(0.418)
Share of certified contract teachers	0.301	(0.459)	0.193	(0.395)
Share of non-certified contract teachers	0.0856	(0.280)	0.0326	(0.177)
Competency score (certified teachers)	0.0904	(0.978)	0.529	(1.024)
Panel B: Student characteristics				
Math test scores (std)	-0.455	(0.997)	0.127	(0.965)
Math test scores: % Below basic	0.218	(0.413)	0.0649	(0.246)
Math test scores: % Proficient	0.157	(0.364)	0.312	(0.463)
Spanish test scores (std)	-0.561	(0.945)	0.141	(0.953)
Spanish test scores: % Below basic	0.231	(0.421)	0.0565	(0.231)
Spanish test scores: % Proficient	0.156	(0.363)	0.374	(0.484)
Panel C: School infrastructure				
Library	0.304	(0.460)	0.655	(0.476)
Computer	0.724	(0.447)	0.961	(0.193)
Internet	0.308	(0.462)	0.942	(0.233)
Sport facility	0.317	(0.465)	0.692	(0.462)
Gym	0.0230	(0.150)	0.118	(0.323)
Cafeteria	0.307	(0.461)	0.208	(0.406)
Teachers' room	0.996	(0.0650)	0.994	(0.0771)
Panel D: Locality infrastructure				
Electricity	0.805	(0.396)	0.991	(0.0940)
Drinking water	0.585	(0.493)	0.940	(0.238)
Sewage	0.262	(0.440)	0.910	(0.287)
Medical clinic	0.328	(0.470)	0.863	(0.344)
Police	0.0845	(0.278)	0.530	(0.499)
Village phone	0.0501	(0.218)	0.0918	(0.289)
Internet access point	0.0583	(0.234)	0.837	(0.369)

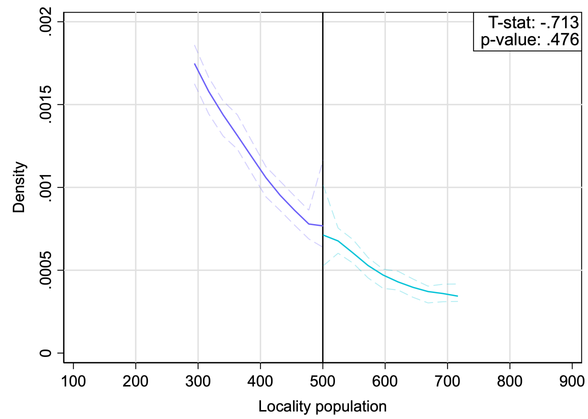
NOTES. This table reports the summary statistics for the universe of rural and urban primary schools in Perú over the period 2016-2018. The first panel describes the baseline characteristics of each type of school (size, bilingual spanish/indigenous language curriculum) for the year 2016, and the teaching staff composition (pooling together the post-recruitment drives years 2016 and 2018). The second panel summarizes students' achievement in the 2016 and 2018 standardized test. The third and the fourth panel describes the quality and quantity of school infrastructures and locality amenities, as measured by the 2016 school census.

C RD Evidence

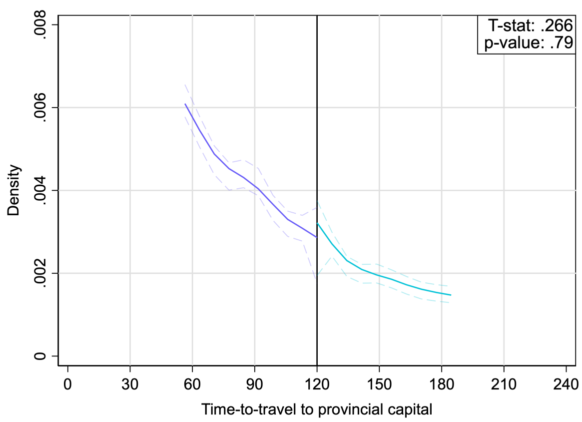
Figure C.1: Manipulation charts



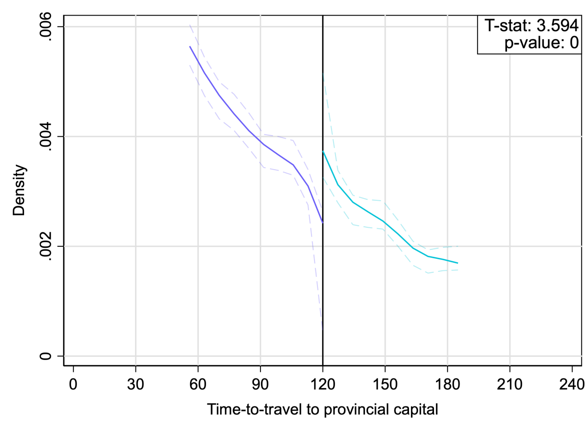
a. Population (2016)



b. Population (2018)



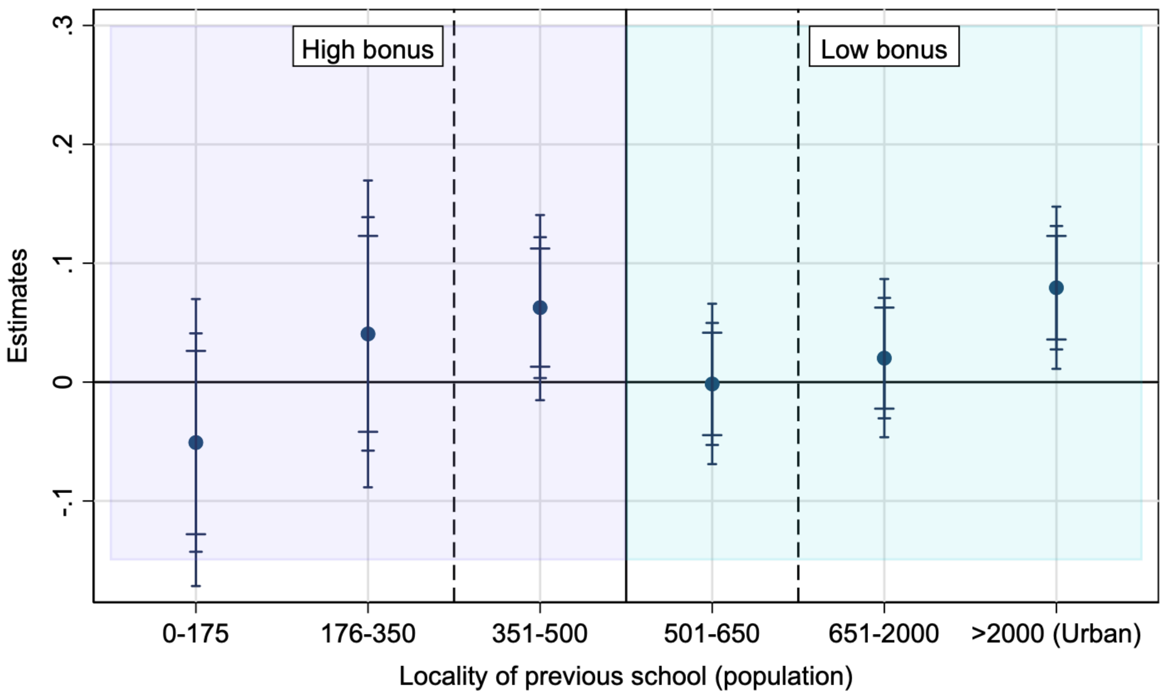
c. Time-to-travel (2016)



d. Time-to-travel (2018)

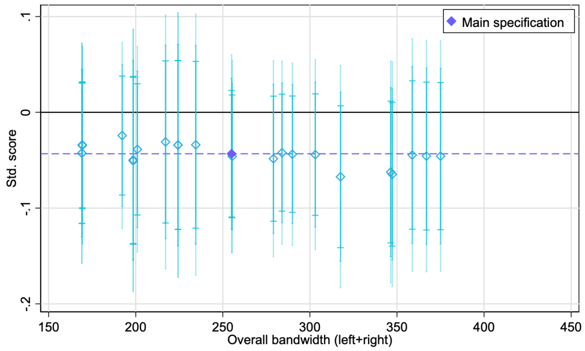
Notes. The figure displays the empirical densities with the corresponding confidence intervals for two running variables (population and time-to-travel) for each of the years in which the teacher recruitment drive was conducted (2016 and 2018). The density is computed using the local-polynomial estimator proposed in Cattaneo et al. (2020), and the figures show the 95% confidence intervals. The sample includes all schools with a contract teacher opening in the corresponding year.

Figure C.2: Rural Bonus and the Origin of Newly Recruited Teachers

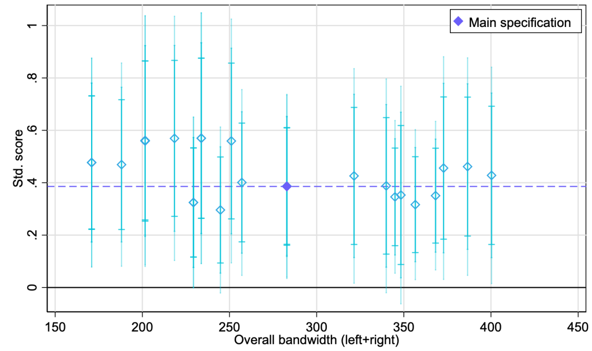


Notes. Each dot represents the effect of crossing the population threshold on the probability that the newly recruited teacher in 2015 – before the wage bonuses for contract teachers were introduced – was working in a school whose location falls in the population bin indicated in the x-axis. The sample includes the contract teacher vacancies assigned to certified teachers in the 2016 and 2018 processes. Bars report the bias-corrected regression-discontinuity estimates along with confidence intervals at the 90, 95, and 99% level obtained using the robust estimator proposed in Calonico et al. (2014). The dashed vertical lines indicate the population bin falling within the optimal bandwidths.

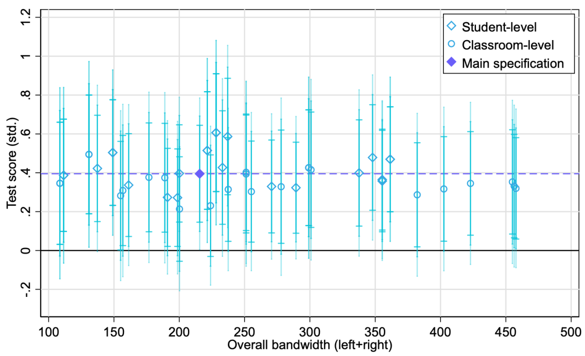
Figure C.3: Robustness to Alternative RD Specifications



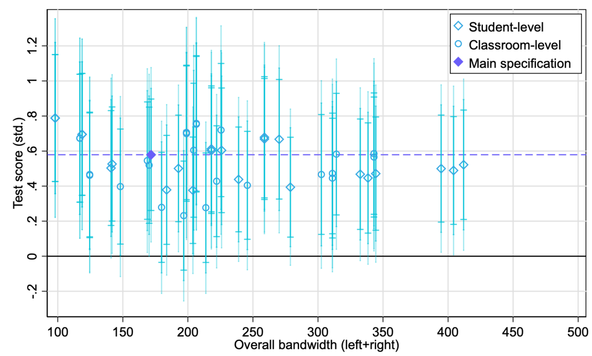
a. Vacancy filled



b. Teacher scores (std)



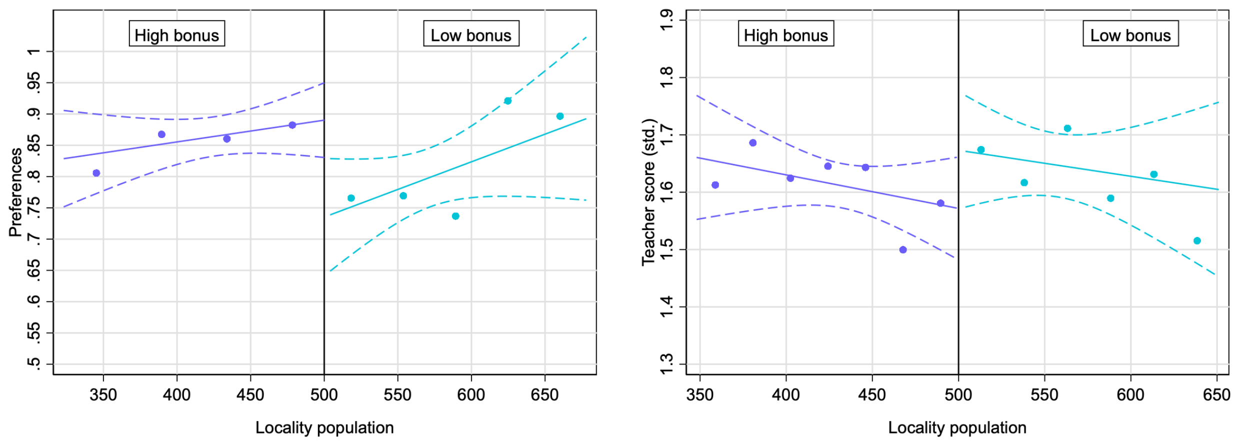
c. Spanish scores



d. Math scores

NOTES. These figures illustrate the robustness of the regression discontinuity estimates to alternative specifications and estimation choices. In Panel A, the outcome variable is whether the vacancy is filled, in Panel B is the teacher competency score, while in Panel C and D is the student achievement in Spanish and math respectively. Markers indicate how the robust bias-corrected regression-discontinuity estimates obtained using the robust estimator proposed in Calonico et al. (2014) varies depending on whether i) the bandwidth is the common mean-square error (MSE) optimal bandwidth, the MSE-optimal bandwidth for the sum of regression estimates, the common coverage error rate (CER) optimal bandwidth, or the CER-optimal bandwidth selector for the sum of regression estimates; ii) the kernel functions used to construct the local-polynomial estimator is uniform or triangular; iii) whether the unit of observation is a student, or the classroom (in which case the outcome is the classroom-level average of the test scores). Specification iii) only applies to Panel C and D. The horizontal dashed line indicates the estimates obtained under the main specification, which uses the common mean-square error (MSE) optimal bandwidth, a triangular kernel, and where the unit of observation is the vacancy (panel A and B) or the student (panel C and D).

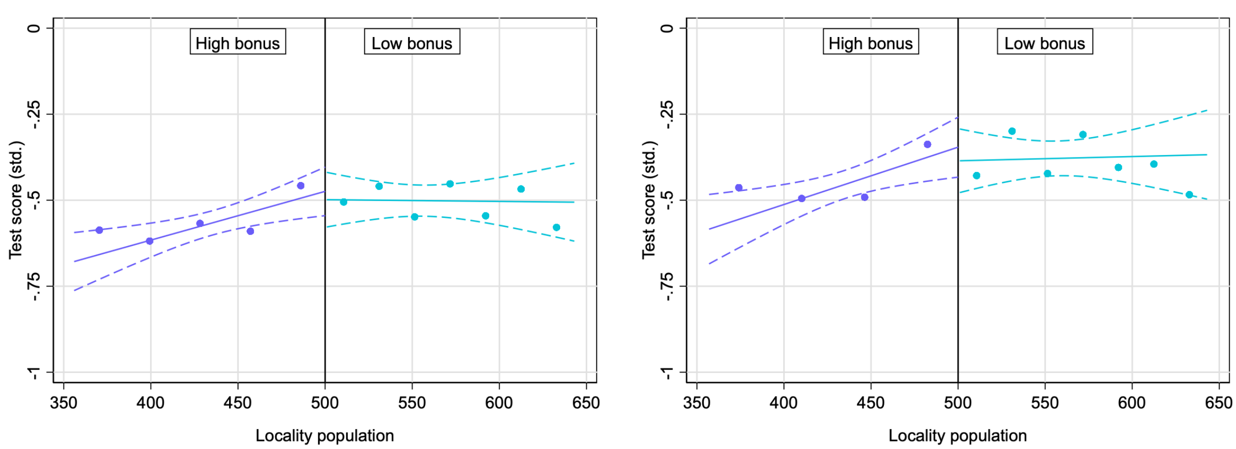
Figure C.4: Permanent Teacher Choices over Job Postings



a) Stated Preferences **b) Competency Score**

Notes. This figure shows how applicants' preferences and quality vary based on the difference between the 500-inhabitants cutoff and the population of the community where the school is located. In Panel A, the unit of analysis is the school, and the outcome variable is a dummy equal to one if a school was mentioned in at least one application. In Panel B, the unit of analysis is a (filled) vacancy and the outcome variable is the standardized (total) score obtained in the centralized test by the newly assigned permanent teacher. Both figures are obtained by pooling the data across two school years (2016 and 2018). Each marker indicates the median of the outcome variable within each bin, defined following the IMSE-optimal evenly spaced method by Calonico et al. (2015). Solid lines represent the predictions from linear regressions estimated separately for observations to the left and to the right of the cutoff, assuming a triangular kernel function. Dashed lines are 95% asymptotic confidence intervals.

Figure C.5: Student Achievement in Schools with Open Positions for Permanent Teachers



a) Spanish **b) Math**

Notes. This figure shows how applicants' preferences and quality vary based on the difference between the 500-inhabitants cutoff and the population of the community where the school is located. In Panel A, the outcome variable is a dummy equal to one if a school was mentioned in at least one application, while in Panel B the outcome variable is the standardized (total) score obtained in the centralized test by the newly-assigned permanent teacher. Each marker indicates the median of the outcome variable within each bin, defined following the IMSE-optimal evenly spaced method by Calonico et al. (2015). Solid lines represent the predictions from linear regressions estimated separately for observations to the left and to the right of the cutoff, assuming a triangular kernel function. Dashed lines are 95% asymptotic confidence intervals.

Table C.1: Covariate Smoothness around the Population Cutoff

	2016			2018		
	Mean (BW)	RD estimate		Mean (BW)	RD estimate	
	(1)	(2)	(3)	(4)	(5)	(6)
School characteristics						
Other wage bonuses (S/.)	91.631	-7.793	(23.028)	67.797	3.566	(19.659)
Distance to the provincial capital	134.337	62.455	(41.435)	172.070	5.913	(48.251)
Distance >120 min.	0.352	0.103	(0.098)	0.464	-0.017	(0.081)
Number of students	102.255	-1.662	(11.310)	105.186	4.734	(8.033)
Indigenous language students	0.259	-0.125	(0.118)	0.253	0.084	(0.080)
% indigenous language students	0.179	-0.127	(0.080)	0.160	0.014	(0.055)
% proficient students (math)	10.799	0.816	(5.347)	10.537	0.167	(3.673)
% proficient students (spanish)	14.734	2.237	(5.582)	14.569	0.663	(3.560)
Village amenities						
Electricity	0.858	0.001	(0.093)	0.851	-0.054	(0.069)
Drinking water	0.688	0.167	(0.135)	0.656	0.003	(0.102)
Sewage	0.407	0.080	(0.127)	0.368	0.008	(0.095)
Medical clinic	0.718	0.269	(0.118)	0.716	0.034	(0.088)
Meal center	0.257	0.115	(0.103)	0.245	0.017	(0.071)
Community phone	0.335	0.112	(0.130)	0.330	-0.025	(0.114)
Internet access point	0.135	0.051	(0.092)	0.125	0.035	(0.070)
Bank	0.010	0.021	(0.013)	0.007	0.014	(0.010)
Public library	0.036	0.005	(0.042)	0.034	-0.044	(0.033)
Police	0.167	-0.071	(0.089)	0.183	-0.172	(0.086)
School amenities						
Distance from district municipality (min.)	202.249	225.645	(294.152)	211.283	-34.191	(110.058)
Teachers room	0.176	0.074	(0.091)	0.164	-0.023	(0.072)
Sport pitch	0.208	-0.065	(0.105)	0.206	-0.032	(0.082)
Courtyard	0.212	-0.134	(0.116)	0.203	-0.097	(0.094)
Administrative office	0.538	-0.048	(0.110)	0.511	-0.107	(0.096)
Courtyard	0.012	-0.026	(0.032)	0.007	0.007	(0.005)
Computer lab	0.404	0.101	(0.126)	0.415	-0.007	(0.102)
Workshop	0.064	0.029	(0.034)	0.056	-0.019	(0.036)
Science lab	0.124	0.104	(0.089)	0.090	0.107	(0.052)
Library	0.501	0.155	(0.127)	0.480	-0.040	(0.114)
At least a personal computer	0.816	0.013	(0.101)	0.828	0.143	(0.088)
Electricity	0.807	0.070	(0.096)	0.804	-0.009	(0.083)
Water supply	0.636	0.153	(0.143)	0.601	0.014	(0.094)
Sewage	0.401	0.072	(0.118)	0.350	0.052	(0.088)

Notes. Notes. This table studies whether schools in localities just above or below the population threshold differ in terms of village and school amenities (as of 2013). Columns (1) to (3) focus on the 2016 assignment process. They report the mean of the variable considered within the bandwidth (Column 1), and the robust bias-corrected regression-discontinuity estimates obtained using the robust estimator proposed in Calonico et al. (2014), along with the standard error (Column 2 and 3). Columns (4) to (6) are the analogous of columns (1)-(3) but focus on the 2018 assignment process. Regressions are defined within a mean-square error optimal bandwidth.

Table C.2: Decomposition of wage increases around the population cutoff

	(1) Overall	(2) Time < 120	(3) Time > 120
Bonus	223.439	24.144	371.608
	(38.817)	(12.175)	(39.521)
Mean dep. var. (Lower Bonus)	1947.481	1887.325	2022.793
Bandwidth	142.494	158.856	172.479
Schools	793	554	374
Observations	2012	1240	945

Notes. This table reports the effect of crossing the population threshold on teacher wages. In all columns, the dependent variable is defined as the sum of the base wage and the wage bonuses described in Figure 1 (for schools that satisfy the criteria). In Column (2), the sample is limited to schools that comply with the time-to-travel cutoff, that is, farther than 120 minutes from the provincial capital. Column (3) only considers schools that do not comply with the time-to-travel cutoff (closer than 120 minutes). In all columns, the sample does not consider schools closer than 30 minutes from the provincial capital and schools located in urban areas (population $\geq$ 2,000 inhabitants). Cells report the bias-corrected regression-discontinuity estimates obtained using the robust estimator proposed in Calonico et al. (2014). Regressions are defined within a mean-square error optimal bandwidth (BW), reported at the bottom part of the table. Standard errors are clustered at the school$\times$year level.

Table C.3: Sharp RD Bounds Under Potential Manipulation

	(1) Preferences	(2) Teacher Score (std.)
Upper Bound	0.121	0.424
Lower Bound	0.086	0.291
95% Confidence Interval [LB - UB]	[0.054 - 0.153]	[0.165 - 0.549]
Bandwidth	157.452	141.447
Schools	850	764
Observations	2080	1870

Notes. This table reports the RD bounds (Gerard et al., 2020) for the threshold crossing effect on two outcomes that are subject to potential censorship due to the assignment of applicants to vacancies. The sample only considers teaching positions that were assigned to a contract teacher during the 2016 or 2018 recruitment process (pooling together the two). In Column (1) the outcome variable is the rank in which a vacancy was chosen in the deferred acceptance mechanism (normalized so that it takes values from zero to one); in Column (2) it is the standardized competency score obtained by the teachers in the centralized test. The 95% confidence intervals are obtained through 1000 bootstrap replications.

Table C.4: Teacher Score (Std.)–Difference-in-discontinuity

	(1) 2016	(2) 2018	(3) Pooled
High Bonus	0.061	-0.045	-0.008
	(0.171)	(0.155)	(0.146)
Post-policy	0.052	0.032	0.040
	(0.124)	(0.111)	(0.105)
Post × High Bonus	0.254	0.397	0.346
	(0.158)	(0.144)	(0.136)
Bandwidth	172.779	174.872	176.099
Schools	432	867	955
Observations	1172	1937	2735

NOTES. This table reports the effect of crossing the population threshold on the (standardized) teacher score on a sample of vacancies observed before and after the introduction of the wage rurality bonuses for contract teachers. The sample includes all contract teacher vacancies assigned to certified teachers in 2015 (pre-policies) and in the 2016 and 2018 recruitment drives. For 2016 and 2018, the outcome variable is the standardized score obtained in the centralized test by the newly assigned teacher; for 2015, it is the standardized score that teachers obtained in the 2016 assignment process (thus available only for those who applied to the test the following year). $HighBonus$ is the binary indicator taking the value one for vacancies in localities below the population cutoff; $Post - Policy$ is the indicator taking the value of one when the year considered is 2016 (Column 1), 2018 (Column 2), or both (Column 3), and zero for the pre-policies year (2015). Each cell reports the conventional regression-discontinuity estimates obtained using a triangular kernel. Regressions are estimated within a mean-square error optimal bandwidth(BW), reported at the bottom of the table, defined – separately for each column – on the pooled sample of vacancies. Standard errors are clustered at the school level.

Table C.5: Sample Selection Around the Population Cutoff

	Schools	Teaching positions			
	(1) N. of vacancies	(2) Vacant	(3) w/student score	(4) w/class-teach link	(5) (2)&(3)&(4)
High Bonus	-0.063	-0.018	-0.048	-0.024	-0.017
	(0.126)	(0.021)	(0.033)	(0.032)	(0.018)
Mean dep. var. (Low Bonus)	1.599	0.261	0.837	0.412	0.106
Bandwidth	183.861	183.571	107.920	162.403	161.243
Schools	3303	3303	1737	2835	2813
Observations	6100	29813	16622	26167	25962

NOTES. This table studies the sample selection around the population threshold. In Column (1), the dependent variable is the number of teaching positions that were open as contract teacher vacancies in each school. Columns (2) to (5) consider a set of binary indicators defining the different samples used. These are: i) a dummy taking value 1 for the positions that, among the universe of teaching positions, were open as a contract teacher vacancy in the centralized assignment processes (Column 2); ii) a dummy taking value 1 for teaching positions in schools were students took the standardized ECE test assessing students' achievement (Column 3); iii) a dummy taking value 1 for teaching positions for which it is possible to link students to teachers through a classroom-teacher matching (Column 4); iv) a dummy taking value 1 when all the previous conditions i), ii), and iii) are satisfied. All columns pool together the recruitment drives years 2016 and 2018. Cells report the bias-corrected regression-discontinuity estimates obtained using the robust estimator proposed in Calonico et al. (2014). Regressions are defined within a mean-square error optimal bandwidth (BW), reported at the bottom part of the table. The table also reports the mean of the dependent variable computed within the intervals $(0, +BW)$ (right-hand-side of the cutoff, corresponding to Low Bonus schools). SE are clustered at the school×year level.

Table C.6: Monetary Incentives and Teacher Characteristics

	(1) Female	(2) Age	(3) Experience	(4) Quechua	(5) Other indig.	(6) Degree
High Bonus	0.013	-0.120	0.023	-0.059	0.007	0.075
	(0.057)	(0.766)	(0.045)	(0.073)	(0.010)	(0.052)
Mean dep. var. (Low Bonus)	0.641	36.056	0.833	0.253	0.009	0.310
Bandwidth	128.130	138.817	162.189	193.026	179.129	198.893
Schools	715	777	911	1086	991	1126
Observations	1725	1849	2159	2541	2342	2622

NOTES. This table reports the effect of crossing the population threshold on several teachers' characteristics. These are a female dummy (column 1), age (column 2), a dummy taking value 1 for teachers with at least 3 years of teaching experience (column 3), a dummy equal to 1 if the teacher speaks a Peruvian indigenous language (column 4 and 5), an indicator for university or technical institute education (column 5). The sample includes all contract teacher vacancies assigned in the 2016 and 2018 processes. Cells report the bias-corrected regression-discontinuity estimates obtained using the robust estimator proposed in Calonico et al. (2014). Regressions are defined within a mean-square error optimal bandwidth (BW), reported at the bottom part of the table. The table also reports the mean of the dependent variable computed within the intervals $(0, +BW)$ (right-hand-side of the cutoff). SE are clustered at the school×year level.

Table C.7: Monetary Incentives and Teaching Staff Composition

	(1) N. of teachers	(2) Student/Teacher	(3) % Contract teachers
High Bonus	0.298	-0.311	-0.063
	(0.364)	(0.218)	(0.051)
Mean dep. var. (Low Bonus)	6.055	2.890	0.504
Bandwidth	155.448	128.807	143.647
Schools	871	721	800
Observations	1113	922	1011

NOTES. This table reports the effect of crossing the population threshold on several teachers' characteristics. These are a female dummy (column 1), age (column 2), a dummy taking value 1 for teachers with at least 3 years of teaching experience (column 3), a dummy equal to 1 if the teacher speaks a Peruvian indigenous language (column 4), an indicator for university or technical institute education (column 5). The sample includes all contract teacher vacancies assigned in the 2015 and 2017 processes, regardless of whether they were assigned to certified or non-certified teachers. In column (4) the sample includes only vacancies assigned during the 2015 assignment process, as the same information is not available for 2017. Cells report the bias-corrected regression-discontinuity estimates obtained using the robust estimator proposed in Calonico et al. (2014). Regressions are defined within a mean-square error optimal bandwidth (BW), reported at the bottom part of the table. The table also reports the mean of the dependent variable computed within the intervals $(0, +BW)$ (right-hand-side of the cutoff) and $(-BW, 0]$ (left-hand-side of the cutoff, corresponding to Low Bonus schools). SE are clustered at the school×year level.

Table C.8: Monetary Incentives and Teacher Retention

	(1) Within-year retention	(2) Same school in t+1	(3) Same school in t+2
High Bonus	0.010	0.019	0.008
	(0.028)	(0.041)	(0.047)
Mean dep. var. (Low Bonus)	0.926	0.495	0.191
Bandwidth	191.245	237.799	160.354
Schools	1074	1360	900
Observations	2523	3131	2133

NOTES. This table reports the effect of crossing the population threshold on several measures of teacher retention. These are a set of binary indicators for whether a contract teacher assigned to a certain school through the centralized process - -observed as of the beginning of the school year (March) – is also observed in the same school at the end of the school year (Column 1), at the beginning of the following school year (Column 2), or at the beginning of the two-years-after school year. Cells report the bias-corrected regression-discontinuity estimates obtained using the robust estimator proposed in Calonico et al. (2014). Regressions are defined within a mean-square error optimal bandwidth (BW), reported at the bottom part of the table. The table also reports the mean of the dependent variable computed within the intervals $(0, +BW)$ (right-hand-side of the cutoff) and $(-BW, 0]$ (left-hand-side of the cutoff, corresponding to Low Bonus schools). SE are clustered at the school×year level.

D Sorting Model

D.1 Identification of $\Sigma_{\delta,\theta}$: Proof

Under Assumption 1 and 3, we can write the probability of observing the matching history $\{\mu^*(i,t)\}_{t=1}^T$ conditional on observed teacher and school characteristics, teachers' choice sets and their value-added coefficients δ_i:

$$\mathbb{P}(\{\mu^*(i,t)\}_{t=1}^T | x_{it}, \boldsymbol{w}, \boldsymbol{a}, \Omega(s_{it}), \delta_i) = \int \prod_{t=1}^{T} \frac{\exp\{\alpha_i w_{\mu^*(i,t)t} + u(a_{\mu^*(i,t)t}, x_{it})\}}{\exp\{\beta_i\} + \exp\{x'_{it}\beta_p\} + \sum_{k\in\Omega(s_{it})} \exp\{\alpha_i w_{kt} + u(a_{kt}, x_{it})\}}$$
$$\times\, \phi(\theta_i | x'_{1it}\gamma^\theta + \Sigma_{\theta,\delta}\Sigma^{-1}_{\delta,\delta}(\delta_i - x'_{2it}\gamma^\delta), \Sigma_{\theta,\theta} - \Sigma_{\theta,\delta}\Sigma^{-1}_{\delta,\delta}\Sigma_{\delta,\theta}) d\theta_i$$

From Fox et al. (2012) we know that this relationship can be inverted to identify the mean and variance of the mixing distribution. The following function $\tilde{\gamma}$ is thus identified:

$$\tilde{\gamma}(x_{1it}, x_{2it}, \delta_i) = x'_{1it}\gamma^\theta + \Sigma_{\theta,\delta}\Sigma^{-1}_{\delta,\delta}(\delta_i - x'_{2it}\gamma^\delta)$$

From there, we can use variation in δ_i to identify $\Sigma_{\theta,\delta}\Sigma^{-1}_{\delta,\delta}$. Conditional on knowing $\Sigma_{\delta,\delta}$, we can recover $\Sigma_{\theta,\delta}$.

D.2 Estimation Procedure

Estimation is done in two steps. First, we take the empirical counterparts of Equations (9) and (10) and solve for $\hat{\beta}$ and $\hat{\delta}_i$ for all i. When the matrix $\sum_{l\in\mu(i,t)} z_{lt}z'_{lt}$ is not invertible due to a lack of within-teacher variation in matched students' characteristics, we follow the procedure described in Ahn et al. (2023) and replace z_{lt} by a subset of students' characteristics that we call z^*_{lt} such that $\sum_{l\in\mu(i,t)} z^*_{lt}z^{*\prime}_{lt}$ is invertible to solve for $\hat{\delta}_i$. From there, assuming that $\mathbb{E}\left[\nu^2_{lijt} | c_{jt}, z_{lt}, i = \mu_s(l,t), j = \mu_w(i,t)\right] = \sigma^2$, estimates of teacher value added coefficients $\hat{\delta}_i$ are asymptotically distributed as follows:

$$\hat{\delta}_i \sim \mathcal{N}\left(W_i\delta_i, \Sigma_i\right)$$

where $W_i = \mathbb{E}\left[z^*_{\mu(i,t)t} z^{*\prime}_{\mu(i,t)t}\right]^{-1} \mathbb{E}\left[z^*_{\mu(i,t)t} z'_{\mu(i,t)t}\right]$ and $\Sigma_i = \sigma^2 \mathbb{E}\left[z^*_{\mu(i,t)t} z^{*\prime}_{\mu(i,t)t}\right]$. Note that W_i collapses to the identity matrix when $z^*_{\mu(i,t)t} = z_{\mu(i,t)t}$. Under Assumption 3 and using Equation (12), this implies that the conditional probability density function of $\hat{\delta}_i$ can be written as:

$$f(\hat{\delta}_i|x_{1it}, x_{2it}, \theta_i) = \phi\left(\hat{\delta}_i|W_i(x'_{2it}\gamma^{\delta} + \Sigma_{\delta,\theta}\Sigma^{-1}_{\theta,\theta}(\theta_i - x'_{1it}\gamma^{\theta})), W_i(\Sigma_{\delta,\delta} - \Sigma_{\delta,\theta}\Sigma^{-1}_{\theta,\theta}\Sigma_{\theta,\delta})W'_i + \Sigma_i\right)$$

Second, we take the empirical counterparts of W_i and Σ_i using an unbiased estimator for σ^2 and construct the likelihood function described in Equation (15) such that:

$$\mathbb{P}(\{\mu^*(i,t)\}_{t=1}^T, \hat{\delta}_i|x_{it}, \boldsymbol{w}, \boldsymbol{a}, \Omega(s_{it})) =$$
$$\int \prod_{t=1}^{T} \frac{\exp\{\alpha_i w_{\mu^*(i,t)t} + u(a_{\mu^*(i,t)t}, x_{it})\}}{\exp\{\beta_i\} + \exp\{x'_{it}\beta_p\} + \sum_{k\in\Omega(s_{it})}\exp\{\alpha_i w_{kt} + u(a_{kt}, x_{it})\}} f(\hat{\delta}_i|x_{1it}, x_{2it}, \theta_i)\phi(\theta_i|\gamma^{\theta}, \Sigma_{\theta,\theta})d\theta_i,$$

where the integral is approximated using Halton sequences.

D.3 Prior and Posterior Distribution of δ_i

From our estimates of $\Sigma_{\theta,\theta}$, $\Sigma_{\delta,\theta}$, $\Sigma_{\delta,\delta}$, γ^{θ} and γ^{δ} we can construct the prior and posterior distribution of δ_i conditional on observable teacher characteristics (x_{i1}, x_{i2}) and random coefficients θ_i using Bayes rule.

$$f(\delta_i|x_{1it}, x_{2it}, \theta_i) = \phi\left(\delta_i|\hat{E}^p_i, \hat{V}^p_i\right)$$

where $\hat{E}^p_i = x'_{2it}\hat{\gamma}^{\delta} + \hat{\Sigma}_{\delta,\theta}\hat{\Sigma}^{-1}_{\theta,\theta}(\theta_i - x'_{1it}\hat{\gamma}^{\theta})$ and $\hat{V}^p_i = \hat{\Sigma}_{\delta,\delta} - \hat{\Sigma}_{\delta,\theta}\hat{\Sigma}^{-1}_{\theta,\theta}\hat{\Sigma}_{\theta,\delta}$ are respectively the expectation and the variance of the prior distribution of δ_i.

$$f(\delta_i|, x_{1it}, x_{2it}, \theta_i, (Y_{lijt}, z_{lijt})_{l\in\mu(i,t)}) = \phi\left(\delta_i|\hat{E}^P_i, \hat{V}^P_i\right)$$

where $\hat{E}^P_i = \hat{E}^p_i + \hat{V}^p_i Z'_i(Z_i\hat{V}^p_i Z'_i + \hat{\sigma}^2 I)^{-1}(Y_i - Z_i\hat{E}^p_i)$ and $\hat{V}^P_i = \hat{V}^p_i - \hat{V}^p_i Z'_i(Z_i\hat{V}^p_i Z'_i + \hat{\sigma}^2 I)^{-1}Z_i\hat{V}^p_i$ are respectively the expectation and the variance of the posterior distribution of δ_i. Note that (Y_i, Z_i) are, respectively, the vector and matrix collecting Y_{lt} and z_{lt} for all students l matched with teacher i where the number of rows corresponds to the number of students.

D.4 Monte Carlo Simulations

We fix the draw of θ_i, assume that the true distribution of the teacher value-added coefficients is $\delta_i | x_{1it}, x_{2it}, \theta_i \sim \mathcal{N}(\hat{E}_i^p, \hat{V}_i^p)$, where $\hat{E}_i^p, \hat{V}_i^p$ are defined in Appendix D.3, and assume that $\nu_{lit} \sim \mathcal{N}(0, \hat{\sigma}^2)$. We also fix the teacher and student population, the realized teacher-school match, as well as their observed characteristics.

For each iteration, we draw a vector of teacher value-added coefficients δ_i and a draw of shocks ν_{lit} to construct students' residualized test scores. We then derive the OLS estimates of δ_i and construct their posterior distribution as described in Appendix D.3. We then solve problem (16) in three scenarios: using the true coefficients δ_i, the shrunken estimates of δ_i, and the OLS estimates of δ_i. We then compute the true total efficiency gains with respect to the Status Quo allocation achieved in the solution of the three problems. We average out across 100 draws.

D.5 Additional Results

Table D.1: Teacher Preferences – γ^θ

		Male	Teacher Score	Urban	Exp Public > 3	Exp Private > 0
	(1)	(2)	(3)	(4)	(5)	(6)
Wage	-0.484	0.200	0.171	-0.001	0.024	-0.047
	(0.095)	(0.041)	(0.024)	(0.036)	(0.041)	(0.040)
Outside Option	3.045	-0.813	-1.441	0.955	-1.727	1.249
	(0.356)	(0.206)	(0.118)	(0.198)	(0.222)	(0.220)

Notes. This Table displays the estimates of the parameters of the conditional mean of θ_i. Standard errors are in parentheses.

Table D.2: Teacher Preferences – $\Sigma_{\theta,\theta}$

	Wage	Outside Option
	(1)	(2)
Wage	0.918	0.584
	(0.048)	(0.084)
Outside Option	-	1.839
		(0.038)

Notes. This Table displays the estimates of the variance covariance matrix of θ_i. The diagonal elements are the standard deviations of θ_i while the off diagonal elements display corr(θ_i). Standard errors are in parentheses.

Table D.3: Teacher Preferences – Γ

	(1)	Male (2)	Teacher Score (3)	Urban (4)	Exp Public > 3 (5)	Exp Private > 0 (6)
Amenity Index	0.108	-0.033	0.001	-0.043	0.014	-0.017
	(0.027)	(0.021)	(0.011)	(0.020)	(0.024)	(0.024)
Border	-0.106	0.206	0.126	0.049	-0.029	-0.166
	(0.099)	(0.075)	(0.043)	(0.074)	(0.086)	(0.087)
VRAEM	-0.214	0.129	0.115	0.087	0.273	0.122
	(0.124)	(0.081)	(0.050)	(0.082)	(0.100)	(0.089)
Multigrade	-0.556	0.204	-0.021	0.012	0.165	0.041
	(0.060)	(0.047)	(0.028)	(0.045)	(0.054)	(0.051)
Single Teacher	-1.085	0.719	0.034	-0.121	0.237	0.291
	(0.100)	(0.076)	(0.044)	(0.074)	(0.090)	(0.084)
Bilingual	-1.137	0.018	-0.445	-0.282	0.325	-0.339
	(0.082)	(0.060)	(0.035)	(0.060)	(0.075)	(0.073)
Time	-0.253	0.064	-0.070	0.059	0.079	-0.122
	(0.024)	(0.018)	(0.011)	(0.017)	(0.022)	(0.021)
Time^2	-0.001	-0.001	0.001	0.000	0.001	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log(Population)	0.148	-0.064	-0.002	0.045	-0.043	0.025
	(0.051)	(0.040)	(0.022)	(0.038)	(0.041)	(0.041)
log(Population)^2	-0.024	0.001	0.004	0.001	0.001	0.003
	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
Time $\times$ log(Pop)	0.043	-0.003	-0.002	-0.007	-0.015	0.022
	(0.004)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
Outside School	2.413	-0.842	-1.261	0.940	-0.207	0.887
	(0.357)	(0.205)	(0.117)	(0.196)	(0.220)	(0.218)
Previous School	5.130	-0.283	-0.166	-0.631	-0.053	0.684
	(0.059)	(0.045)	(0.022)	(0.042)	(0.052)	(0.047)
Same Ethnicity: Quechua	1.701	-0.115	0.310	0.052	-0.372	0.222
	(0.105)	(0.078)	(0.047)	(0.078)	(0.097)	(0.093)
Same Ethnicity: Aymara	2.641	-0.521	0.472	-0.691	0.053	0.649
	(0.244)	(0.189)	(0.135)	(0.193)	(0.199)	(0.201)
Distance Spline: $[0, 20)$	-0.112	-0.009	-0.024	0.049	0.001	0.011
	(0.004)	(0.003)	(0.002)	(0.003)	(0.004)	(0.003)
Distance Spline: $[20, 100)$	-0.045	0.000	-0.007	-0.007	-0.002	0.003
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
Distance Spline: ≥ 100	-0.006	-0.001	0.001	-0.001	-0.001	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes. This Table displays the estimates of the parameters of the function capturing preferences over non-pecuniary amenities u. Standard errors are in parentheses.

Table D.4: Student Achievement – β

	Math	Spanish
	(1)	(2)
School Lagged Score	0.270 (0.066)	0.319 (0.066)
School Lagged Score^2	0.120 (0.060)	0.142 (0.053)
School Female	-0.733 (0.286)	-1.302 (0.293)
School Quechua-Aymara	-0.291 (0.336)	-0.332 (0.336)
School Age	-0.747 (0.161)	-0.274 (0.157)
School Infrastructure Index	0.039 (0.019)	0.038 (0.019)
Classroom Lagged Score	-0.188 (0.043)	-0.233 (0.046)
Classroom Lagged Score^2	-0.070 (0.035)	-0.094 (0.032)
Classroom Female	0.972 (0.189)	1.119 (0.196)
Classroom Quechua-Aymara	-0.302 (0.198)	-0.343 (0.195)
Classroom Age	0.039 (0.097)	-0.134 (0.093)
Class Size	-0.003 (0.003)	-0.009 (0.003)

Notes. This table displays the estimates of the parameters associated with school and classroom characteristics in the student achievement production function. Standard errors are in parentheses.

Table D.5: Student Achievement – $\Sigma_{\delta,\delta}$

	ATE	Lagged Score	Lagged Score^2	Female	Quechua-Aymara	Age
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Math						
ATE	0.465	0.036	-0.201	0.182	0.068	0.204
	(0.006)	(0.030)	(0.055)	(0.067)	(0.188)	(0.061)
Lagged Score	-	0.145	0.171	0.023	0.133	-0.119
		(0.005)	(0.069)	(0.085)	(0.245)	(0.085)
Lagged Score^2	-	-	0.049	0.387	0.749	0.217
			(0.004)	(0.157)	(0.660)	(0.164)
Female	-	-	-	0.098	-0.192	0.414
				(0.010)	(0.449)	(0.168)
Quechua-Aymara	-	-	-	-	0.040	0.001
					(0.030)	(0.501)
Age	-	-	-	-	-	0.115
						(0.007)
Panel B: Spanish						
ATE	0.408	0.129	-0.182	-0.042	0.025	0.135
	(0.006)	(0.031)	(0.045)	(0.076)	(0.114)	(0.070)
Lagged Score	-	0.150	0.262	-0.250	-0.180	-0.036
		(0.005)	(0.058)	(0.105)	(0.155)	(0.090)
Lagged Score^2	-	-	0.061	0.251	0.211	0.319
			(0.003)	(0.147)	(0.205)	(0.123)
Female	-	-	-	0.083	0.046	0.071
				(0.013)	(0.296)	(0.205)
Quechua-Aymara	-	-	-	-	0.067	0.387
					(0.019)	(0.330)
Age	-	-	-	-	-	0.110
						(0.008)

Notes. This Table displays the estimates of the variance covariance matrix of the prior distribution of teacher value added coefficients δ_i. The diagonal elements are the standard deviations of δ_i while the off diagonal elements display corr(δ_i). Standard errors are in parentheses.

Table D.6: Student Achievement – γ^{δ}

	$\bar{\delta}$	Score 1	Score 2	Score 3	Female	Quechua-Aymara	Non Certified
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Math							
ATE	0	-0.004	0.070	0.073	0.059	0.080	-0.147
		(0.011)	(0.011)	(0.011)	(0.018)	(0.023)	(0.037)
Lagged Score	0.487	0.011	0.015	0.008	0.013	-0.004	-0.018
	(0.008)	(0.006)	(0.005)	(0.005)	(0.009)	(0.011)	(0.018)
Lagged Score^2	0.002	0.001	0.004	-0.012	-0.007	-0.019	0.008
	(0.005)	(0.003)	(0.003)	(0.003)	(0.006)	(0.007)	(0.011)
Female	-0.091	-0.001	-0.004	0.000	0.012	0.010	0.055
	(0.011)	(0.007)	(0.007)	(0.007)	(0.012)	(0.016)	(0.025)
Quechua-Aymara	0.017	0.005	-0.016	0.000	0.002	-0.000	-0.006
	(0.013)	(0.008)	(0.008)	(0.007)	(0.014)	(0.016)	(0.030)
Age	-0.121	-0.005	-0.007	-0.016	-0.005	-0.002	0.022
	(0.013)	(0.009)	(0.008)	(0.008)	(0.014)	(0.020)	(0.025)
Panel B: Spanish							
ATE	0	-0.004	0.050	0.074	0.109	0.016	-0.074
		(0.011)	(0.010)	(0.010)	(0.017)	(0.021)	(0.037)
Lagged Score	0.545	0.001	0.018	0.014	0.030	-0.018	-0.017
	(0.008)	(0.006)	(0.005)	(0.005)	(0.009)	(0.012)	(0.019)
Lagged Score^2	-0.007	-0.004	-0.001	-0.009	-0.002	-0.008	0.027
	(0.005)	(0.003)	(0.003)	(0.003)	(0.005)	(0.007)	(0.011)
Female	-0.008	0.006	0.003	-0.007	-0.002	-0.001	0.072
	(0.011)	(0.008)	(0.007)	(0.007)	(0.012)	(0.016)	(0.029)
Quechua-Aymara	0.012	0.005	-0.017	-0.005	0.001	0.010	-0.024
	(0.013)	(0.009)	(0.008)	(0.008)	(0.014)	(0.016)	(0.031)
Age	-0.116	0.002	-0.003	-0.015	0.022	-0.013	0.049
	(0.012)	(0.009)	(0.008)	(0.008)	(0.014)	(0.019)	(0.025)

Notes. This Table displays the estimates of the parameters of the conditional mean of the prior distribution of the teacher value added coefficients. Standard errors are in parentheses.

Table D.7: Covariance Matrix – Cov(θ, δ)

	ATE	Lagged Score	Lagged Score^2	Female	Quechua-Aymara	Age
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Math						
Wage	-0.286	0.044	-0.001	-0.006	-0.025	0.023
	(0.008)	(0.005)	(0.003)	(0.008)	(0.008)	(0.008)
Outside Option	0.137	0.140	-0.008	0.013	-0.074	0.127
	(0.021)	(0.014)	(0.009)	(0.021)	(0.023)	(0.021)
Panel B: Spanish						
Wage	-0.243	0.047	-0.003	-0.009	-0.027	0.025
	(0.007)	(0.005)	(0.003)	(0.007)	(0.008)	(0.008)
Outside Option	0.151	0.148	-0.006	-0.030	-0.081	0.135
	(0.019)	(0.013)	(0.009)	(0.019)	(0.022)	(0.022)

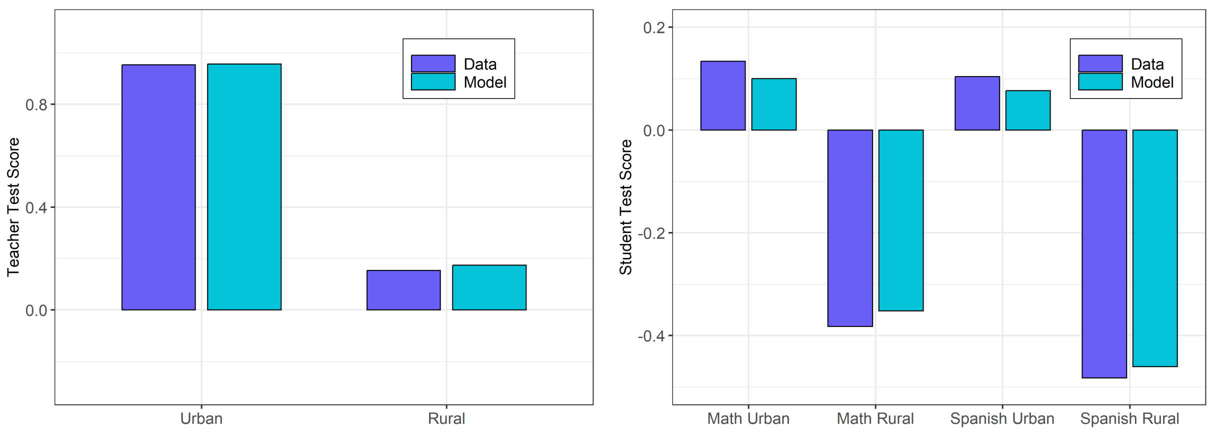
Notes. This Table displays the estimates of the covariances between the teacher value added coefficients δ_i and the random coefficients θ_i. Standard errors are in parentheses.

Table D.8: Model Fit, Matched Characteristics

	Mean		SD	
	Data	Model	Data	Model
	(1)	(2)	(3)	(4)
Wage	1.984	1.993	0.413	0.424
Distance	74.770	80.121	154.979	126.329
Amenity Index	0.065	0.050	0.923	0.930
Border	0.074	0.073	0.261	0.261
Vraem	0.050	0.052	0.218	0.221
Multigrado	0.337	0.346	0.473	0.476
Unidocente	0.081	0.082	0.273	0.275
Bilingue	0.240	0.241	0.427	0.428
Time	2.707	2.729	4.996	4.833
log(Pop)	7.733	7.708	3.580	3.564
Previous School	0.176	0.166	0.380	0.372
Same Ethnicity: Quechua	0.127	0.122	0.333	0.327
Same Ethnicity: Aymara	0.011	0.010	0.105	0.100
Outside School	0.305	0.302	0.460	0.459
$\times$ Male	0.250	0.248	0.433	0.432
$\times$ Std Score	-0.380	-0.362	0.897	0.880
$\times$ Urban Teacher	0.434	0.435	0.496	0.496
$\times$ Exp Public > 3	0.637	0.636	0.481	0.481
$\times$ Exp Private > 0	0.431	0.434	0.495	0.496
Outside Option	0.377	0.393	0.485	0.488
$\times$ Male	0.223	0.227	0.416	0.419
$\times$ Std Score	-0.368	-0.368	0.878	0.893
$\times$ Urban Teacher	0.424	0.427	0.494	0.495
$\times$ Exp Public > 3	0.362	0.382	0.481	0.486
$\times$ Exp Private > 0	0.539	0.528	0.498	0.499

NOTES. This Table assesses the fit of the model by comparing moments of the distribution of matched teacher and school characteristics in the data and in a simulated teacher-school status quo match. We simulate a teacher-school match using the teacher-proposing DA algorithm given the estimated preference parameters and a random draw of ϵ and θ to construct u_{ij} for all teacher-school pairs (i, j).

Figure D.1: Model Fit, Teacher Sorting and Spatial Inequalities



Notes. This Figure assesses the fit of the model by comparing the urban-rural gap in teacher competency and student achievement in the data and in a simulated teacher-school status quo match. We simulate a teacher-school match using the teacher-proposing DA algorithm given the estimated preference parameters and a random draw of ϵ and θ to construct u_{ij} for all teacher-school pairs (i, j). We predict student achievement in the simulated match using shrunken estimates of teacher value-added (see Appendix D.2 for details on constructing this posterior distribution).

Table D.9: Evaluation of the Rural Wage Bonus, Spanish Test Scores

	Status Quo (1)	No Rural Bonus (2)	Policy Effect (3)
Panel A: Total Value Added			
Urban-Rural Gap	0.105	0.175	-0.069
Urban	0.032	0.050	-0.018
Rural	-0.073	-0.125	0.051
Moderately Rural	-0.031	-0.038	0.007
Rural	-0.149	-0.109	-0.040
Extremely Rural	0.034	-0.117	0.150
Panel B: Match Effects			
Urban	-0.000	-0.003	-0.004
Rural	0.027	0.026	0.000

NOTES. Column (1) and Column (2) summarize the spatial distribution of the part of teacher value added attributed to match effects in the simulated teacher-school match in the status quo and in the absence of the rural wage bonus policy, respectively. Each simulated teacher-school match is constructed using the teacher-proposing DA algorithm given the estimated preference parameters and a fixed random draw of ϵ and θ to construct u_{ij} for all teacher-school pairs (i, j). Teacher value-added corresponds to shrunken estimates of $z'_{lt}\delta_i$ (see Appendix D.3 for details on how to construct the posterior distribution of δ_i). Column (3) is the difference between Column (1) and (2).

Table D.10: Evaluation of the Rural Wage Bonus, Other Matching Outcomes

	Status Quo (1)	No Rural Bonus (2)	Policy Effect (3)
Panel A: Share Filled Vacancies			
Urban	0.818	0.852	-0.034
Moderately Rural	0.852	0.882	-0.030
Rural	0.773	0.784	-0.011
Extremely Rural	0.775	0.636	0.139
Panel B: Teacher Competency Score			
Urban Rural Gap	0.501	0.712	-0.211
Urban	1.225	1.298	-0.073
Rural	0.724	0.586	0.138
Moderately Rural	1.091	1.104	-0.013
Rural	0.323	0.553	-0.230
Extremely Rural	0.869	0.383	0.486

NOTES. Column (1) and Column (2) summarize the spatial distribution of vacancies filling rate (Panel A) and teacher competency (Panel B) in the simulated teacher-school match in the status quo and in the absence of the rural wage bonus policy, respectively. Each simulated teacher-school match is constructed using the teacher-proposing DA algorithm given the estimated preference parameters and a fixed random draw of ϵ and θ to construct u_{ij} for all teacher-school pairs (i, j). Teacher value-added corresponds to shrunken estimates of $z'_{lt}\delta_i$ (see Appendix D.3 for details on how to construct the posterior distribution of δ_i). Column (3) is the difference between Column (1) and (2).

E Optimal Teacher Compensation Policies

E.1 School Preferences Satisfy the Substitutes Condition

We show that the preference ordering described by (P1)-(P2) satisfies the substitutes condition defined in Hatfield and Milgrom (2005). Denote the set of all possible contracts $X = S \times T \times W$ where S is the set of schools, T the set of teachers we consider and W the set of wages that schools can propose. We assume that wages range discretely from the minimum wage proposed to teachers in Perú to an arbitrarily large upper bound. Define $C_s(X)$ and $R_s(X)$ the chosen set and the rejected set of school s from the set of contracts X. Elements of X are substitutes for school s if for all subsets $X' \subset X'' \subset X$ we have $R_s(X') \subset R_s(X'')$.

Consider X' a subset of X. Define w^* the wage offered in $C_s(X')$, $\underline{t}^*$ as the teacher with the lowest value added in $C_s(X')$ and $\bar{t}^*$ as the teacher with the highest value added in $C_s(X')$. Consider that we add an additional contract to X' such that $X'' = X' \cup \{(s,t,w)\}$.

We first look at the case where $Y_{\bar{t}^*s} < c_s$. If $Y_{ts} \geq c_s$ then $C_s(X'') = \{(s,t,w)\}$ and $R_s(X'') = C_s(X') \cup R_s(X')$ for any w. If $Y_{ts} < c_s$ and $w > w^*$ then $C_s(X'') = C_s(X')$ and $R_s(X'') = C_s(X') \cup \{(s,t,w)\}$. If $Y_{ts} < c_s$ and $w < w^*$ then $C_s(X'') = \{(s,t,w)\}$ and $R_s(X'') = C_s(X') \cup R_s(X')$. Finally, if $Y_{ts} < c_s$ and $w = w^*$ two cases may arise:

- If the size of $C_s(X')$ is strictly smaller than school s capacities, under (P1), we have that $C_s(X'') = C_s(X') \cup \{(s,t,w)\}$ and $R_s(X') = R_s(X'')$.

- If the size of $C_s(X')$ is equal to school s capacities (school s is at max capacity), under (P1) we have: (i) $C_s(X'') = C_s(X')$ and $R_s(X'') = R_s(X') \cup \{(s,t,w)\}$ if $Y_{ts} < Y_{\underline{t}^*s}$, or (ii) $C_s(X'') = C_s(X') \setminus \{(s,\underline{t}^*,w)\} \cup \{(s,t,w)\}$ and $R_s(X'') = R_s(X') \cup \{(s,\underline{t}^*,w)\}$ if $Y_{ts} > Y_{\underline{t}^*s}$.

In any case, $R_s(X') \subseteq R_s(X'')$.

We then look at the case where $Y_{\bar{t}^*s} \geq c_s$. If $Y_{ts} < c_s$ then $C_s(X'') = \{(s,t,w)\}$ and $R_s(X'') = C_s(X') \cup R_s(X')$ for any w. If $Y_{ts} < c_s$ and $w > w^*$ then $C_s(X'') = C_s(X')$ and $R_s(X'') = C_s(X') \cup \{(s,t,w)\}$. If $Y_{ts} < c_s$ and $w < w^*$ then $C_s(X'') = \{(s,t,w)\}$ and $R_s(X'') = C_s(X') \cup R_s(X')$. Finally, if $Y_{ts} < c_s$ and $w = w^*$ two cases may arise:

- If the size of $C_s(X')$ is strictly smaller than school s capacities, under (P1), we have that $C_s(X'') = C_s(X') \cup \{(s,t,w)\}$ and $R_s(X') = R_s(X'')$.

- If the size of $C_s(X')$ is equal to school s capacities (school s is at max capacity), under (P1) we have: (i) $C_s(X'') = C_s(X')$ and $R_s(X'') = R_s(X') \cup \{(s,t,w)\}$ if $Y_{ts} < Y_{\underline{t}^* s}$, or (ii) $C_s(X'') = C_s(X') \setminus \{(s,\underline{t}^*,w)\} \cup \{(s,t,w)\}$ and $R_s(X'') = R_s(X') \cup \{(s,\underline{t}^*,w)\}$ if if $Y_{ts} > Y_{\underline{t}^* s}$.

In any case, $R_s(X') \subseteq R_s(X'')$.

E.2 Proof Proposition 1

Let us denote the school-optimal stable set of contracts given (P1)-(P2) as (μ^*, w^*). We first show that condition (C1) is satisfied under (μ^*, w^*). Assume that (C1) does not hold, this implies that there would exist a school k such that $Y_{\mu^*(k)k} < c_k$ which would be a direct contradiction of stability given that school k would be willing to keep increasing w_k above w_k^* until $Y_{\mu(k)k} \geq c_j$. A violation of (C2) would be a direct violation of stability as there would exist a teacher-school pair that would prefer to rematch given w^* under (P1)-(P2). This implies that (C2) holds under (μ^*, w^*). Finally, we know that the school-optimal stable set of contracts is unanimously preferred by all schools conditional on stability (Hatfield and Milgrom, 2005). This implies that, conditional on stability, the sum of the wages offered is minimal, which finishes to prove Proposition 1.

E.3 Counterfactual Teacher Compensation Policies

Table E.1: Fixed Priorities

	Status Quo	Teacher Value Added Threshold				
		$c = -0.4$	$c = -0.3$	$c = -0.2$	$c = -0.1$	$c = 0$
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Teacher Value Added						
Urban	0.024	0.033	0.026	0.035	0.036	0.010
Rural	-0.053	-0.053	-0.014	0.012	0.014	0.082
Moderately Rural	-0.033	-0.050	-0.050	-0.052	-0.025	0.116
Rural	-0.111	-0.020	0.021	0.013	-0.034	0.007
Extremely Rural	0.067	-0.046	0.003	0.048	0.100	0.113
Panel B: Match Effects						
Urban	-0.007	-0.005	-0.008	-0.006	-0.004	-0.002
Rural	0.008	0.015	0.022	0.039	0.033	0.031
Moderately Rural	-0.045	-0.037	-0.031	-0.025	-0.001	-0.019
Rural	0.031	0.039	0.043	0.052	0.039	0.059
Extremely Rural	0.047	0.033	0.047	0.073	0.047	0.050
Panel C: Monthly Total Cost (in Soles)						
% Base Wage	0.111	0.086	0.140	0.234	0.379	0.621
Mean Bonus per School	223	171	279	467	759	1,242
SD Bonus per School	220	407	576	839	1,184	1,698

Notes. This table summarizes the spatial distribution of teacher value added (Panel A) and its match effect component (Panel B) under the simulated equilibrium teacher-school match resulting from different wage bonus policies. Column (1) corresponds to the absence of the rural wage bonus, Columns (2) to (6) correspond to the wage policy solving Equation (17) when condition (C1) is to have at least one teacher with value-added above c in every rural school where c takes different values. Teachers' preferences over schools are constructed from the estimated preference parameters and a fixed random draw of ϵ and θ to construct U_{ijt} for all teacher-school pairs (i, j). Teacher value-added corresponds to shrunken estimates of $z'_{lt}\delta_i$ (see Appendix D.3 for details on how to construct the posterior distribution of δ_i). Panel C summarizes the cost of each policy.

Table E.2: All Certified Teachers

	No Rural Bonus	Status Quo	All Certified
	(1)	(2)	(3)
Panel A: Teacher Value Added			
Urban	0.059	0.024	0.044
Rural	-0.105	-0.053	-0.063
Moderately Rural	-0.055	-0.033	-0.054
Rural	-0.049	-0.111	-0.073
Extremely Rural	-0.099	0.067	-0.004
Panel B: Match Effects			
Urban	0.002	-0.007	-0.005
Rural	0.001	0.008	0.008
Moderately Rural	-0.052	-0.045	-0.045
Rural	0.051	0.031	0.024
Extremely Rural	0.020	0.047	0.057
Panel C: Monthly Total Cost (in Soles			
% Base Wage	0	0.111	0.065
Mean Bonus per School	0	223	130
SD Bonus per School	0	220	322

Notes. This table summarizes the spatial distribution of teacher value added (Panel A) and its match effect component (Panel B) under the simulated equilibrium teacher-school match resulting from different wage bonus policies. Column (1) corresponds to the absence of the rural wage bonus, Column (2) corresponds to the status quo wage bonus policy and Column (3) corresponds to the wage policy solving Equation (17) when condition (C1) is to have at least one certified teacher in every rural school. Teachers' preferences over schools are constructed from the estimated preference parameters and a fixed random draw of ϵ and θ to construct U_{ijt} for all teacher-school pairs (i, j). Teacher value-added corresponds to shrunken estimates of $z'_{lt}\delta_i$ (see Appendix D.3 for details on how to construct the posterior distribution of δ_i). Panel C summarizes the cost of each policy.