Improving the Distribution of Quality in Health Provision: Two Incentive Reforms in Peru's Healthcare System *

José Flor-Toro — *Northwestern University*

Matteo Magnaricotte — *University of Chicago*

This version: December 15, 2025

Abstract

We examine the labor supply decisions of healthcare providers in Peru and their consequences in a large sector characterized by staff shortages and inequitable supply. The Peruvian government implemented two incentive reforms to reduce unfilled posts and increase healthcare quality in health centers serving poorer, primarily rural populations. Using two regression discontinuity designs, we find that career advancement incentives attract significantly more skilled physicians to the targeted areas but are almost ineffective for nurses; the opposite is true for monetary incentives, which attract more skilled nurses to the neediest areas. However, applying the same empirical approach to rich survey data covering the pre- and post-reform periods, we don't find any change in utilization and insignificant or negative effects on the perceived quality of healthcare services received in the targeted areas.

JEL Classification: H51, I14, I18, J45.

*Lori Beaman, Gaston Illanes, Matthew Notowidigdo, and Chris Udry provided invaluable support and guidance for this project. Fabiola Alba-Vivar and Paula Natalia Barreto Parra contributed with extensive feedback and suggestions. The authors would like to thank Mario González at the Peruvian Ministry of Finance for helping coordinate with different Peruvian institutions relevant to this paper and its data. We are deeply grateful to Juan Manuel García for sharing data about the 2006 FONCODES Poverty Score with us and to Luis Palomino and Jorge Aguero for sharing DHS/ENDES data. The authors also thank Jimena Díaz Najera for her excellent research assistance in this paper. The Northwestern Development Group and Applied Micro Lunch participants have provided support and suggestions for this project.

1 Introduction

Access to primary healthcare has been increasing globally, but pockets with little to no access persist in developed and especially developing countries (Weiss et al., 2020). At the same time, reflecting the disappointing results of several past initiatives to increase access (Ansah et al., 2009; Powell-Jackson et al., 2015; Chari and Okeke, 2014; Godlonton and Okeke, 2016), attention has started to focus on the quality of services (Das et al., 2018; Kruk et al., 2016; Lagarde et al., 2019). Recent experimental evidence shows that having access to a physician instead of other mid-level healthcare providers plays an essential role in determining health outcomes (Okeke, 2023) and that more skilled physicians can improve health outcomes at birth Posso et al. (2024). However, how to build incentive-compatible systems that induce better providers to locate in high-need areas is unclear.

We study two incentive reforms that Peru introduced, aiming to decrease unfilled vacancies at rural health centers and increase the average quality of healthcare providers working there. Both reforms featured incentive schedules with discontinuous rewards based on the location of the health center posting a job: the first, introduced in 2009, provided career advancement incentives depending on the area's poverty level; the second reform, in 2014, offered monetary bonuses based on the location's elevation. Discontinuities in the incentives provided allow us to produce causal estimates of the reforms' effects on the supply of quality and quantity of physicians and nurses, as well as on relevant health-related outcomes.

The incentive reforms we study are incorporated into the structure of a semi-compulsory service program called *Rural and Marginal Urban Health Service* (SERUMS).[1] The SERUMS is a one-year program that all graduating health professionals have to undergo as a requirement for any career opportunity within the public sector. Due to the significance of the public sector, graduates in nursing and medicine, which we study in this paper, have near-universal participation rates. The system has been designed to reduce health-related gaps between the more rural, poorer areas and the more urbanized, richer ones. Staffing the poorer areas of Peru according to WHO's standards would require 8,500 physicians and 10,500 nurses, or, respectively, 21% and 23% of the existing stock of health professionals in each major.

[1] See Frehywot et al. (2010) for a survey of similar programs around the world and a description of common features.

SERUMS allocates health providers to health centers in a way that interacts with the two reforms of interest. Participants are ordered according to a combination of their college GPA and a national exam score, with higher-ranked participants choosing first among the available health centers. In 2008, a reform awarded bonus points for residency exams and for competitive public sector positions to providers according to the poverty quintile of the area of the selected health center. Then, in 2014, a new reform awarded health professionals who selected health centers above 3,800 meters above sea level a monetary bonus between 32% and 49% of their base pay. In both cases, high-quality providers could choose desirable locations, while lower-ranked providers could only select the remaining positions.

We use a regression discontinuity design based on two different running variables and three discontinuities. For the 2008 reform, we compare average provider quality in districts just below and just above the 60th (third to fourth quintile) or 80th (fourth to fifth quintile) percentile of the selected poverty measure. These two thresholds determined a meaningful increase in awarded career advancement incentives. Similarly, we compare average provider quality at health centers just below and above the elevation threshold of 3,800, determining eligibility for additional pay. Controlling for the running variables and continuity assumptions allows us to identify the causal effect of the two schemes. To do this, we collect several characteristics of all physicians and nurses participating in SERUMS between 2008 and 2018, including their college GPA, national assessment scores, and the selected health center. For each health center, we obtain the poverty measure used in the first incentive reform, elevation, and other characteristics.

Our findings show that both incentive types were at least partially successful in attracting more and better providers to disadvantaged areas. Career advancement incentives attracted better physicians to poorer areas but were unsuccessful in motivating nurses; on the other hand, monetary incentives attracted more and better nurses but had smaller, non-significant effects on doctors. As the reforms did not affect the number of job postings, increases in the number of providers came from a reduction in unfilled vacancies. We assess the robustness of our findings to changes in specifications, bandwidth choice, and aggregation level and confirm the overall pattern. We present different hypotheses for the difference in monetary and career incentives' effectiveness for different providers. One explanation lies in the differences in physicians' and nurses' lifetime gains and career trajectories: an early career boost

has a greater net present value for a physician than for a nurse. Preference heterogeneity and the different demographic composition in the two professions can also account for such observation. However, we find very little evidence of major differences in the responsiveness to incentives along the gender and ethnic dimensions.

We then turn to health-related outcomes. Based on prior evidence and the reforms' goals, we focus on healthcare utilization and the overall quality of healthcare services. We measure both using the Peruvian Demographic and Family Health Survey waves between 1990 and 2019. We use the geolocation of respondents to replicate the previously introduced regression discontinuity analysis. We don't find any change in utilization and mostly insignificant effects on the quality of services received, with career incentives showing significantly lower perceived quality of services at one of the two analyzed discontinuities along the poverty measure.

This study relates to several strands of literature in economics. The public sector has employed many different personnel and organizational policies to achieve its goals. Finan et al. (2015) provide stylized facts for the differences between public and private sector employment in low-, medium-, and high-income countries and surveys several field experiments related to selection, incentive structures, and monitoring in public employment. As an example of the surveyed work, Dal Bó et al. (2013) shows that higher compensation can increase applications and the quality of the final hires. A trade-off highlighted in this literature is that between prosociality and talent. Deserranno (2019) and Ashraf et al. (2020) find that higher compensation attracts more applicants, but can discourage or crowd out more prosocial individuals. Whether the final outcomes are positively or negatively affected depends on multiple context-dependent factors.

Most similarly, in Peru, Bobba et al. (2021) study the impact of an unconditional change in the structure of teacher compensation inequality in access to qualified teachers and student test scores. Unlike this study, they find that the reduction in inequality was effective in reducing students' achievement gaps. Using market design tools, they show that the optimal policy is more cost-effective.

Other features of the SERUMS program have been the object of scholarly research. Recently, Agte and Bedoya (2023) have studied the effects of participation on psychologists' preferences, beliefs, and career choices. Most relevant to our work, some studies have ana-

lyzed the motivation of nurses in their location choice (see, e.g., Huicho et al., 2012, 2015). Overall trends in the labor market (including its relation to SERUMS) for physicians, nurses, and midwives during 2007-2013 are described in Jimenez et al. (2017).

Finally, this paper relates to the literature pertaining provider quality and quantity and health outcomes. Okeke (2023) provides experimental evidence that receiving a physician rather than a mid-level healthcare provider significantly decreases mortality. Posso et al. (2024) focuses on physicians and leverages experimental variation in their assignment to different hospitals in Colombia, showing that more skilled physicians improve neonatal health outcomes. Das and Hammer (2014) discusses the multiple constraints in providing high-quality healthcare that are prevalent in developing countries. Friedrich and Hackmann (2021) find detrimental impacts of a nurse shortage in Denmark on hospital care as measured by readmission rates and nursing home mortality. Our work shows the limitations of even partially successful incentive schemes in affecting access to improved health services.

These results have several practical implications. First, they suggest that incentives should be tailored to each category of health professionals, prioritizing career incentives for physicians and monetary incentives for nurses. The latter observation is validated by prior studies on provider preferences in Peru: Huicho et al. (2012) use a discrete choice experiment to evaluate nurses' location preferences, showing the primacy of financial considerations. Second, the lack of impact on healthcare utilization and perceived quality of services suggests that the targeted measure of quality might be too noisy or not relevant to the outcomes of the study. Even substantial increases in provider quality, measured by their academic abilities, appear insufficient to increase utilization or the perception of service quality. Lastly, cost efficiencies could be gained by removing ineffective incentives to attract providers to the desired locations.

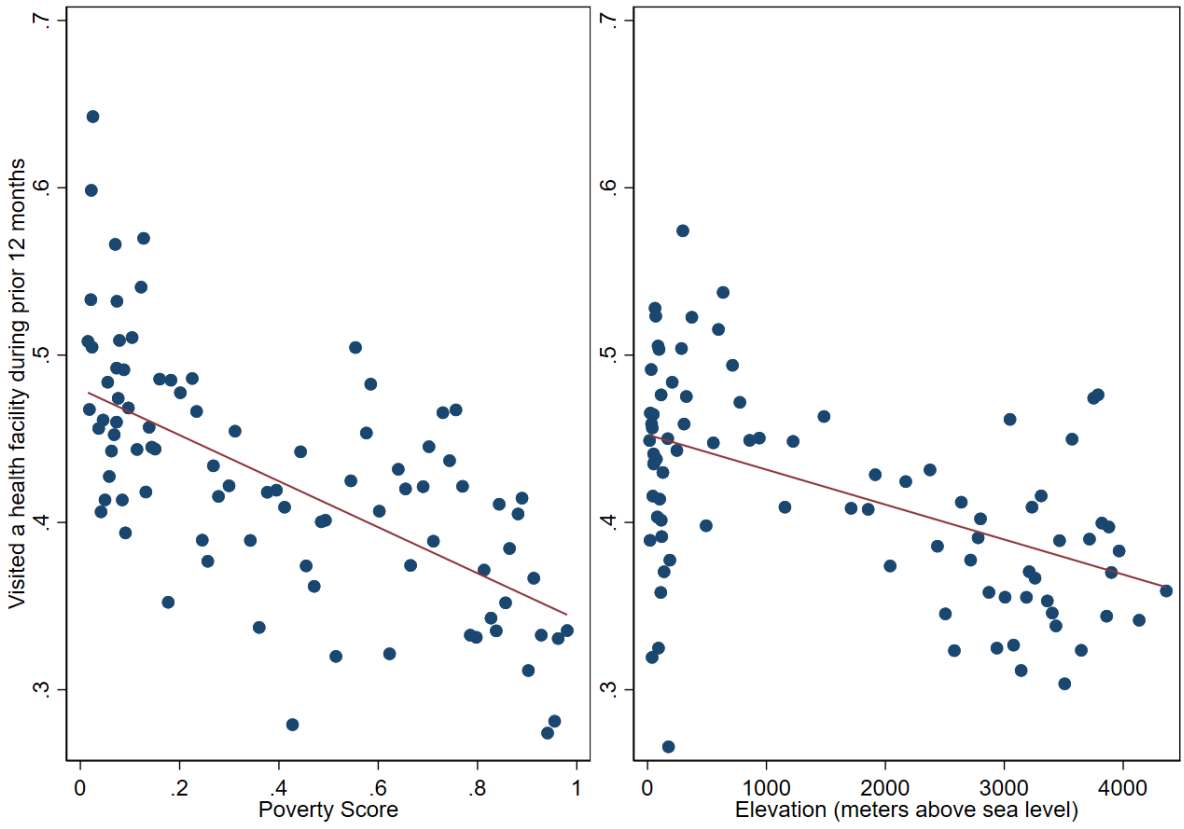
2 Context

2.1 Access to healthcare in Peru

Peru presents wide gaps and inequalities in access to healthcare. This is observed in the difference in development outcomes, including health-related outcomes, between the more

rural, poorer highlands and the more urbanized, richer coastal areas.[2] Figure 1 shows one example of this inequality using administrative birth records: women in poorer or higher elevation areas are systematically less likely to have visited a health facility in the previous twelve months.

Figure 1: Poverty and Elevation Gradient for Healthcare Access



Notes: The figure presents a binned scatter plot displaying the relation between access to healthcare and poverty (left panel) and elevation (right panel) from DHS data in the period 1996-2007. A 10p.p. increase in the poverty score implies a 3% decrease in the probability of having visited a health facility, while a 1,000 meters increase in elevation decreases it by 4%.

An important aspect of these gaps and inequalities is related to human resources. In general, poor and remote areas have difficulty attracting health professionals through regular hiring mechanisms,[3] which leads to an unequal distribution of human resources. As a

[2]The Peruvian jungle region, which accounts for more than a third of the surface area but around a tenth of the population, is also a low-elevation area. This region is characterized by very rural and remote areas and very poor health outcomes.

[3]As in many developing countries, tenured public jobs are a substantial share of public employment. Health

result, in 2007 only the richest quintile of districts in Peru met the World Health Organization's standards for physician and nurse staffing of 10 per 10,000 people; the other districts in Peru only had between a fifth and a half of this number (Observatorio Nacional de Recursos Humanos en Salud, 2009). The number of health professionals needed to close this gap is substantial: 8,500 physicians and 10,500 nurses, which are, respectively, 21% and 23% of the existing stock of health professionals in each major. Peru heavily relies on professional labor outside regular hiring mechanisms. SERUMS, a compulsory service program, is one such alternative way to bring providers to rural and remote public health centers.

2.2 The SERUMS system

The *Rural and Marginal Urban Health Service*, or SERUMS by its Spanish acronym, is a one-year semi-compulsory service program that all graduating health professionals must undergo as a requirement for any career opportunity within the public sector.[4] Since the public sector is the main employer in health professions and offers most learning positions, such as residencies, participation rates in the program are very high. While all kinds of health professionals participate in the program (physicians, dentists, nurses, midwives, pharmacists, nutritionists, social workers, biologists, psychologists, public health engineers, and veterinarians), we focus on graduates in nursing and medicine, which represent the largest share of participants and whose participation is nearly universal.

Graduating health professionals can opt for either paid or unpaid positions in the SERUMS system. Paid positions offer competitive monthly wages for a 36-hour workweek, and since a public institution funds them, they are usually located in very remote and poor areas. Unpaid positions only demand an 18-hour workweek. Both types of positions offer non-monetary incentives (e.g., access to residency scholarships) and provide reimbursement for moving expenditures. Regardless of the position they choose, the service is 12 months long.

Allocation to these positions happens through a centralized allocation system. Each year

providers can also be hired through temporary contracts (which lack the usual labor benefits but can pay higher wages) or internships (very limited in number and the scope of responsibilities). Private provision of healthcare is limited in poor, remote areas.

[4]This includes positions posted by the Peruvian Ministry of Health and by EsSalud, the Peruvian public health insurance. SERUMS is akin to existing partially compulsory systems for allocating human resources to health care in other countries. As can be seen in Appendix Figure 5, at least 70 countries used such systems to guarantee the deployment of health professionals to rural and remote areas (Frehywot et al., 2010).

in March and October, SERUMS participants are sorted into positions in each of the 25 administrative regions of the country. Each sorting round is run by the region's Health Directorate. All health institutions communicate their staffing needs to the Health Directorate (e.g., how many nurses, physicians, midwives), which checks budget availability and other conditions and then posts the list of all available positions with their relevant features (e.g., name and location of health center, funding institution, paid/unpaid type, wages and other incentives).

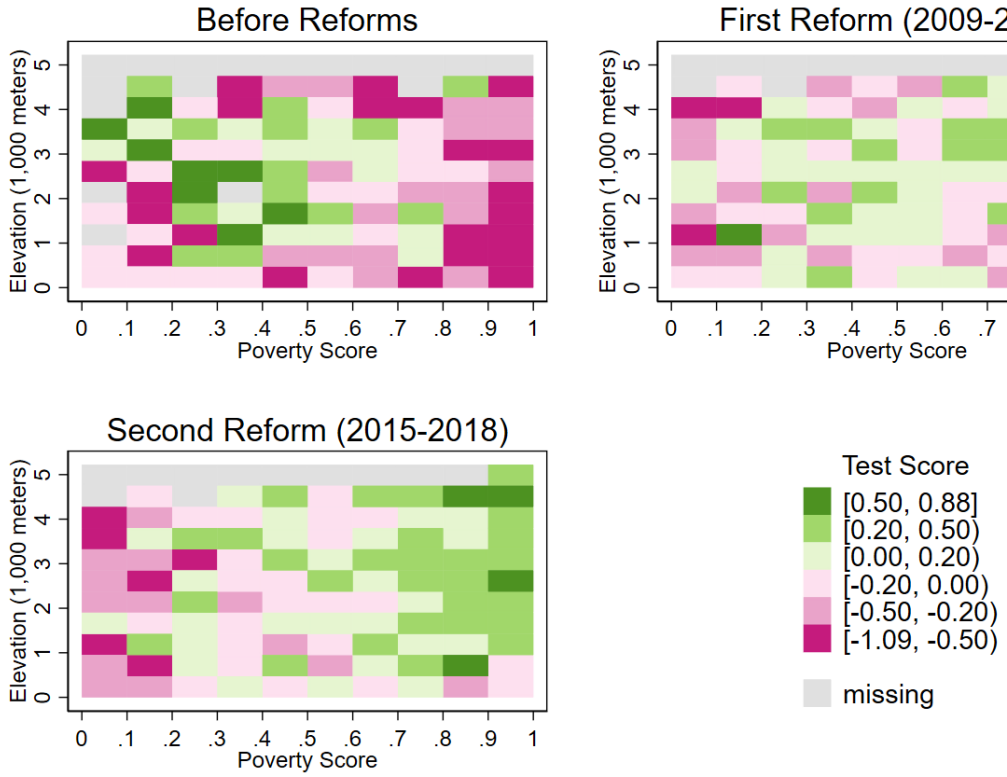
To participate in SERUMS, health professionals must have graduated from college[5] and joined the respective professional board. Graduates register in the same region of the college where they graduated. In the case of physicians and nurses, applicants must have taken a national exam, and disclose both their score in that exam and their college GPA. Both academic scores are then combined to create a priority number ρ_i, defined as $\rho_i = 0.7 \cdot \text{NatExam}_i + 0.3 \cdot \text{GPA}_i$.

On the day of the sorting, which is the same across all regions for each round and major, physicians and nurses applying to SERUMS gather and are called in order of their priority number ρ_i to publicly select a paid position. The list of each position's attributes (including any incentive provided, as discussed in the next sections) is made public prior to the sorting process. Participants can either choose a position or refuse to choose at all. At the end of this stage, professionals who have been allocated a position are given two days to give up their assigned position or swap it with another willing professional. In some cases, the remaining paid positions undergo a second sorting. Once all paid positions are allocated, unpaid ones are allocated among remaining applicants, which includes those who refused to choose in the first step.[6] Thanks to their higher priority in the sorting process, high-skilled professionals tend to select more urban, amenity-rich areas. The top-left panel in Figure 2 shows that until 2008 providers with higher national exam scores, who were granted higher priority in their choice of health center, were very unlikely to select health centers in the poorest or highest areas.

[5] In Peru, universities award a bachelor's degree for completion of all relevant coursework, and a licentiate's degree for successfully defending a dissertation (as well as any other major-specific requirements, such as special internships). Participation in SERUMS requires both.

[6] When some applicants are unable to get a position in their region and other regions still have leftover positions, a national sorting is carried out in the main branch of the Ministry of Health.

Figure 2: Average Scores in National Assessment for Physicians and Nurses



Notes: The three panels represent the distribution of provider quality as measured by their score in a national assessment test. Scores are standardized by provider type for each period. Average standardized scores are calculated over all providers employed through SERUMS in health centers for each elevation and poverty score cell. Gray areas represent cells without data, i.e., combinations of poverty score and elevation that don't have any provider in our dataset.

The Peruvian health system has become highly dependent on SERUMS participants, who represent more than 50% of all health professionals in districts in the bottom SES quintile (Observatorio Nacional de Recursos Humanos en Salud, 2009). Despite recently graduating, participants are often tasked with key health services, ranging from early childhood vaccination to births. In recent years the total number of participants in the program has roughly doubled, going from 6,000 to 10,000 per year: this amounts to 3/4 of the healthcare staffing gap estimated in 2009 (Observatorio Nacional de Recursos Humanos en Salud, 2009).

2.3 Non-Monetary Incentives: The 2008 Reform

In March 2008, a set of incentives was introduced to motivate high-skilled participants to choose more underserved, poorer areas. Table 1 shows the incentives schedule, which is composed of two parts: bonus points for the residency exam and for competitive public sector positions.

Residency exams are highly competitive examinations that provide specialized hands-on training: obtaining a residency in a well-reputed public hospital usually leads to better labor market opportunities in the public or private sector.[7] The incentive schedule rewards with more points those professionals performing SERUMS in poorer districts as defined by the 2006 poverty score classification. The 10 points awarded for working in the poorest areas amount to around 0.2 SD in the residency exam. The bonus points for competitive public-sector positions contribute to the overall evaluation that applicants to public jobs have to go through. This comprises a written exam of basic skills, an interview, and a check-up of professional experience. This evaluation is scored on a 0-100 scale, and SERUMS participants can obtain up to 15 points by choosing a health center in the poorest, rather than the richest, areas.

Table 1 shows that bonus points increase discontinuously based on the quintiles of a poverty score built in 2006 by the official entity tasked with funding development infrastructure in Peru (FONCODES, in Spanish).[8] See Section 3 for additional details.

Table 1: 2008 Reform - Points for Open Competitive Exams

Quintile	For Public Jobs	For Residencies
1 (Richest)	0	0
2	2	2
3	5	6
4	10	8
5 (Poorest)	15	10

Notes: Points awarded to health professionals for participating in SERUMS depending on the poverty of the location.

[7] Participation in residency programs is common for physicians and well-remunerated. However, it is less common for other health professions. For nurses, government-run residency programs have been available only since 2015 (DS-031-2015-SA).

[8] Other uses of this index are discussed in Section 5.5. No other policy we are aware of was based on the same discontinuities.

The 2008 incentives were part of a substantial revamping of the SERUMS system and were accompanied by a steady expansion in the number of SERUMS positions financed by the Government. However, to our best knowledge, none of these other changes were implemented differentially across poverty score cut-offs. The top-right panel in Figure 2 shows that more providers with higher national exam scores, who were granted higher priority in their choice of health center, chose to work in poorer areas. This is particularly evident when comparing this Figure to the pre-reform distribution in the top-left panel.

2.4 Monetary Incentives: The 2014 Reform

In August 2014, to motivate health professionals to serve in the most remote areas, new monetary incentives were added to SERUMS.[9]

Table 2 shows that these incentives benefited areas located in the coca-growing region (VRAEM, or *Valles de los Ríos Apurímac, Ene, y Mantaro* in Spanish) as well as areas considered to be remote or border zones (ZAF for the acronym in Spanish, *Zonas Alejadas o de Frontera*). For coca-growing regions, physicians were awarded an additional 26% monthly wage and for other health professionals (including nurses) the bonus was 54%. For ZAF regions, which we will focus on, bonuses for physicians and nurses were respectively 32% and 49%. Since bonus were cumulative, physicians going to health centers which were remote, high-elevation, and located in a border district were given a 96% increase ($3\times32\%$) over the base wage.[10] In practice, this is very uncommon.

Cartographers and the national statistical office previously defined the VRAEM area, while ZAF health centers were those 6 hours away or more from the nearest city, located in a border district, or above the 3,800 meters above the sea level (MSL) cut-off. As discussed later in the paper, we will focus our attention on the elevation threshold. The bottom panel in Figure 2 shows that more physicians with higher national exam scores, who were granted higher priority in their choice of health center, were fairly unlikely to select to work in low-elevation, richer areas. This distribution is in stark contrast to that shown in

[9] In September 2013 (DL 1153), the concept of additional bonus ('Valorización priorizada') was introduced. In July 2014 (DS 015-2014-SA), the criteria by which bonuses were to be assigned were established. This was after the assignment for SERUMS' first application round of 2014 but before the second application round. In August (DS 226-2014-EF), a month before the application process, the incentive amounts were defined. We therefore regard these extra incentives as being introduced in August.

[10] The base wage was increased uniformly in 2018 for each specialization.

Table 2: 2014 Reform - Monetary Incentives

Area	Physicians	Nurses
VRAEM (Coca-growing)	26%	54%
Remote ($\geq$ 6h. from cities)	32%	49%
Near a Border	32%	49%
High Elevation ($\geq$ 3,800 MSL)	32%	49%
Base Pay (2013)	USD 1,340	USD 660

Notes: Bonuses are defined as a percentage of the base pay, e.g. a nurse working in a high elevation area would receive $323 in addition to their $660 base pay, a 49% bonus.

the top-left panel (before any reform) and shows further progress in the goal of providing more skilled providers to the neediest areas when compared to the situation in the period 2009-2013 shown in the top-right panel.

3 Data

We use data produced by the Ministry of Education on SERUMS describing the final allocation of each health professional to a health center for each round of the program between 2008 and 2018. This data includes applicants' information, such as major and college of graduation, GPA, and the score in the major-specific national exam.[11] We obtain several time-invariant characteristics for each public health center, such as date of creation, precise geographical location, and elevation, from the publicly available National Registry of Health Institutions (RENAPRESS). We match more than 90% of all allocated applicants to a health center in RENAPRESS.

We augment this dataset by linking each health center to the poverty map produced by FONCODES in 2006 and used to determine the intensity of career incentives awarded through the 2008 reform. The poverty map calculates an index of poverty (or poverty score) based on several factors obtained from the 2005 Population Census and other data sources.[12]

[11] The data allows us to track the entire choice history of the applicants. For instance, we observe whether a candidate who eventually obtained an unpaid position had previously chosen a paid one and later rejected it.

[12] This score is based on a combination of the 2005 Population Census and the 1999 Height Census, which surveyed school-age children). The poverty score was determined for each one of the nearly 1,900 districts, based on indicators constructed from these sources (Díaz, 2006).[13] These variables were then combined through a Principal-Component Analysis, and the resulting first principal component was used to create population-weighted quintiles of districts.

Appendix Figure 6 shows the quintile of the poverty score for each Peruvian district. Finally, we impute missing values for elevation by combining the health centers' geolocation and an elevation raster.

Our preferred measure of quality for nurses and physicians will be the standardized SERUMS score. This score is a weighted average composed for 70% by a major-specific national exam (*Examen Nacional de Medicina* and *Examen Nacional de Enfermería*) and for 30% by the college GPA.[14] The national exams are high-stakes exams, relevant for career opportunities (e.g. SERUMS and residencies), that can only be taken once per year. While 'passing' is *not* a requirement to serve in SERUMS, professionals scoring below that threshold are understood to have insufficient knowledge of their field. We standardize the raw SERUMS score within year and major.

Table 3 reports the number of districts by poverty and elevation, along with several relevant features over the period 2008-2018.

Table 3: District-level Features

	(1) All	(2) Q1	(3) Q2	(4) Q3	(5) Q4	(6) Q5	(7) <3,800	(8) >3,800
N	1759	45	48	217	637	791	1531	207
Poverty Score		.015	.039	.095	.36	.803	.250	.412
Elevation		391	609	1014	1781	2370	1048	3943
Quality								
Physicians		-.22	-.29	-.29	-.08	.1	-.04	.06
Nurses		-.07	-.30	-.01	.13	-.05	0	.12
Quantity								
Physicians		1.7	1.4	1.6	1.4	1.5	1.5	1.3
Nurses		4.9	4.8	1.9	1.4	1.5	1.7	1.5

Notes: The table reports district-level features. Columns (2)-(6) report statistics by poverty quintile (see Section 1.3 for more details), while columns (7) and (8) report statistics according to whether the average elevation of health centers in the district is above or below 3,800 meters over sea level. We observe 1,759 unique districts out of the current 1,874. Poverty score quintiles are obtained using the 2005 Census counts for each district as weights, hence their uneven frequency. Poverty Score and Elevation are fixed over time for each district and we report averages weighted by the 2005 Census counts for each district. Quality and Quantity change over time for each district. We calculate the average quality (see Section 3 for more details) and number of health professionals in a given district over the period 2008-2018 and report a simple average of the districts within each group.

We use two health-related outcomes from a cross-sectional, yearly survey of the Peruvian

[14]While a measure like college GPA may depend on differences in electives or may be subject to grading differences across colleges, such standardized exams are not subject to these challenges. Nonetheless, the unconditional correlation of the two measures is very high (0.8).

population, the Demographic and Family Health Survey (ENDES in its Spanish acronym) conducted by the National Institute of Statistics and Informatics (INEI in Spanish). We use all the available waves of this survey for the years 1990 to 2019. The survey is designed to have representativeness at the national, regional, and urban/rural levels in all the latest waves. Additional details about the survey methodology can be found on INEI's ENDES website at proyectos.inei.gob.pe/endes/.

Women between 12 and 49 are eligible to answer the questionnaire we analyze. We build two outcomes from the data to measure healthcare utilization and the quality of healthcare provided. The first is obtained using a binary variable with a value equal to 1 if the interviewed woman has visited a health center in the last 12 months.[15] Quality of healthcare received is measured through the average of three binary indicators taking value equal to 1 if (1) they rated the treatment received from health personnel in their visits to a health center "Good" (the available options are "Good," "Regular," "Bad," "Don't know"), (2) they rated the treatment received from the doctors in their visits to a health center "Good" (same options), and (3) they rated the explanations received from the doctors in their visits to a health center "Good" (same options). We label the average of those three indicators "Good Healthcare Index".

4 Empirical Strategy

In this section, we present our basic hypotheses, show some descriptive facts motivating our research design, and describe the estimating equations that will identify the parameters of interest.

4.1 Hypotheses and Descriptive Facts

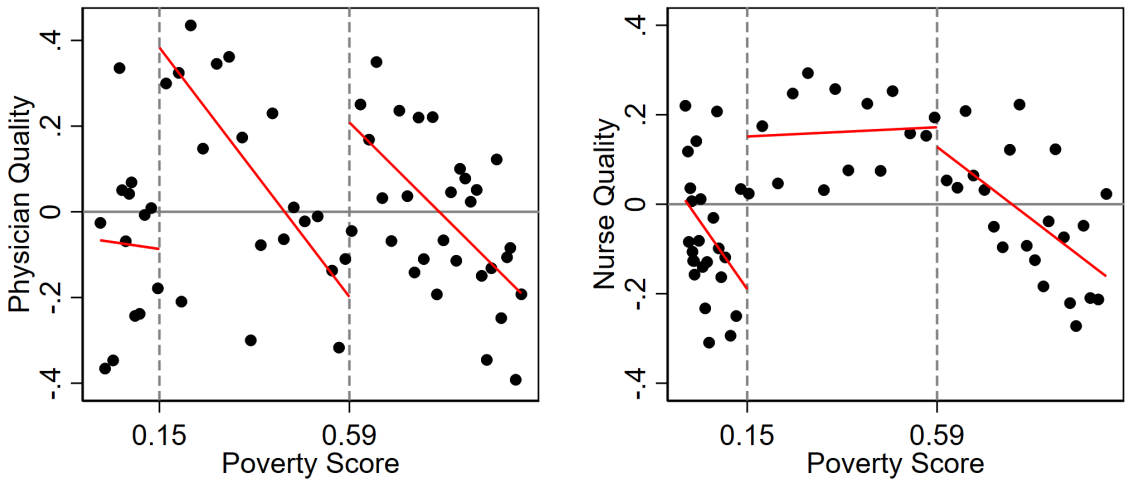
Our baseline hypothesis is that providing either non-monetary or monetary incentives will increase the attractiveness of job positions. In turn, we will observe that the average quality of health professionals at health centers providing these incentives will increase, due to the allocation system giving higher priority to better professionals. Secondarily, we hypothesize

[15] We choose this outcome due to its importance in prior work in Peru (Bernal et al., 2017 finds that utilization leads to supplier-induced demand for health services).

that the increase in quality will be reflected in an improvement in health outcomes.

Figure 3 plots local polynomials estimating the relationship between the standardized scores of physicians and the poverty of their selected SERUMS location. We estimate separately these polynomials for three bins of the 2006 poverty score: the first three quintiles, the fourth, and the fifth. We observe discontinuous jumps at the thresholds between these three bins and an overall negative relationship between average provider quality and poverty. The two discontinuities that emerge will be at the center of the two regression discontinuity approaches described in the next section. Figure 3 shows the same relationship for nurses: we observe a smaller but clear increase in the average exam score at the first threshold, but not at the second. It is worth noting that neither figure should be interpreted causally as they pool together scores of professionals involved in SERUMS at different moments in time and in different regions.

Figure 3: Relation between Provider Quality and Poverty

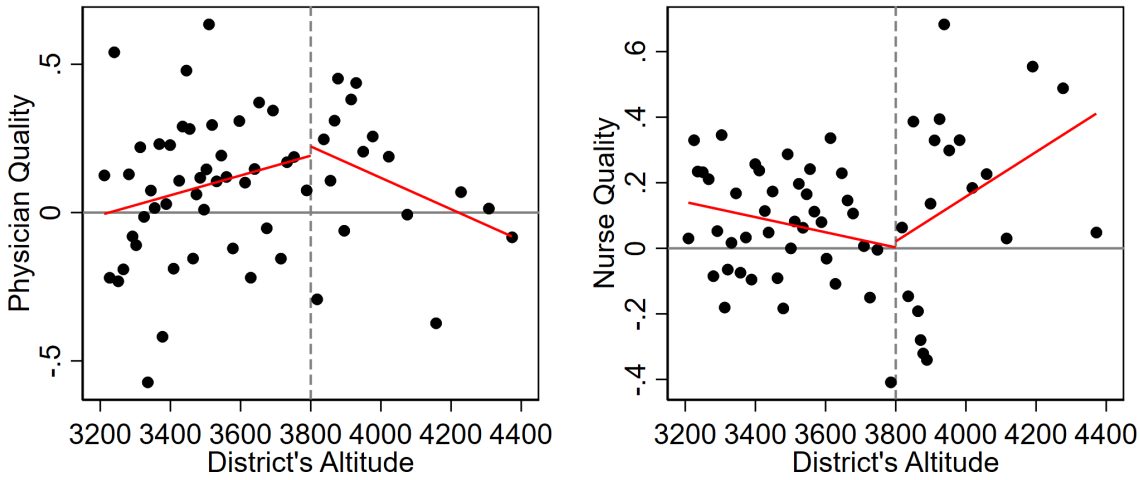


Notes: The figure presents binned scatter plots with fitted lines of physician quality on the left and nurse quality on the right relative to the poverty of the district of the health center they chose. Districts are arranged from the richest on the left to the poorest on the right. Separate lines are produced for health centers in the poorest district quintile, second poorest district quintile, and others. Quintiles are population-weighted.

Figure 4 uses a similar approach but plots the standardized scores of physicians and nurses against the elevation of their selected SERUMS location. The patterns we observe are almost opposite to those previously described for poverty: we observe an increase in the quality of nurses at the threshold for the application of the monetary incentives reform,

but a (small) decrease in the quality of physicians. Once more, these figures shouldn't be interpreted causally, but motivate a more formal analysis in the next sections.

Figure 4: Relation between Quality and Elevation



Notes: The figure presents binned scatter plots with fitted lines of physician quality on the left and nurse quality on the right relative to the elevation of the health center they chose. Health centers are arranged from the lowest elevation on the left to the highest on the right. Separate lines are produced for health centers below and above the 3,800-meter elevation threshold for monetary incentives. Health centers further than 100km from the 3,800-meter contour line or within 30km from the country's border are excluded.

4.2 Identification Strategy

4.2.1 Non-Monetary Incentives

We first estimate the effect that the introduction of non-monetary incentives had on the quality and quantity of the attracted health providers and on health-related outcomes. An ideal experiment would provide non-monetary incentives to randomly chosen districts and compare the outcome of those areas against that of other locations. However, incentives are provided to a selected group of districts by the 2008 reform, with stronger incentives for the poorer ones.

To estimate the effect of non-monetary incentives, we use a regression discontinuity design that leverages two discontinuities in the incentive schedule.[16] The first discontinuity is

[16]The discontinuities around the thresholds between the first and second quintiles and the second and third quintiles have a much smaller number of observations around them. This is due to the fact that quintiles are

between the third and the fourth poverty quintile: choosing a health center in a poorer area doubles the number of points awarded for public jobs from 5 to 10 and increases from 6 to 8 the number of points awarded for residency exams. We compare health centers located in districts with a poverty score just below the threshold of 0.1507 to those just above while controlling for the poverty score. To increase power, we pool the outcomes of the SERUMS allocation process in different regions and different years and control for related fixed effects. The estimating equation is:

$$y_{dt} = (\beta^{[3-4]} + \alpha_1^{[3-4]} poverty_d)\mathbf{1}[poverty_d \geq 0.1507] + \alpha_2^{[3-4]} X_{dt} + \varepsilon_{dt}^{[3-4]} \quad (1)$$

where y_{dt} is the selected outcome for district d at time t, $poverty_d$ is the 2006 poverty score of the district, and X_{dt} are controls that include $poverty_d$, and region-specific time fixed effects.[17] Following Gelman and Imbens (2019), we control for $poverty_d$ linearly above and below the threshold to avoid overfitting. $\beta^{[3-4]}$ is the main parameter of interest, estimating the local average treatment effect of the non-monetary incentives on the average quality of health professionals for districts near the threshold. We restrict our sample to the districts with poverty scores near the threshold, i.e. those in the interval defined by 0.1507 ± 0.09.[18]

The second discontinuity that we study is that between the fourth and fifth poverty score quintiles: choosing a health center in a poorer area further increases the number of points awarded for public jobs by 5 (from 10) and increases from 8 to 10 the number of points awarded for residency exams. We use the same approach described above to study the change around the threshold of 0.5911. For consistency, we also restrict this sample to the districts with poverty scores in the interval defined by 0.5911 ± 0.09. Our outcome is, again, the average standardized SERUMS score of the health professionals who chose to work in the district d at time t. The estimating equation is:

population-weighted, and denser areas tend to be less poor. This results in only 48 districts in the richest quintile, 55 in the second quintile, 229 in the third, 689 in the fourth, and 811 in the fifth and poorest. This can be observed in Appendix Figure 6.

[17] The relevant region is determined by the health directorate that allocates providers to a given health center.

[18] We choose this bandwidth because the threshold between the second and third quintile is at 0.0557. A larger bandwidth would include health centers in the second poverty quintile, where non-monetary incentives are even lower, and a smaller bandwidth would reduce our precision. In the Appendix, we show that the choice of bandwidth does not drive our findings.

$$y_{dt} = (\beta^{[4-5]} + \alpha_1^{[4-5]} poverty_d)\mathbf{1}[poverty_d \geq 0.5911] + \alpha_2^{[4-5]} X_{dt} + \varepsilon_{dt}^{[4-5]} \quad (2)$$

where variables have similar interpretations to those presented above.

Both models rely on standard Regression Discontinuity Design assumptions. We need the error to be uncorrelated with our discontinuity indicator after conditioning on the running variable (*poverty*) and other controls. Intuitively, health centers just below each threshold should work as an appropriate control group for those just above it that received the incentive treatment.

4.2.2 Monetary Incentives

We now estimate the effect of the introduction of monetary incentives on the quality and quantity of the attracted health providers and on health-related outcomes. Once again, an ideal experiment would provide monetary incentives to randomly chosen health centers and compare the outcome for those centers against that of the other locations. However, incentives are provided to a selected group of health centers by the 2014 reform, with stronger incentives for those in higher elevation areas.

To estimate the effect of monetary incentives, we use a regression discontinuity design similar to that used for non-monetary incentives that was described previously. The discontinuity that we leverage is that introduced by the 2014 reform at 3,800 meters of elevation: professionals that selected health centers above this threshold for their service would receive a bonus between 32 and 49% of their base pay. We then compare the average quality of professionals that selected health centers with elevations just above and just below 3,800 meters. Appendix Figure 7 shows the relevant threshold, i.e., the contour line. Again, we pool together data from different years and different regions to increase power and control for the related fixed effects. The estimating equation is:

$$y_{ht} = (\beta^L + \alpha_1^L elevation_h)\mathbf{1}[elevation_h \geq 3800] + \alpha_2^L X_{ht} + \varepsilon_{ht}^L \quad (3)$$

where y_{ht} is the selected outcome for health center h at time t,[19] $elevation_h$ is the elevation

[19]Notice that, while our prior analysis used district as the cross-sectional unit dimension, we now use health centers as units. This is because elevation is defined at the health center level and not at the district level like the

measured at the health center, and X_{ht} are controls that include $elevation_h$, and region-specific time fixed effects. To account for the presence of non-monetary incentives and increase precision, X_{ht} includes binary indicators for each poverty quintile. Following Gelman and Imbens (2019), we control for $elevation_h$ linearly above and below the 3,800 meters threshold. β^L is the main parameter of interest, estimating the local average treatment effect of the monetary incentives on the average quality of health professionals for health centers near the threshold. We restrict our sample to the health centers with an elevation near the threshold, i.e., those in the interval defined by $3,800 \pm 200$ meters.

4.3 Bandwidth, Sample Selection, and Aggregation

For the analysis of career (non-monetary) incentives, we use data for the years 2009 to 2013. This is consistent with the fact that both 2008 and 2014 were years with changes to SERUMS's incentive schemes. The exclusion of the years following the second reform allows us to reduce concerns that the second policy might have attenuated the impacts of the first. For similar reasons, we use data for 2015-2018 for the analysis of the second reform, which introduced monetary incentives.

When analyzing the impact of the 2014 reform, we further restrict our sample to those areas within 30km from the 3,800 meters contour line (i.e. we select municipalities near the place where elevation crosses the incentive threshold). This is particularly important as Peru has a large and populous high plain with an average elevation of around 3,750 meters, and we wouldn't expect balance or continuity when including non-proximate areas. We also exclude districts within 30km from the border with other countries as a specific monetary incentive was introduced for border-adjacent districts.

Finally, we aggregate our data at the level of meaningful variation in the running variable. This means that, when analyzing career incentives, the unit of analysis is the district, while individual health centers are the unit of analysis for monetary incentive regressions. Aggregation level is discussed in the notes of each table. Consistently with these choices and the fact that the units of analysis are observed multiple times in the period studied, we cluster standard errors at the unit level. We provide results using alternative aggregation in

poverty score. The Appendix reports results for all discontinuities using different aggregations.

the Appendix and discuss them among the robustness checks.

5 Results

In this section, we study the impact of non-monetary incentives introduced by the 2008 reform and of monetary incentives introduced by the 2014 reform on the quality of health providers attracted by the affected health centers.

5.1 Impact of Incentives on Providers' Quality and Quantity

First, we report the impact of non-monetary incentives on the average score of the health professionals attracted by treated districts. Table 4 reports our estimates for the causal parameter of interest β for models 1, 2, and 3. We can observe that an increase in non-monetary incentives had a positive and significant impact for physicians, with an estimate of about half a standard deviation at each of the two thresholds examined (columns 1 and 3). This clearly indicates that non-monetary incentives were able to attract physicians to poorer areas, as suggested by the descriptive evidence previously presented. However, columns 2 and 4 of the same table show that those same incentives didn't significantly affect the quantity of physicians in the treated areas. Similarly, we don't find smaller, non-significant impact from monetary incentives (columns 5 and 6).

The situation reverses for nurses: columns 1 to 4 of Table 5 show that the estimated effects of non-monetary incentives are both smaller than those for physicians and not statistically significant. Conversely, we find significant increases in the number of nurses choosing high-elevation locations ($p < 0.05$) and higher average quality ($p < 0.1$). This suggests that nurses reacted much more strongly to monetary than non-monetary incentives.

5.2 Impact of Incentives on Healthcare Utilization and Quality

We now study the impact of incentives on the two chosen outcomes: healthcare utilization in the past 12 months and the "Good Healthcare Index", measuring service quality perceptions of the surveyed women as described in Section 3. Odd columns in Table 6 show that utilization was not impacted by the provision of either monetary or non-monetary incentives. This

Table 4: Effects of Non-Monetary and Monetary Incentives on Doctors

	Non-Monetary 1		Non-Monetary 2		Monetary	
	(1) Quality	(2) Quantity	(3) Quality	(4) Quantity	(5) Quality	(6) Quantity
Estimate	0.573***	-0.0402	0.431**	-0.364	0.240	0.0127
	(0.182)	(0.354)	(0.185)	(0.279)	(0.192)	(0.0318)
N	1215	1215	1080	1080	824	824
Threshold	0.151	0.151	0.591	0.591	3800	3800
Bandwidth	0.09	0.09	0.09	0.09	200	200
Baseline	-.0486	1.8292	-.1270	1.6178	.1260	1.0244

Notes: The table reports estimates using physicians' data for model 1 (columns 1 and 2), model 2 (columns 3 and 4), and model 3 (columns 5 and 6). Each regression controls for fixed effects for time fixed effects in each assignment region. Baseline values are reported for the untreated units within the chosen bandwidth. Standard errors, reported in parentheses, are clustered at the district level for columns 1-4 and at the health center for columns 5-6. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

suggests that the quality increases of physicians for non-monetary incentives and of nurses for monetary incentives didn't increase the utilization of healthcare services. Columns 2, 4, and 6 tell an even more pessimistic story: perceptions of service quality were either unaffected (columns 4 and 6) or significantly decreased.

5.3 Robustness of Results

In Appendix Figures 8, 9, and 10, we show that the results are robust to the choice of bandwidth. As expected, with small bandwidths, confidence intervals tend to be larger, and estimates tend to be more extreme. However, none of our results appears to be driven by a specific bandwidth choice.

We also increase the granularity of the fixed effects used. Including province fixed effects (Peru has 196 provinces), we are effectively comparing units above and below the threshold of interest *within each province*. Unfortunately, this implies a substantial loss of observations contributing to the estimate of the discontinuities. This happens when we can only find treated or untreated districts within a province. Appendix Tables 9, 10, and 11 show similar results, although the impact of monetary incentives on nurse quality loses significance while that on physicians increases and gains significance. Finally, the negative impact on the *Good*

Table 5: Effects of Non-Monetary and Monetary Incentives on Nurses

	Non-Monetary 1		Non-Monetary 2		Monetary	
	(1) Quality	(2) Quantity	(3) Quality	(4) Quantity	(5) Quality	(6) Quantity
Estimate	0.134	-0.355	0.0495	-0.231	0.339*	0.0464**
	(0.169)	(0.281)	(0.153)	(0.266)	(0.175)	(0.0185)
N	1191	1191	972	972	1159	1159
Threshold	0.151	0.151	0.591	0.591	3800	3800
Bandwidth	0.09	0.09	0.09	0.09	200	200
Baseline	.0467	2.2128	.1775	1.5108	-.0055	1.0079

Notes: The table reports estimates using nurses' data for model 1 (columns 1 and 2), model 2 (columns 3 and 4), and model 3 (columns 5 and 6). Each regression controls for fixed effects for time fixed effects in each assignment region. Baseline values are reported for the untreated units within the chosen bandwidth. Standard errors, reported in parentheses, are clustered at the district level for columns 1-4 and at the health center for columns 5-6. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Healthcare Index at the first non-monetary threshold becomes much smaller and insignificant.

As discussed in Section 4.2, the Appendix includes the results using different aggregation levels. We show health center level estimates for models using district as the unit of analysis in the main text, and vice versa. These alternative specifications create a mismatch between the level of variation in incentives and the unit of analysis. Nonetheless, we find that our conclusions on physicians are extremely robust, with strong effects for non-monetary incentives but no impact of monetary incentives (see Appendix Table 12). Appendix Table 13, on the other hand, shows that the increase in nurses' quality and quantity loses significance when aggregating at the district level as expected, as elevation can vary greatly within a district.

Finally, we provide placebo tests where we estimate the impact of policies before their implementation. We use health-related outcomes for both incentive reforms thanks to data from the period 1991 to 2007. We only have measures of provider quality starting in 2008, which we use to evaluate the 2014 reform on placebo outcomes for the period 2009-2013. In Appendix Table 14, we reject the null of no impact during placebo years for only one of five placebo specifications ($p - value < 0.05$). We plan to identify and test more outcomes suitable for placebo exercises.

Table 6: Effects of Non-Monetary and Monetary Incentives on Health Outcomes

	Non-Monetary 1		Non-Monetary 2		Monetary	
	(1) Utilization	(2) Care Quality	(3) Utilization	(4) Care Quality	(5) Utilization	(6) Care Quality
Estimate	0.00379	-0.102***	-0.0310	0.00577	0.00168	0.0404
	(0.0271)	(0.0370)	(0.0297)	(0.0605)	(0.0548)	(0.103)
N	21646	3427	8206	1175	6473	385
Threshold	0.151	0.151	0.591	0.591	3800	3800
Bandwidth	0.09	0.09	0.09	0.09	200	200
Baseline	.4319	.7694	.3851	.6988	.5976	.7163

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports estimates using DHS data for model 1 (columns 1 and 2), model 2 (columns 3 and 4), and model 3 (columns 5 and 6). Each regression controls for fixed effects for time fixed effects in each department. Baseline values are reported for the untreated units within the chosen bandwidth. Standard errors, reported in parentheses, are clustered at the district level. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.4 Provider Characteristics and Incentive Impact

Differences in the underlying characteristics of nurses and physicians might explain their different responses to incentives. To study how incentives affected the gender and ethnic composition of providers at health centers affected by the reforms, we use the same empirical strategy and replicate the analysis using measures of Indigenous ethnicity and gender as our outcomes.

Table 7 shows non-significant estimates for non-monetary and monetary incentives among physicians. Whereas Table 4 reported large and significant estimates for non-monetary incentives on provider quality, these results show that the gender and ethnic composition were not as affected by the reforms. It is worth highlighting, however, that point estimates for our measure of Indigenous ethnicity, and to a lesser extent for the gender measure, are fairly large in column (6), especially relative to a low baseline.

In Table 8 we see that the estimated effects of non-monetary incentives among nurses are also mostly small and not statistically significant. The coefficients are much larger for monetary incentives, with a p-value below 10% in column (5). This suggests that female nurses reacted more strongly than male nurses to monetary incentives. We also see a large but non-significant reduction in the Indigenous composition due to monetary incentives.

Table 7: Effects of Incentives on Gender and Ethnic Composition of Doctors

	Non-Monetary 1		Non-Monetary 2		Monetary	
	(1)	(2)	(3)	(4)	(5)	(6)
	Female	Indigenous	Female	Indigenous	Female	Indigenous
Estimate	0.0157	-0.0368	-0.0313	-0.00563	-0.0498	-0.0797
	(0.0595)	(0.0357)	(0.0467)	(0.0367)	(0.0785)	(0.0584)
N	1215	1215	1080	1080	824	824
Threshold	0.151	0.151	0.591	0.591	3800	3800
Bandwidth	0.09	0.09	0.09	0.09	200	200
Baseline	.491	.135	.434	.135	.502	.137

Notes: The table reports estimates using doctors' data for model 1 (columns 1 and 2), model 2 (columns 3 and 4), and model 3 (columns 5 and 6). Each regression controls for fixed effects for time fixed effects in each assignment region. Baseline values are reported for the untreated units within the chosen bandwidth. Standard errors, reported in parentheses, are clustered at the district level for columns 1-4 and at the health center for columns 5-6. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note that these results are informative about the heterogeneity of responses. If, for example, women were more reluctant to work in remote areas, we would expect them to have a weaker response to incentives and a decrease in the share of women at the examined thresholds since men would "crowd out" women by choosing the incentivized areas over others. As an alternative example, if women were generally more responsive to monetary incentives, we would expect to find an increase in the share of women at the threshold where monetary incentives were introduced.

5.5 Limitations

Our results rely on key counterfactual assumptions: that, absent the introduction of the reforms, the outcomes of interest would have been continuous at the studied thresholds. This assumption cannot be directly tested. One main threat to its credibility comes from the possibility that other policies might have been introduced that had similar or contrary effects to those we're studying. For example, the assumption would be violated if the poverty map produced by FONCODES was used to allocate funds differentially to areas in different quintiles of poverty, and those funds affected providers' location choices.

Two main policies leveraged the same poverty map that provides the running variable used to study career incentives. *Juntos* is a conditional cash transfer (CCT) program consist-

Table 8: Effects of Incentives on Gender and Ethnic Composition of Nurses

	Non-Monetary 1		Non-Monetary 2		Monetary	
	(1) Female	(2) Indigenous	(3) Female	(4) Indigenous	(5) Female	(6) Indigenous
Estimate	-0.00284	0.00812	-0.0294	0.0307	0.107*	-0.0727
	(0.0435)	(0.0448)	(0.0549)	(0.0494)	(0.0639)	(0.0624)
N	1191	1191	972	972	1159	1159
Threshold	0.151	0.151	0.591	0.591	3800	3800
Bandwidth	0.09	0.09	0.09	0.09	200	200
Baseline	.815	.140	.745	.237	.758	.215

Notes: The table reports estimates using nurses' data for model 1 (columns 1 and 2), model 2 (columns 3 and 4), and model 3 (columns 5 and 6). Each regression controls for fixed effects for time fixed effects in each assignment region. Baseline values are reported for the untreated units within the chosen bandwidth. Standard errors, reported in parentheses, are clustered at the district level for columns 1-4 and at the health center for columns 5-6. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

ing of a roughly $30 monthly payment (approximately 20% of the average family income). This transfer is conditional on the children having 85% school attendance, complete immunization, and pre- and post-natal care. The rollout of the program was based on an index that includes poverty and the share of civil conflict-affected villages in the district. This index used for rollout did not present discontinuities.

The second policy using the same poverty map is the redistribution of the Canon Minero, a tax on profits coming from mineral extraction. This is distributed to districts according to their production, demographic composition, and the 2006 poverty measure. The distribution shares do not change discontinuously. We are not aware of any other major policy relying on the 2006 poverty map nor of any policy using 3,800 meters of elevation as a cutoff.

A limitation of our work is the lack of information on the effort exerted by the providers. Indeed, one potential reason for the lack of impact on healthcare outcomes is that the two reforms improved selection but did nothing to improve worker monitoring. Better workers exerting low effort might have similar health impacts as any other worker exerting low effort.

Although differences in physicians' and nurses' reactions to incentives provide somewhat independent variation in quality at different thresholds, we can't definitively disentangle the impact of (better) physicians from the impact of (better) nurses.

Finally, our analysis treats the supply of providers as near-infinite. The increase in the

average quality of physicians or nurses comes from the arrival of high-scoring providers to the targeted areas. However, the supply of providers is fixed in the short term, and those high-scoring providers are being diverted from other locations. This diversion has negligible effects if it happens uniformly (i.e. the providers going from the targeted areas wouldn't have concentrated in specific locations absent the incentives).

6 Discussion and Conclusions

We examined two incentive reforms within Peru's SERUMS program aimed at attracting more and better physicians and nurses to poorer and more remote areas, with the final goal of improving healthcare access and service quality. While both career incentives and monetary bonuses were at least partially successful in increasing healthcare provider quality and quantity in underserved areas, their effects varied across physician and nurse cohorts. Notably, despite these improvements, we observe no significant changes in healthcare utilization or perceived service quality, indicating the limitations of these incentive schemes in addressing broader healthcare provision challenges.

The differing responses of physicians and nurses to incentive structures underscore the importance of carefully considering the backgrounds, preferences, and career trajectories of healthcare professionals. While career incentives primarily attracted physicians, monetary bonuses were more effective in incentivizing nurses. This understanding of provider responsiveness to different incentive levers can inform future policy design and be used to better address inequalities in healthcare quality and availability.

What drives the differences between physicians and nurses in the reaction to different incentive types? A possible answer comes from the observation that physicians have a higher base pay and steeper wage growth during their careers, while nurses are paid a lower wage and have fewer growth prospects. This means that a "career boost" is much more valuable to a physician than to a nurse. At the same time, nurses might have higher marginal utility from additional pay due to their lower baseline pay. Additional factors, such as differences in the gender composition of the two professions and heterogeneous preferences, are not satisfactory. As we have seen in Section 5.4, the gender and ethnic composition of providers did not change significantly as a consequence of the incentive introduction. Additionally, a

higher share of women among nurses and a stronger reluctance of women to work in remote locations would not explain why monetary incentives successfully attracted them to high-elevation, remote areas.

The lack of effects on the selected health outcomes also deserves a discussion. Several papers have highlighted the potential for and found evidence of a trade-off between prosociality and talent. Additionally, the lack of monitoring might have affected the effectiveness of increases in provider quality on service quality. Although we don't find statistically significant impacts on the ethnic composition of providers, cultural or linguistic mismatch between health providers and the local population might have also caused the negative impact found in Table 6.

Our findings highlight the need for a deeper understanding of the relationship between the availability of highly qualified health professionals, utilization of healthcare services, and perception of service quality. The possibility of patient preferences over dimensions other than quality as measured by test scores and the discretionality of professionals over their amount of effort suggest that providing incentives might be introducing a trade-off between quality and other desirable characteristics, such as prosociality, or cultural and linguistic match with local patients.

References

Agte, P. and Bedoya, M. (2023). The making of a public sector worker: The causal effects of temporary work assignments to poor areas.

Ansah, E. K., Narh-Bana, S., Asiamah, S., Dzordzordzi, V., Biantey, K., Dickson, K., Gyapong, J. O., Koram, K. A., Greenwood, B. M., Mills, A., et al. (2009). Effect of removing direct payment for health care on utilisation and health outcomes in ghanaian children: a randomised controlled trial. *PLoS medicine*, 6(1):e1000007.

Ashraf, N., Bandiera, O., Davenport, E., and Lee, S. S. (2020). Losing prosociality in the quest for talent? sorting, selection, and productivity in the delivery of public services. *American Economic Review*, 110(5):1355–1394.

Bernal, N., Carpio, M. A., and Klein, T. J. (2017). The effects of access to health insurance: evidence from a regression discontinuity design in peru. *Journal of Public Economics*, 154:122–136.

Bobba, M., Ederer, T., Leon-Ciliotta, G., Neilson, C., and Nieddu, M. G. (2021). Teacher compensation and structural inequality: Evidence from centralized teacher school choice in perú. Technical report, National Bureau of Economic Research.

Chari, A. V. and Okeke, E. N. (2014). Can institutional deliveries reduce newborn mortality? evidence from rwanda.

Dal Bó, E., Finan, F., and Rossi, M. A. (2013). Strengthening state capabilities: The role of financial incentives in the call to public service. *The Quarterly Journal of Economics*, 128(3):1169–1218.

Das, J. and Hammer, J. (2014). Quality of primary care in low-income countries: facts and economics. *Annu. Rev. Econ.*, 6(1):525–553.

Das, J., Woskie, L., Rajbhandari, R., Abbasi, K., and Jha, A. (2018). Rethinking assumptions about delivery of healthcare: implications for universal health coverage. *Bmj*, 361.

Deserranno, E. (2019). Financial incentives as signals: experimental evidence from the recruitment of village promoters in uganda. *American Economic Journal: Applied Economics*, 11(1):277–317.

Díaz, J. (2006). Nuevo Mapa de Pobreza del Fondo de Cooperación para el Desarrollo Social (FONCODES) 2006. Technical report, Fondo de Cooperación para el Desarrollo - FONCODES.

Finan, F., Olken, B. A., and Pande, R. (2015). The personnel economics of the state.

Frehywot, S., Mullan, F., Payne, P. W., and Ross, H. (2010). Compulsory Service Programmes for Recruiting Health Workers in Remote and Rural Areas: Do They Work? *Bulletin of the World Health Organization*, 88(5):364–370.

Friedrich, B. U. and Hackmann, M. B. (2021). The returns to nursing: Evidence from a parental-leave program. *The Review of Economic Studies*, 88(5):2308–2343.

Gelman, A. and Imbens, G. (2019). Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics*, 37(3):447–456.

Godlonton, S. and Okeke, E. N. (2016). Does a ban on informal health providers save lives? evidence from malawi. *Journal of development economics*, 118:112–132.

Huicho, L., Miranda, J. J., Diez-Canseco, F., Lema, C., Lescano, A. G., Lagarde, M., and Blaauw, D. (2012). Job preferences of nurses and midwives for taking up a rural job in peru: a discrete choice experiment. *PloS one*, 7(12):e50315.

Huicho, L., Molina, C., Diez-Canseco, F., Lema, C., Miranda, J. J., Huayanay-Espinoza, C. A., and Lescano, A. G. (2015). Factors behind job preferences of peruvian medical, nursing and midwifery students: a qualitative study focused on rural deployment. *Human resources for health*, 13:1–11.

Jimenez, M. M., Bui, A. L., Mantilla, E., and Miranda, J. J. (2017). Human resources for health in peru: recent trends (2007–2013) in the labour market for physicians, nurses and midwives. *Human Resources for Health*, 15:1–7.

Kruk, M. E., Yamey, G., Angell, S. Y., Beith, A., Cotlear, D., Guanais, F., Jacobs, L., Saxenian, H., Victora, C., and Goosby, E. (2016). Transforming global health by improving the science of scale-up. *PLoS biology*, 14(3):e1002360.

Lagarde, M., Huicho, L., and Papanicolas, I. (2019). Motivating provision of high quality care: it is not all about the money. *bmj*, 366.

Observatorio Nacional de Recursos Humanos en Salud (2009). Experiencias de Planificación de los Recursos Humanos en Salud Perú, 2007-2010. Technical report, Ministerio de Salud.

Okeke, E. N. (2023). When a doctor falls from the sky: The impact of easing doctor supply constraints on mortality. *American economic review*, 113(3):585–627.

Posso, C., Tamayo, J., Guarin, A., and Saravia, E. (2024). Luck of the draw: The causal effect of physicians on birth outcomes. *Borradores de Economía; No. 1269*.

Powell-Jackson, T., Mazumdar, S., and Mills, A. (2015). Financial incentives in health: new evidence from india's janani suraksha yojana. *Journal of health economics*, 43:154–169.

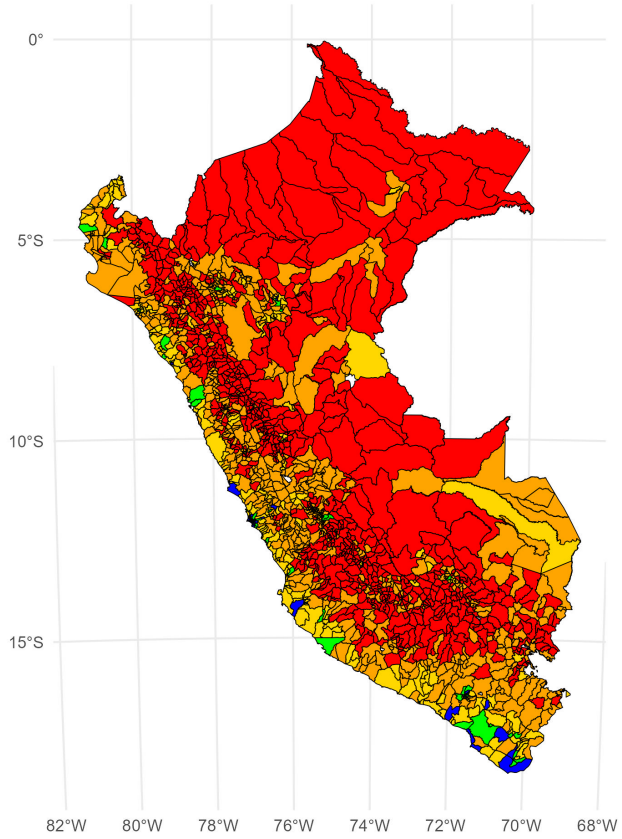
Weiss, D., Nelson, A., Vargas-Ruiz, C., Gligorić, K., Bavadekar, S., Gabrilovich, E., Bertozzi-Villa, A., Rozier, J., Gibson, H., Shekel, T., et al. (2020). Global maps of travel time to healthcare facilities. *Nature Medicine*, 26(12):1835–1838.

7 Additional Figures and Tables

Figure 5: Countries with Civil Service Programs for Health Providers

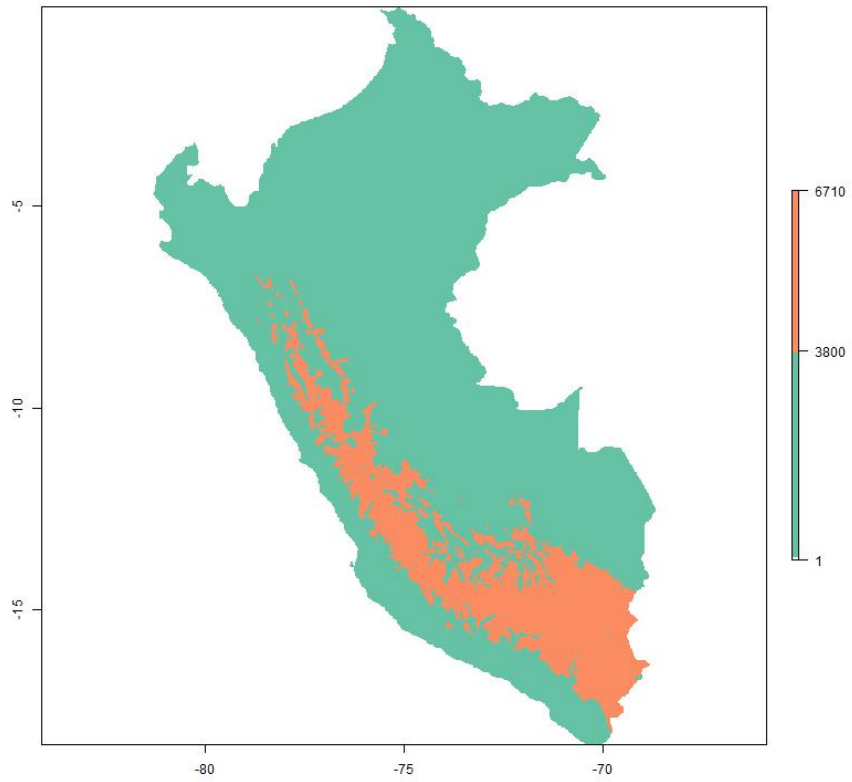
Notes: Countries with civil service programs for health providers similar to SERUMS. Image from Frehywot et al. (2010).

Figure 6: Map of Peru's Districts by Poverty Score Quintile



Notes: Map of the district categorized according to the quintile of the poverty score calculated in the poverty map created by FONCODES. Red districts are in the poorest quintile, orange districts are in the second poorest, and yellow districts are in the third poorest.

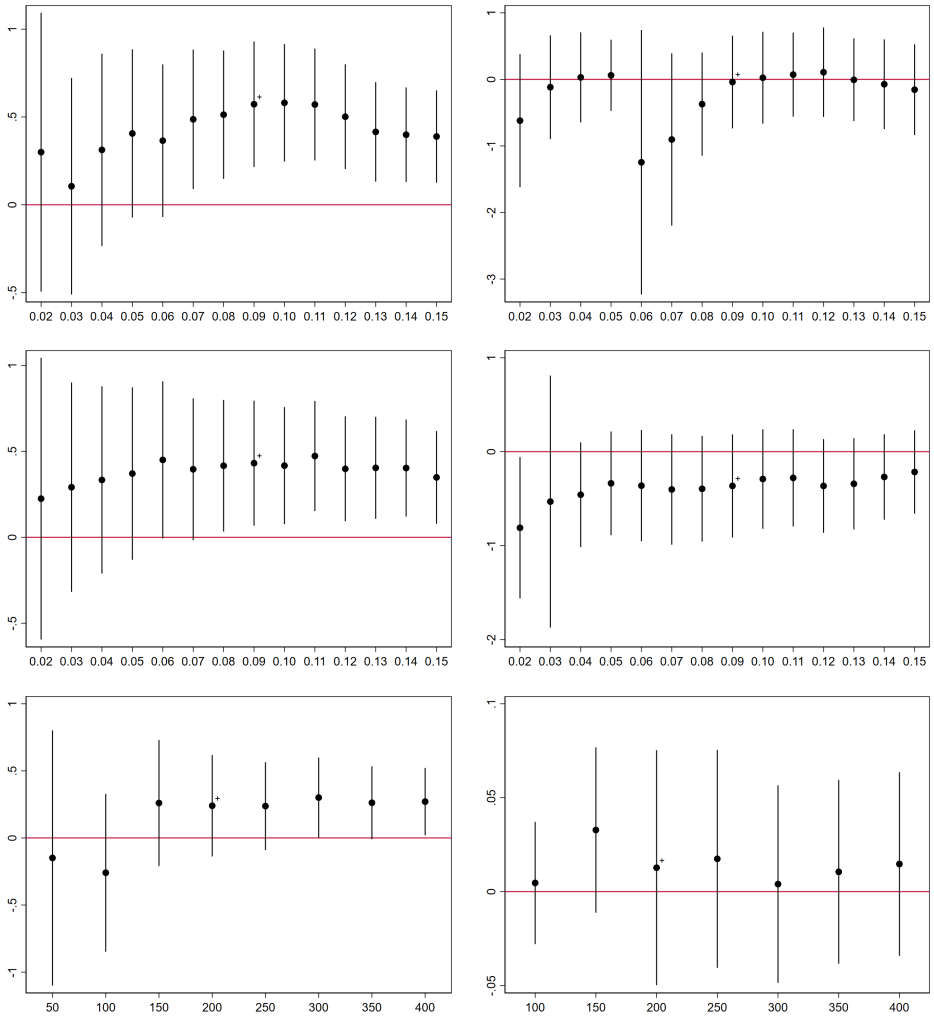
Figure 7: Map of Peru and 3,800 MSL contour line



Notes: Map of Peru. Orange areas are above 3,800 MSL.

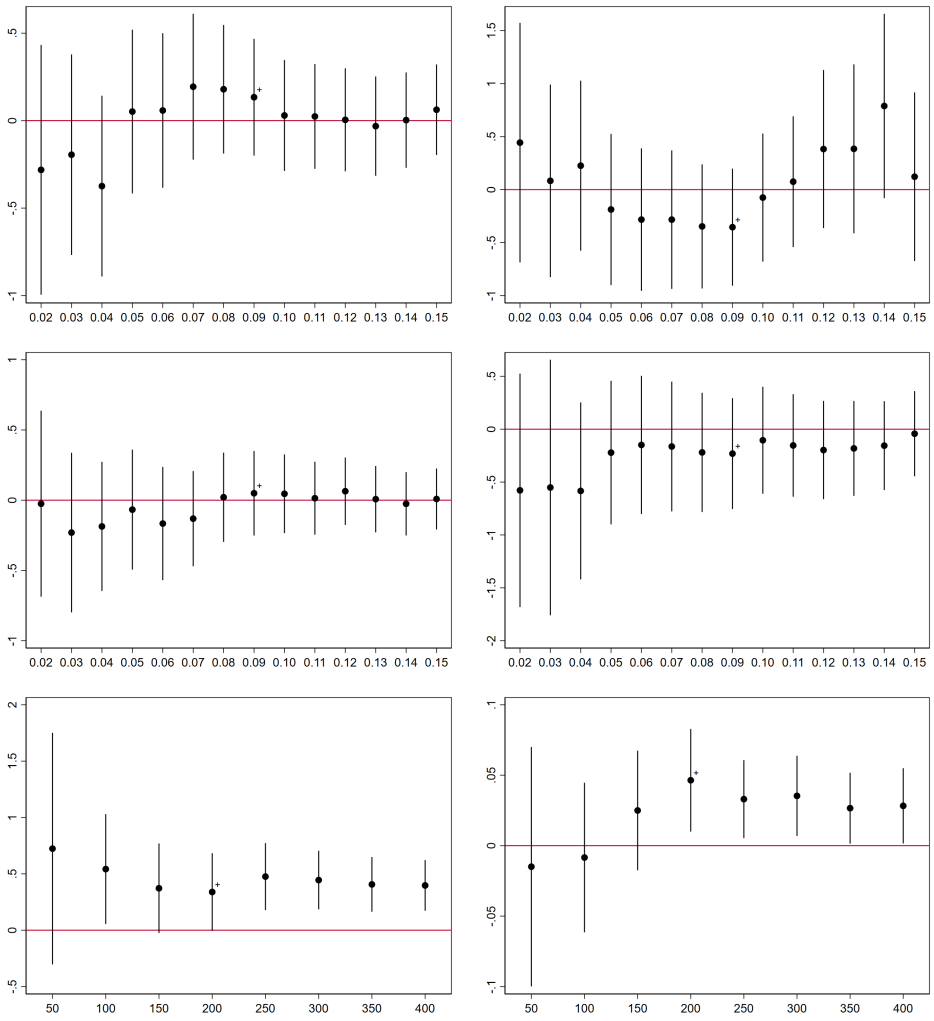
7.1 Robustness of Bandwidth Choice

Figure 8: Bandwidth Robustness of Table 4



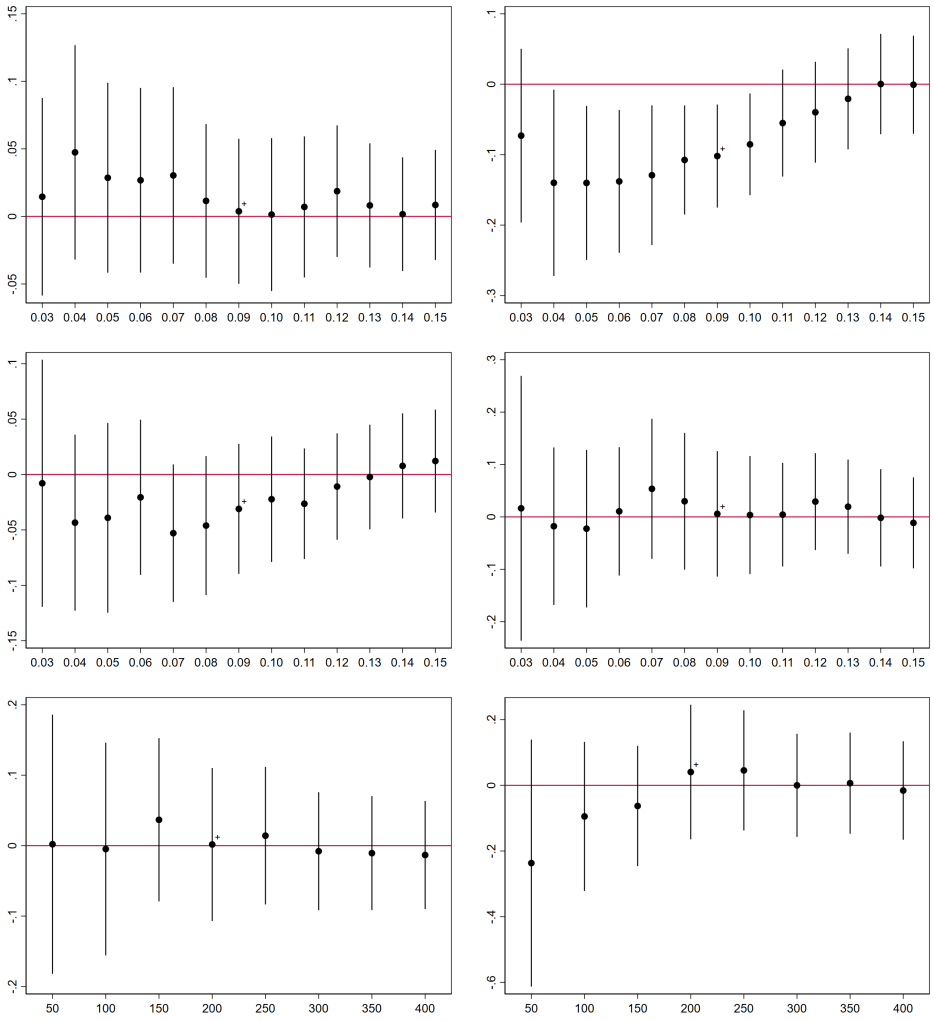
Notes: Replication of Table 4 using different bandwidths. Specifications are the same as the replicated table. The sign "+" indicates the bandwidth presented in the original table. Left column shows models with quality as an outcome, right column shows models with quantity as an outcome; first row shows the estimates for the first non-monetary threshold, second row for the second non-monetary threshold, third row for the the monetary threshold.

Figure 9: Bandwidth Robustness of Table 5



Notes: Replication of Table 4 using different bandwidths. Specifications are the same as the replicated table. The sign "+" indicates the bandwidth presented in the original table. Left column shows models with quality as an outcome, right column shows models with quantity as an outcome; first row shows the estimates for the first non-monetary threshold, second row for the second non-monetary threshold, third row for the the monetary threshold.

Figure 10: Bandwidth Robustness of Table 6



Notes: Replication of Table 4 using different bandwidths. Specifications are the same as the replicated table. The sign "+" indicates the bandwidth presented in the original table. Left column shows models with healthcare utilization as an outcome, right column shows models with healthcare quality perceptions (Good Healthcare Index) as an outcome; first row shows the estimates for the first non-monetary threshold, second row for the second non-monetary threshold, third row for the the monetary threshold.

7.2 Robustness Including Province Fixed Effects

Table 9: Effects of Non-Monetary and Monetary Incentives on Doctors

	Non-Monetary 1		Non-Monetary 2		Monetary	
	(1) Quality	(2) Quantity	(3) Quality	(4) Quantity	(5) Quality	(6) Quantity
Estimate	0.473**	0.441	0.649***	-0.108	0.518***	0.0239
	(0.198)	(0.487)	(0.188)	(0.351)	(0.191)	(0.0320)
N	1210	1210	1078	1078	821	821
Threshold	0.151	0.151	0.591	0.591	3800	3800
Bandwidth	0.09	0.09	0.09	0.09	200	200
Baseline	-.04958	1.83128	-.12436	1.61915	.12590	1.02463

Notes: The table reports estimates using physicians' data for model 1 (columns 1 and 2), model 2 (columns 3 and 4), and model 3 (columns 5 and 6). Each regression controls for fixed effects for time fixed effects in each assignment region and fixed effects for each province. Baseline values are reported for the untreated units within the chosen bandwidth. Standard errors, reported in parentheses, are clustered at the district level for columns 1-4 and at the health center for columns 5-6. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Effects of Non-Monetary and Monetary Incentives on Nurses

	Non-Monetary 1		Non-Monetary 2		Monetary	
	(1) Quality	(2) Quantity	(3) Quality	(4) Quantity	(5) Quality	(6) Quantity
Estimate	0.206	-0.345	0.0170	-0.0231	0.0388	0.0622***
	(0.187)	(0.298)	(0.154)	(0.227)	(0.186)	(0.0224)
N	1190	1190	967	967	1157	1157
Threshold	0.151	0.151	0.591	0.591	3800	3800
Bandwidth	0.09	0.09	0.09	0.09	200	200
Baseline	.04672209	2.2128174	.17513909	1.5144928	-.00753227	1.0079051

Notes: The table reports estimates using nurses' data for model 1 (columns 1 and 2), model 2 (columns 3 and 4), and model 3 (columns 5 and 6). Each regression controls for fixed effects for time fixed effects in each assignment region and fixed effects for each province. Baseline values are reported for the untreated units within the chosen bandwidth. Standard errors, reported in parentheses, are clustered at the district level for columns 1-4 and at the health center for columns 5-6. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Effects of Non-Monetary and Monetary Incentives on Health Outcomes

	Non-Monetary 1		Non-Monetary 2		Monetary	
	(1) Utilization	(2) Care Quality	(3) Utilization	(4) Care Quality	(5) Utilization	(6) Care Quality
Estimate	-0.0361	-0.0335	0.0153	-0.0227	-0.0126	0.0295
	(0.0321)	(0.0424)	(0.0374)	(0.116)	(0.0548)	(0.128)
N	21646	3426	8206	1172	6473	379
Threshold	0.151	0.151	0.591	0.591	3800	3800
Bandwidth	0.09	0.09	0.09	0.09	200	200
Baseline	.43187	.76957	.38508	.69890	.59761	.71569

Notes: The table reports estimates using DHS data for model 1 (columns 1 and 2), model 2 (columns 3 and 4), and model 3 (columns 5 and 6). Each regression controls for fixed effects for time fixed effects in each department. Baseline values are reported for the untreated units within the chosen bandwidth. Standard errors, reported in parentheses, are clustered at the district level. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

7.3 Robustness with Alternative Aggregation

Table 12: Effects of Non-Monetary and Monetary Incentives on Doctors

	Non-Monetary 1		Non-Monetary 2		Monetary	
	(1) Quality	(2) Quantity	(3) Quality	(4) Quantity	(5) Quality	(6) Quantity
Estimate	0.655***	0.177***	0.454**	-0730	0.0136	-0.527
	(0.188)	(0.0649)	(0.198)	(0.0190)	(0.187)	(0.358)
N	1911	1911	1656	1656	564	564
Threshold	0.151	0.151	0.591	0.591	3800	3800
Bandwidth	0.09	0.09	0.09	0.09	200	200
Baseline	-.09316	1.23153	-.13309	1.01465	.10957	1.63504

Notes: The table reports estimates using physicians' data for model 1 (columns 1 and 2), model 2 (columns 3 and 4), and model 3 (columns 5 and 6). Each regression controls for fixed effects for time fixed effects in each assignment region. Baseline values are reported for the untreated units within the chosen bandwidth. Average quality and quantity are obtained for each health center for columns 1-4 and for each district for columns 5 and 6. Standard errors, reported in parentheses, are clustered at the district level for columns 1-4 and at the health center for columns 5-6. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Effects of Non-Monetary and Monetary Incentives on Nurses

	Non-Monetary 1		Non-Monetary 2		Monetary	
	(1) Quality	(2) Quantity	(3) Quality	(4) Quantity	(5) Quality	(6) Quantity
Estimate	0.164	0.0493	0.0371	-0206	0.0519	0.257
	(0.146)	(0.0564)	(0.151)	(0.0146)	(0.201)	(0.328)
N	2202	2202	1494	1494	812	812
Threshold	0.151	0.151	0.591	0.591	3800	3800
Bandwidth	0.09	0.09	0.09	0.09	200	200
Baseline	-.039125	1.22773	.152010	1.02107	.08091	1.5

Notes: The table reports estimates using nurses' data for model 1 (columns 1 and 2), model 2 (columns 3 and 4), and model 3 (columns 5 and 6). Each regression controls for fixed effects for time fixed effects in each assignment region. Baseline values are reported for the untreated units within the chosen bandwidth. Average quality and quantity are obtained for each health center for columns 1-4 and for each district for columns 5 and 6. Standard errors, reported in parentheses, are clustered at the district level for columns 1-4 and at the health center for columns 5-6. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

7.4 Placebo Estimates

Table 14: Effects of Non-Monetary and Monetary Incentives in Placebo Years

	(1) Utilization	(2) Utilization	(3) Utilization	(4) Quality	(5) Quality
Non-Monetary	0.0714	0.138**			
	(0.0434)	(0.0531)			
Monetary			-0.0131	0.00746	0.0978
			(0.0425)	(0.130)	(0.202)
N	16619	5022	4778	762	801
Threshold	0.151	0.591	3800.000	3800.000	3800.000
Bandwidth	0.09	0.09	200.00	200.00	200.00
Baseline	0.46684	0.43799	0.39609	-0.00779	0.06048
Sample	Health	Health	Health	Doctor	Nurse
Years	1991-2007	1991-2007	1991-2007	2009-2013	2009-2013

Notes: The table reports estimates using years in which incentives were not present. Column 1 is a placebo test for the first non-monetary threshold (model 1), while column 2 is a test for the second (model 2). Columns 3-5 are placebo tests for the monetary incentives (model 3). Controls are the same as the main non-placebo specifications. Baseline values are reported for the untreated units within the chosen bandwidth. Standard errors, reported in parentheses, are clustered at the district level for columns 1 and 2 and at the health center for columns 3-5. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$