Incentives, Management, and Public Service Delivery: Evidence from a large field experiment in Peru*

Sarita Oré-Quispe

October 21st, 2026

Click here for the most recent version

Abstract

Preliminary findings, please do not circulate or cite without the author's permission.

Should governments reward local service delivery or the outcomes these services aim to achieve? I study this question in Peru's municipal home visit program, in which volunteer community health workers (CHWs) visit families to promote preventive care for infants. In close collaboration with the central government, I randomly assigned 539 rural municipalities, 85% of the national total, to one of two contracts. The *Delivery-only* contract rewards home visit coverage, while the *Delivery and Output* contract also rewards a child anemia indicator. Both contracts pay the same amount. Because CHWs are volunteers, any response must work through how municipal managers organize, monitor, and motivate them. Adding the output indicator did not raise coverage, but it changed managerial practices. Managers tracked anemia more closely, coordinated referrals with health facilities, and rewarded CHWs with non-monetary recognition. Although the prevalence of anemia did not ultimately improve, children made more visits to health facilities, received more vaccines, and their families doubled the amount of iron supplements picked up for the child. The managerial changes observed persist even after all municipalities reverted to the *Delivery-only* contract.

Keywords: Performance-based transfers, Managerial practices, Public service delivery, Community health workers, Child health.
JEL Classification: H77, D73, M54, I18, I15, C93.

*Teachers College, Columbia University. Email: spo2111@tc.columbia.edu. I appreciate the valuable feedback from Alex Eble, Cristóbal Otero, Micaela Sviatschi, Samuel Berlinski, Pascaline Dupas, Seema Jayachandran, Carolina Concha-Arriagada, Eric Verhoogen, Michael Best, Gianmarco León-Ciliotta, Stefano Caria, Alejandra Ramos, Stanislao Maldonado, Pablo Celhay and participants at the Columbia University and Princeton University development colloquiums. Special thanks to the Peruvian Ministry of Economy and Finance and Ministry of Health for their support in the ongoing implementation and data sharing. I gratefully acknowledge financial support from The Weiss Fund, the CAF—Development Bank of Latin America and the Caribbean, J-PAL's Governance Initiative, and Columbia University. This study has received approval from the Teachers College Institutional Review Board (IRB) and the protocol ID is 24-196. This study is registered in the AEA RCT Registry and the unique identifying number is: AEARCTR-0012869. All errors remain my own.

1 Introduction

Subnational governments deliver many essential public services, including education, health, water, and sanitation. Worldwide, they account for 21.5% of public expenditure and 8.3% of GDP (OECD, 2022). Central government transfers fund more than half of their revenues (OECD, 2022). Transfer design is therefore one of the main levers the center has to shape the quantity and quality of local service provision. That recognition has pushed governments increasingly toward performance-based transfer schemes.[1] These schemes must choose what to reward. Service delivery indicators, such as visit coverage or teacher attendance, track tasks that local governments control. They are easy to verify and are observable as soon as the service is delivered. The downside is that they often reward only task completion, not the quality of the service delivered (Holmström et al., 1991; Mohanan et al., 2021). On the other hand, output indicators, such as child health or student learning, tie the reward directly to the central government's goal. They let local governments innovate, use local information, and choose how to reach that goal. Outputs, however, are noisier signals of effort and often take time before they are observable (Celhay et al., 2019; Grant Miller et al., 2020). This trade-off raises a central question for the efficiency of public spending: Should the central government reward service delivery or the outputs the transfer is meant to improve, or both?

In this paper, I address this question in the context of a home visit program that promotes preventive health care and child health in rural Peru. First, I examine whether adding an output indicator to a service delivery contract changes managerial practices, service delivery, and child health. Second, I examine whether these changes persist once the output indicator is removed. Rural Peru is a natural setting for these questions. In 2023, only 32.8% of rural infants under 12 months completed their scheduled checkups, and 60.5% of those aged 6 to 12 months were anemic (INEI, 2023b). Iron deficiency is the leading cause of childhood anemia, and the main policy response is iron supplements that families pick up at health facilities (Chong et al., 2016; WHO, 2016; WHO, 2023). To close these gaps, since 2019 the *Incentive Program for Municipal Management Improvement* (hereafter, Incentive Program) has rewarded service delivery alone, paying municipalities for the coverage of home visits. Coverage reached about 90%, yet the program had no effect on preventive service use or anemia. Policymakers saw an incentive problem: the contract paid for visits, whether or not they worked.

In partnership with the Ministry of Economy and Finance and the Ministry of Health, I designed a field experiment within the Incentive Program to test whether rewarding an output would solve the problem. The experiment covers 539 municipalities, 85% of all rural municipalities in the country. I randomly assigned each municipality to one of two contracts. 270 municipalities kept the status quo *Delivery-only* contract, which

[1] Over the past two decades, at least 84 countries have adopted performance-based transfers in education and health (Muralidharan et al., 2020).

rewards the share of young children and pregnant women who receive home visits. The remaining 269 moved to a new *Delivery and Output* contract, which adds an output indicator, the share of children who are not anemic or whose hemoglobin does not fall.[2] This mixed contract is a natural middle ground. It keeps the insurance that service delivery indicators give municipalities while tying part of the reward to the principal's objective. To isolate the role of the output indicator, I hold payments constant across contracts.

The study has two phases, one for each research question. In the trial phase, from August to December 2024, municipalities operated under their randomly assigned contracts. The government announced each assignment through official resolutions and workshops. I complemented these with a nationwide campaign that sent each municipality its own indicators and thresholds. By October 2024, 98.2% of managers correctly identified their contract type. In the reversion phase, starting in January 2025, an administrative change unrelated to the trial returned all municipalities to the *Delivery-only* contract. This reversion created a unique opportunity to study persistence of the effects of the *Delivery and Output* contract on managerial practices and health outcomes.

Both contracts reward the municipality as an institution, not individual workers. The transfer therefore reaches families only through a chain of agents. The mayor receives the transfer on behalf of the municipality and appoints a manager to run the Home Visit Program. The average manager oversees 11 community health workers (CHWs). Each CHW visits up to 15 families, up to three times a month, to promote checkups and iron supplementation and to refer children at risk of anemia to the local health facility. CHWs cannot provide care or dispense supplements, so families must obtain both at the health facility. Health facility staff train CHWs and treat the children they refer as part of their regular duties. CHWs are mission-motivated volunteers who receive only a transport stipend (Besley et al., 2005). Because they are not municipal employees, managers lack the usual levers of hiring, firing, and pay. Any response to the contract must therefore come from how managers organize, monitor, and motivate CHWs, and how they coordinate with health facilities.[3]

A key strength of this study is the combination of complementary data sources that follow the incentive from the manager's office to the child's hemoglobin test. My main sources are administrative records that cover every municipality, child, and health facility in the experiment. To measure service delivery, I use the Home Visit Program's own records, which log every visit to targeted children and pregnant women. These records also identify everyone involved in the program, including managers, supervisors, CHWs, and the health facility staff responsible for the program. They let me track this workforce over time, including who is

[2]The output indicator counts a child as a success if hemoglobin does not fall between the 6- and 12-month screening tests. This definition reflects infant physiology. The iron stores a child accumulates before birth run out by about six months, just as rapid growth raises iron needs. Without adequate iron intake, hemoglobin tends to decline over this period (Cerami, 2017).

[3]At baseline, 80% of managers reported that missing the contract targets would affect their professional standing, and 87% reported high or medium autonomy over program implementation.

active, how many visits each CHW makes, and when personnel leave the program. To measure service use and child health, I use Ministry of Health records that cover every visit by children to a public health facility, including each checkup, vaccine dose, iron supplement pick-up, hemoglobin test, and diagnosis. Municipal budget records cover spending across all municipal offices and spending categories. They let me test whether the contract changed resources rather than practices, both through reallocation across municipal units and through changes in how the home visit program and health budget were spent.

I complement these records with original surveys that capture what administrative data cannot, namely how managers run the program. I surveyed program managers at baseline, at the end of the trial, and again after the reversion. These surveys capture how managers monitor performance, manage referrals, and motivate CHWs, and how these practices evolve once the output indicator is removed. To corroborate these reports, I surveyed staff at health facilities at endline. Health facility staff work alongside managers and CHWs but outside the incentive scheme, so their reports provide third-party measures of coordination and referrals that are not subject to managers' reporting bias or to experimenter demand. Together, these sources create a uniquely rich dataset that links the contract to managerial practices, frontline delivery, families' use of preventive care, and child health.

I present four sets of results, each comparing municipalities under the *Delivery and Output* contract with those under the *Delivery-only* contract. First, the *Delivery and Output* contract did not change the resources devoted to the program. During the trial, neither total municipal spending nor health spending changed. Within the health budget, spending did not shift across CHW stipends, other program spending (including personnel wages), or maternal and child health outside the home visit program. Staffing did not change either. The number of CHWs per municipality was unchanged relative to a *Delivery-only* mean of 11.7, and so was the share with higher education, which suggests that CHW composition was also stable. Turnover among program managers, supervisors, and the health facility staff responsible for the program was rare in both arms, below 2% during the trial. On the supply side, health facility staff report no change in the equipment and supplies available for anemia screening and treatment. With the same budget, staff, and health facility resources, municipalities also delivered the same number of home visits. Coverage was already close to universal, reaching 88% of targeted children and 94% of pregnant women under the *Delivery-only* contract.

Second, the *Delivery and Output* contract changed how managers used these resources. Managers reorganized their work around the path from the home visit to the health facility. They monitored the output more closely. Managers under the *Delivery and Output* contract were 13.7 p.p. (19%) more likely to monitor the anemia indicator, while their supervision of CHWs rose only slightly, by 4.6 p.p. (8%). They also managed referrals more actively and were 12.4 p.p. (24%) more likely to improve care for children at risk of anemia.

Finally, they found ways to motivate volunteers they could not pay. They were 17.1 p.p. (46%) more likely to give CHWs non-monetary incentives such as public recognition and municipal services. Health facility staff, who work alongside the program as part of their regular duties, corroborate these changes. Because they are not part of the incentive scheme, their reports serve as third-party measures that are less likely to reflect experimenter demand or reporting bias. They report that referrals were 19.5 p.p. (83%) more likely to be coordinated with the facility, and that CHWs were 14.4 p.p. (55%) more likely to accompany mothers there. They also report that CHWs received 1.2 more training sessions (31%), although this estimate is less precise.

Third, these managerial changes increased families' use of preventive care. Children in *Delivery and Output* municipalities made 0.37 more health facility visits (3.9%) and received 0.28 more vaccine doses (8.8%). These gains matter for child health because routine checkups catch developmental delays and illness early, and vaccines protect against the leading infectious causes of child death (Jones et al., 2003; Black et al., 2003). Neither facility visits nor vaccines were rewarded by the contract. Once families reached the health facility, they received its standard bundle of preventive services, so the output indicator crowded in care that the contract did not target.

Iron supplement pick-ups also doubled, rising by 0.36 bottles from a *Delivery-only* mean of 0.36. Each bottle covers one month of treatment, so even children in *Delivery and Output* municipalities received less than one month of iron over the five-month trial. Clinical trials, by contrast, require about three months of daily supplementation to raise hemoglobin by 0.56 g/dL (Gera et al., 2007). Scaling this benchmark by the additional supply implies an expected gain of at most 0.07 g/dL, even with perfect adherence. At observed adherence, where only 24.4% of children reach 90% adherence (INEI, 2023a), the expected gain falls to about 0.02 g/dL. As this benchmark predicts, the estimated effect on hemoglobin is small and statistically indistinguishable from zero. The estimate is −0.03 g/dL (s.e. 0.07), with a 95% confidence interval from −0.17 to 0.10 g/dL.[4] This interval contains both zero and the expected gains, and the minimum detectable effect of 0.20 g/dL is roughly three times the upper-bound benchmark. The experiment therefore lacks the power to detect changes in hemoglobin and anemia of the size the increase in iron implies.

Fourth, the managerial changes outlasted the contract. After all municipalities returned to the *Delivery-only* contract, managers in former *Delivery and Output* municipalities remained 13.4 p.p. more likely to track the anemia indicator, nearly the same effect as during the trial. They kept coordinating referrals with facility staff (10.3 p.p.) and rewarding CHWs (13.4 p.p.), and their CHWs kept accompanying mothers (11.3 p.p.). Only the gain in CHW training faded. These managers also served a new cohort of younger children, born

[4] Selection into testing does not explain this result. Testing rates, children's distance to the district capital, and prior contact with the health facility were similar across arms.

too late to be reached during the trial. Children in this cohort made 1.1 more health facility visits, but their health outcomes did not change. These families never experienced the *Delivery and Output* contract, so this persistence cannot reflect habits that households formed during the trial. It points instead to lasting changes in how managers run the program.

This paper contributes to three strands of literature. First, it contributes to the personnel economics of the state and to the design of public sector contracts. A large literature evaluates performance incentives for individual frontline workers (Muralidharan et al., 2011; Duflo et al., 2012; Ashraf et al., 2014). Another evaluates institutional transfers to subnational governments, schools, and health facilities (Lavy, 2002; Basinga et al., 2011; Gertler et al., 2012; Olken et al., 2014; Celhay et al., 2019; Luo et al., 2020; Bandiera et al., 2021). Few studies, however, compare contracts that weigh service delivery and outputs differently within the same institutional setting (Fryer, 2011; Mohanan et al., 2021). I show that adding an output indicator to a service delivery contract changes how local governments deliver services at scale. Managers reorganized their work, and children used more preventive services, without new resources. The contract succeeded along every margin that managers and CHWs control.

Second, the paper contributes to the literature on management in the public sector. Management quality shapes state capacity and public sector productivity (Bloom et al., 2007; Rasul et al., 2018; Fenizia, 2022). Yet civil service rules often limit standard personnel tools such as hiring, firing, and merit pay (Bandiera et al., 2021; Rasul et al., 2021; Otero et al., 2025). Managers in this setting face a sharper constraint. Because CHWs are volunteers, managers cannot pass monetary incentives down to them. I show that an institutional transfer can still change how managers work. Managers under the *Delivery and Output* contract monitored performance more closely, coordinated with other public providers, and motivated frontline workers through non-monetary means. I also provide evidence on whether such practices persist (Celhay et al., 2019; Dunsch et al., 2023). They outlasted the contract and reached a new cohort of children never exposed to it, which points to lasting changes in managerial routines.

Finally, the paper contributes to the literature on community health workers in low- and middle-income countries. Prior work studies incentives aimed at CHWs themselves, through financial compensation (Deserranno, 2019; Björkman Nyqvist et al., 2019), social recognition (Ashraf et al., 2014; Bandiera et al., 2022; Fracchia et al., 2023), or career progression (Ashraf et al., 2016). I study a different lever. The contract rewards the municipality, not the CHW, and reaches mission-motivated CHWs only through their manager. Under the *Delivery and Output* contract, CHWs accompanied more mothers to the health facility, and children made more visits and picked up more iron supplements. Incentives for local governments can therefore deploy volunteer health workers more effectively without changing what they are paid.

The rest of the paper proceeds as follows. Section 2 describes Peru's Incentive Program and childhood anemia. Section 3 describes the Home Visit Program and who delivers it. Section 4 describes the administrative and survey data. Section 5 presents the experimental design. Section 6 reports results during the trial phase. Section 7 examines whether results persisted after all municipalities reverted to the *Delivery-only* contract. Section 8 concludes.

2 Background

This section describes the institutional and public health setting of the study. It first describes Peru's administrative structure and the Incentive Program, a national performance-based scheme linking performance transfers to service delivery targets. It then documents the prevalence and consequences of childhood anemia in Peru.

2.1 Administrative organization and municipal incentives

Administratively, Peru is divided into 25 regions, 196 provinces, and 1,891 districts. Regional governments oversee the regions, while local governments, or municipalities, administer districts. Municipalities' core functions are basic local services — water and sanitation, waste collection, maintenance of public spaces, local infrastructure, business licensing, and local tax collection. Health is formally a shared competence, but the transfer of functions to municipalities was never completed. Regional governments deliver primary healthcare and hospital services, with little municipal participation (Alves et al., 2025). Municipalities do not run health facilities, employ medical personnel, or deliver medical care; their involvement in health runs through the health sector. They can organize and finance community-based activities, but diagnosis and treatment remain with the health facility.

In 2009, the Ministry of Economy and Finance launched the Incentive Program for Municipal Management Improvement[5] to improve the quality of local service delivery. The Incentive Program operates as an institutional performance-based financing scheme, allowing municipalities to access additional budget conditional on achieving specific service delivery targets. The transfer goes to the municipality as an institution, not to the individuals who implement the program. Governments often favor institutional over individual incentives because they limit discretion, enhance transparency, and reduce corruption risks. The design has a cost, however: it can generate free-riding and weak internal accountability, since the rewards

[5] In Spanish, *Programa de Incentivos a la Mejora de la Gestión Municipal*. Before 2016, it was known as *Plan de Incentivos a la Mejora de la Gestión Municipal*. For more information, see www.mef.gob.pe.

need not translate into direct benefits for frontline workers (G. Miller et al., 2014).[6]

Participation in the Incentive Program is voluntary; however, in practice, all municipalities participate because the potential transfer represents an important share of their expected annual tax revenue. On average, municipalities can obtain resources equivalent to 15.7 times their expected annual tax revenue (Figure A.3), with a median ratio of 3.4. The incentive payments are disbursed after performance verification, typically during the following fiscal year, and can be used freely to finance the provision of public services.

2.2 Children's anemia in Peru

Anemia is a condition in which the number of red blood cells or the hemoglobin concentration is lower than normal. Anemia during early childhood is broadly associated with growth retardation, reduced resistance to infections, delayed psychomotor development, lower language acquisition, and impaired cognitive and socioemotional development (INS, 2024). These developmental deficits can become a significant deterrent to human capital accumulation in developing countries, and hence a potentially important barrier to economic development (Chong et al., 2016).[7]

Yet anemia continues to be a pressing global concern. Nearly half of all children under five worldwide are affected (42.8%), and the burden is even more severe in low- and middle-income countries, where prevalence reaches 56.5% (WHO, 2025). In Peru, the prevalence of anemia among children under five years of age is 33.4%, and 60.5% among those aged 6 to 12 months (INEI, 2023b). Prevalence is consistently higher in rural areas where households are poorer and have less access to basic services: among children aged 6 to 12 months, it reached 64.9% in rural areas in 2023, compared to 58.7% in urban areas.[8]

Anemia is primarily caused by iron deficiency (WHO, 2023) and it is the most prevalent micronutrient deficiency in the world (Chong et al., 2016). For children under one year of age, the most appropriate interventions are iron supplementation and caregiver nutritional education, given their age-specific nutritional needs and feeding practices.[9] Supplementation is administered daily in drops or syrup for at least three consecutive months, and adherence is critical to its effectiveness (WHO, 2016). Clinical trials show that when

[6] Performance-based transfers of this kind have become increasingly common across Latin America and the Caribbean (Alves et al., 2025). A survey of 17 Latin American and Caribbean countries finds that performance-based financing is in place in 14 of them (Pimenta et al., 2023).

[7] Chong et al. (2016) find that, among Peruvian teenage students who were anemic at baseline, three months of iron supplementation raised average test scores by 0.4 standard deviations and increased educational aspirations by 16 percent.

[8] 36% of rural households live in poverty or extreme poverty, compared to 15% in urban areas; 85% have access to sanitation, compared to 96%; and 90% rely on government-sponsored health insurance, compared to 59% (ENAHO, 2024).

[9] The World Health Organization recommends four strategies to combat anemia: iron supplementation, food fortification, nutrition education, and dietary diversity (WHO, 2006). Although food fortification is generally considered the most effective and cost-efficient strategy (Mithra et al., 2020), it is poorly suited to this age group. Infants are typically exclusively breastfed until six months, and even after complementary feeding begins, small portion sizes limit the intake of fortified and diverse foods.

adherence is enforced, supplementation can raise hemoglobin by 0.56 g/dL among children under 24 months (Gera et al., 2007).

Although iron deficiency is the leading cause, anemia is a multi-causal condition. Beyond dietary iron intake, hemoglobin status in early childhood depends on maternal iron stores transferred in utero (Cerami, 2017) and children's capacity to absorb ingested iron, which is shaped by environmental conditions such as access to safe water and adequate sanitation (Kothari et al., 2019; Larsen et al., 2017).[10] Other micronutrient deficiencies (folate, vitamin B12, vitamin A) and inherited disorders contribute as well (WHO, 2001). As a result, even sustained iron supplementation need not translate into measurable hemoglobin gains when these complementary conditions are not met.

3 The Home Visit Program

Home visit programs are a widely used service delivery model in low- and middle-income countries. They bring services into the household, reaching families that have little contact with health, education, and social programs. A growing body of evidence shows they improve caregiver knowledge (Stansert Katzen et al., 2020), feeding practices (Galasso et al., 2019; Janmohamed et al., 2020), and child development (Yousafzai et al., 2014; Grantham-McGregor et al., 2020; Zhou et al., 2023). Peru adopted this model against childhood anemia in 2019, when the Incentive Program began rewarding municipalities for implementing a new home visit program, *Compromiso 1*. The program seeks to increase adherence to iron supplementation among children under 12 months of age.[11]

The Home Visit Program is a community education initiative that delivers information to caregivers on how to prevent and reduce anemia in children. The program targets 879 municipalities, including all 245 urban municipalities located in main cities and 634 of the most rural municipalities.[12] Municipalities, in collaboration with healthcare staff, conduct home visits to eligible households. The program reaches the most vulnerable populations: in urban districts, the program targets low-income families; in rural districts, coverage is universal.[13]

[10] Diarrhea and intestinal parasites cause infection and inflammation that impair iron absorption, so a child can take the full dose and still absorb little of the iron (Pasricha et al., 2021).

[11] The Ministry of Health designed the program. Its full name in Spanish is *Compromiso 1: Mejora del estado nutricional y de salud de las niñas y niños de 3 a 12 meses*. It builds on *Cuna Más*, a home visit program run centrally by the Ministry of Development and Social Inclusion since 2012. *Compromiso 1* instead assigns delivery to local governments. It thus marked a first step toward the decentralization of public health functions, giving municipalities an operational role in preventive, community-based care.

[12] Municipalities are classified into seven types (A to G) based on their urban status, whether they are the main district of a province, and their share of urban population (Table B.1). According to this classification, the Home Visit Program targets the urban municipalities from main cities and the group of the most rural municipalities.

[13] Eligibility in urban districts is determined by health insurance status. Targeted children are those without any health insurance or enrolled in the *Seguro Integral de Salud* (SIS), a government-sponsored insurance program for low-income populations, equivalent to

Home visits are tailored to the child's age and health status. Visits for children aged 1–5 months are preventive, targeting a critical developmental window in which neurological development accelerates and iron needs rise sharply (MIDIS, 2018). For children aged 6–12 months, visits may be preventive or recuperative depending on whether anemia has already been diagnosed. Each household receives up to three visits per month, each lasting approximately 20 minutes. During these visits, CHWs deliver standardized messages on nutrition, hygiene, preventive care, and early childhood development.

3.1 Home visit organization

The Home Visit Program operates through a hierarchy of five actors, illustrated in Figure 1. The Central Government designs the incentive contract and disburses the conditional bonus to the Municipality, headed by the Mayor. The Mayor receives the bonus on behalf of the municipality and holds unilateral authority to hire, dismiss, or reassign the Municipal Manager, who runs the Home Visit Program on the ground.[14] The Manager in turn leads the Community Health Workers (CHWs), who deliver home visits to Families with young children. Health outcomes are ultimately realized at the family level, and verifiable indicators—visit coverage and, under the *Delivery and Output contract, the output indicator*—feed back to the Central Government to determine the disbursement of the bonus. The rest of this section describes the main actors.

The Municipal Manager

Within each municipality, the Home Visit Program is run by a single Municipal Manager. The manager is appointed by the mayor—who receives the program's conditional transfer from the central government on behalf of the municipality and retains the authority to dismiss or reassign the manager—and in turn leads the Community Health Workers (CHWs) who carry out the visits.

Managers combine intrinsic and extrinsic motives. Baseline survey data show that 55% are motivated primarily by the program's contribution to population wellbeing, while the remaining 45% are motivated primarily by meeting the targets that secure the incentive payment. Institutionally, the manager's key private stake is continued employment: the mayor holds unilateral hiring and dismissal authority, and 80% of managers report that failing to meet the performance thresholds set by the incentive contract would jeopardize their position.

The municipal manager oversees all health and social programs within the municipality, including the

Medicaid in the United States.

[14]The majority of service delivery agents (80%) are CHWs, so, for simplicity, I use the terms service delivery agents and CHW interchangeably. About half of municipalities are large enough—by target-population size—to add an intermediate organizational layer of supervisors between the manager and the CHWs; each supervisor oversees roughly 20 CHWs.

implementation of the Home Visit Program. Within the program, managerial effort consists of four practices. *Monitoring* covers both tracking the performance indicators on which the municipality is evaluated and verifying CHWs' visits directly, through accompaniment, household calls, or re-visits. *Referral management* runs from identifying children at risk of anemia to acting on their care, which places the manager as the main point of contact with health facility staff and requires following referred children through to treatment. *Motivating CHWs* involves distributing discretionary rewards and cultivating a sense of mission, complementing the pro-social motives that already drive them. *Training* involves organizing refresher sessions, coordinated with health facility staff, beyond the mandatory initial training. Managers also hold personnel (i.e., hiring, dismissals, promotions) and budget authority.

Managers exercise substantial discretion over how they allocate effort across these activities: 46% report a high level of overall autonomy, 41% a medium level, and only 11% a low level. As Figure A.4 shows, autonomy is highest in coordinating with health facilities (71%) and training CHWs (59%), and remains substantial in hiring CHWs (52%), managing personnel payments (47%), planning activities (44%), and incentivizing CHWs (42%). A large majority (87%) believe the program's success depends on their own effort.

The manager observes the number of home visits each CHW delivers to each child, since visit records are logged in the administrative platform. The manager does not directly observe the quality of individual visits—whether the CHW counseled the mother effectively, assessed the child correctly, and left the family motivated to comply. Quality is inferred indirectly, through occasional accompaniment and through feedback from health facility staff, who verify a random 10% sample of home visits.

The Community Health Worker

CHWs work under a quasi-volunteer arrangement, receiving only a transport stipend of approximately USD 3 per visit that essentially covers commuting costs rather than serving as labor compensation. This arrangement is consistent with the broader CHW workforce in low- and middle-income countries, where engagement relies heavily on pro-social motivations rather than financial returns (Ashraf et al., 2020). To perform their tasks, CHWs receive mandatory two-day in-person training from Health Facility staff on essential topics and learn to recognize children's danger signs for referral. They also receive discretionary on-the-field training from the manager.

CHWs are responsible for scheduling and conducting the home visits.[15] Each home visit involves two distinct margins of effort: quantity and quality. On the quantity margin, the CHW decides how many visits to

[15] If a caregiver is not home at the scheduled time, the visit is rescheduled rather than skipped; outright refusals are rare, occurring in fewer than 5% of cases.

deliver to each assigned child during the reference period, subject to the program's target of up to three visits per month. On the quality margin, the CHW decides how carefully to follow the visit protocol—that is, how thoroughly to assess the child, how attentively to counsel the caregiver , and whether to refer children showing danger signs.[16] The visit protocol is standardized through a checklist and a flipchart that guide the delivery of each topic. Family adherence to iron supplementation and healthy practices is influenced by the CHW's effort, but compliance ultimately rests on the family's own choice.[17]

The Health Facility

Peru's Ministry of Health classifies health facilities by the complexity of the care they deliver. The categories range from Type I-1 health posts (postas), which offer only basic preventive and outpatient care, to Type III hospitals, which offer specialized treatment and surgery. Type I-1 posts sit at the base of this hierarchy. Type I-1 posts deliver the routine preventive services that children receive during their first year of life—growth-monitoring visits, vaccinations, anemia screening, and iron supplementation through their child health unit. These posts are also the facilities that participate in the Home Visit Program.

Although health facilities are not part of the incentive scheme, they play a key supporting role through three functions beyond direct service provision. First, child health unit staff train CHWs, delivering both the initial mandatory two-day training and discretionary follow-up sessions coordinated with municipalities. Second, they verify the veracity of a random 10% sample of home visits. Third, they manage referrals from CHWs, ensuring that children identified as at risk of anemia or in need of further care are promptly seen and treated.

4 Data

4.1 Administrative data

I draw on four sources of administrative data: the Public Healthcare Information System, which records all health services delivered to children; the Home Visit Program dataset, which tracks CHW characteristics and completed home visits; the Public Budget Information System, which records municipal expenditures and

[16] CHWs are not permitted to provide medical care or dispense supplements themselves; when they identify a child at risk, they refer the case to the health facility and encourage the caregiver to bring the child in. Also, by national regulation, CHWs cannot directly dispense iron supplements; medication must be obtained from a health facility.

[17] Compliance is multidimensional—administering iron supplementation daily at the correct dose and time, attending health facility appointments, adopting iron-rich dietary changes, and applying hygiene and sanitation practices—and costly to families in time and effort, particularly in rural areas where health facilities are distant.

transfers; and district-level poverty and census indicators used to characterize baseline conditions and examine treatment effect heterogeneity.

4.1.1 Public Healthcare Information System – HIS

I use anonymized administrative data from the public healthcare information system, which records all medical services delivered to children, including diagnoses. These services include routine growth and development check-ups as well as vaccinations. Over the first year of life, a total of 17 vaccine doses are scheduled at specific ages in accordance with national immunization guidelines (Figure A.7).[18]

The database also contains detailed information on anemia screening. Routine hemoglobin tests are conducted at two recommended ages—between 6 and 9 months and between 12 and 13 months—and include measured hemoglobin concentrations as well as the corresponding anemia diagnoses. Children are classified as anemic if their altitude-adjusted hemoglobin concentration is below 10.5 g/dL.[19] Consistent with this, preventive iron supplementation is provided starting at four months of age and continued during the first year of life. When anemia is detected through routine screening, children are prescribed therapeutic (higher-dose) iron supplementation. Figure A.6 summarizes the timing of health-facility–provided services during the first year of life.

The dataset also includes measurements of child weight and height collected during routine visits, from which anthropometric indicators can be constructed. Given the five-month duration of the intervention, I focus on weight-for-age z-scores (WAZ) rather than height-for-age z-scores (HAZ), as weight is more responsive to short-term nutritional changes while length and height reflect longer-term growth processes (Frison et al., 2016).

4.1.2 Home Visit Program

I use the Home Visit Program administrative dataset, which contains detailed information on CHW characteristics, payroll, and the timing of all completed home visits. This database enables the identification of the CHW assigned to each family and allows me to assess CHW turnover. Because CHWs are volunteers receiving only a transport stipend, with no performance pay or career advancement tied to logged visits, the

[18] During the first month of life, children attend four growth-monitoring and counseling visits at health facilities, followed by monthly visits through month 11.

[19] Hemoglobin is measured at health facilities using capillary blood obtained via finger prick and analyzed with portable hemoglobinometers. In line with national guidelines and World Health Organization (WHO) standards, hemoglobin concentrations are adjusted for altitude above 500 meters because individuals living at higher elevations typically exhibit higher hemoglobin levels as a physiological response to reduced oxygen availability (WHO, 2011). Without this adjustment, children residing at higher altitudes would be mechanically less likely to be classified as anemic despite facing comparable or higher levels of iron deficiency.

scope for strategic manipulation of these records is limited. This is reinforced by the verification protocol, in which health-facility staff cross-check a random 10% sample of visits against household reports.

4.1.3 Public Budget Information System – SIAF

I use administrative data from the *Sistema Integrado de Administración Financiera* (SIAF), Peru's Integrated Financial Management System administered by the Ministry of Economy and Finance. SIAF records all public expenditures and transfers executed by national, regional, and local governments, including the funds municipalities receive through the Incentive Program. The data are available at the monthly transaction level and provide detailed information on expenditure amounts, timing, budget classifications, and revenue sources. I use these data to track municipal spending patterns across functional categories and to assess whether municipalities reallocate resources toward activities related to the Home Visit Program.

4.1.4 District-Level Data

I use district-level poverty indicators from the official Poverty Map produced by the *Instituto Nacional de Estadística e Informática* (INEI). The poverty map provides estimates of monetary poverty at the district level using a small-area estimation methodology that combines household survey data with information from the 2017 Demographic Census Data. The incidence of monetary poverty allows me to characterize baseline socioeconomic conditions. In addition, I use district-level indicators of access to safe drinking water and sanitation from the census. Together, these indicators allow me to examine heterogeneity in treatment effects by baseline poverty and environmental conditions.

4.2 Survey Data

I collected primary data through two survey instruments, one targeting municipality managers and the other targeting health facility staff.[20] The manager survey was administered at three points in time — baseline (March–April 2024), a short midline follow-up (October 2024), and endline (June–July 2025). The health facility survey was administered only at endline (August–September 2025). Endline surveys were designed to retrospectively capture information on two distinct periods: the trial phase (August–December 2024), when municipalities operated under their assigned incentive contracts, and the reversion phase (January–May 2025), when an unanticipated administrative delay prevented renewal of the *Delivery and Output* contract and all municipalities reverted to the *Delivery-only* contract. Figure 2 summarizes the project timeline. This

[20] I did not administer a household or caregiver survey, so I do not observe caregiver attitudes, knowledge, satisfaction, or perceived value of the home visits.

retrospective two-period structure allows me to assess both the contemporaneous effects and the persistence of any behavioral changes.

4.2.1 Survey to Managers

Because the municipal incentive operates through municipal managers rather than directly through frontline CHWs, the manager survey is the central instrument for documenting the behavioral mechanisms through which the contract affects implementation. The baseline survey collected data on managers' demographic backgrounds and Home Visit Program staff characteristics. It also elicited beliefs and expectations about the incentive scheme and gathered detailed data on CHW training (e.g., number of sessions and trainers), implemented monitoring practices, incentives to personnel, complementary anemia-reduction strategies, and partnerships with other institutions.

Since the effectiveness of incentive contracts depends on municipalities fully understanding their terms, the midline follow-up assessed managers' comprehension of the incentive scheme and their assigned treatment arm, along with a small set of questions on the implementation of the Home Visit Program. At endline, I successfully tracked 442 managers (82%) who were able to answer questions related to both study periods. Attrition rates were balanced across treatment arms (see Table A.5 and Table A.2).[21]

4.2.2 Survey to Health Facilities

The endline survey was administered to the health staff member responsible for implementing the Home Visit Program at each health facility.[22] The questionnaire collected information on facility-level management practices, supply availability, coordination with municipalities, and verification of home visits. Specifically, it included modules on logistics and availability of equipment, supplies, and iron supplements; coordination mechanisms with municipal teams and CHWs; frequency and modalities of CHW training sessions; and procedures for prioritizing and attending community referrals. Additional sections gathered information on the CHWs' work environment and collaboration dynamics within the health facility.

[21] A potential concern is that the incentive scheme could differentially affect managerial turnover, creating selection into treatment implementation. If managers in *Delivery and Output* municipalities exit or enter at higher rates, estimated treatment effects might partly capture changes in implementer composition rather than behavioral responses to the contract. However, Table A.6 shows that the probability a manager remained in the municipality throughout the trial phase is statistically indistinguishable across treatment arms, mitigating concerns about differential turnover.

[22] The endline survey was completed by around 40% of the universe of health facilities located in the municipalities in the experiment. This survey take-up rate is broadly in line with meta-analytic benchmarks for online surveys (Wu et al., 2022). I was not able to collect baseline information on health facilities. In Table A.3, I use health facilities' predetermined characteristics that should not have been affected by the intervention to test for balance in covariates. None of the differences in variables is statistically significant, confirming that health facilities in municipalities under the *Delivery-only* contract constitute a valid counterfactual for those in municipalities under the *Delivery and Output* contract.

5 Research design

This section describes the experimental design. I begin by presenting the incentive payment structure and the two contract variants at the center of the experiment. I then document baseline manager motivation and perceived stakes, which shape the behavioral environment in which the contracts operate. Finally, I describe the randomization procedure, the econometric specification, and present checks on incentive awareness and baseline balance.

5.1 Incentive Payment Structure

The incentive program rewards municipalities through a step-function payment schedule defined separately for each performance indicator. For every indicator, the central government sets two pre-determined thresholds: a *partial-bonus threshold* and a *full-bonus threshold*, where the partial-bonus threshold is a lower performance level than the full-bonus threshold. The partial-bonus threshold is the minimum performance level a municipality must reach to receive 75% of the bonus tied to that indicator. The full-bonus threshold is the performance level at which the entire bonus is paid. Municipalities performing below the partial-bonus threshold receive nothing for that indicator. Performance above the full-bonus threshold is feasible, but generates no additional payment. Both thresholds are set in advance based on each municipality's baseline performance, are common across treatment arms, and are not modified by the assignment of the incentive contract.

Formally, each municipality d is eligible for a total potential bonus B_d, distributed equally across K performance indicators, so that each indicator k carries a potential bonus of $B_{dk} = B_d/K$. Municipality d's payment for indicator k, denoted as P_{dk}, follows the step function below, where x_{dk} denotes performance, $\underline{x}_{dk}$ denotes the partial-bonus threshold, and $\bar{x}_{dk}$ denotes the full-bonus threshold:

$$P_{dk} = \begin{cases} 0, & \text{for } x_{dk} < \underline{x}_{dk} \\ 0.75 * B_{dk}, & \text{for } \underline{x}_{dk} \leq x_{dk} < \bar{x}_{dk} \\ B_{dk}, & \text{for } x_{dk} \geq \bar{x}_{dk} \end{cases} \tag{1}$$

Figure A.8 plots the payment schedule for a single indicator following Equation 1. The municipal incentive payment $P_d = \sum_{k=1}^{K} P_{dk}$ is disbursed after performance verification and at the beginning of the following fiscal year, without affecting the ongoing implementation. Before the trial began, from January to July 2024, municipalities were evaluated solely on whether they met the predetermined target for the indicator measuring the share of children under 12 months who received a home visit. Consequently, the total bonus

that municipalities could obtain depended entirely on their performance in that single indicator.

5.2 The Intervention

A new incentive scheme for the period August – December 2024 was developed in joint collaboration with the Ministry of Economy and Finance and the Ministry of Health. It was formalized through a series of legal resolutions.[23] The incentive scheme considers up to three indicators:[24]

(i) The percentage of children under 12 months old who receive home visit

(ii) The percentage of pregnant women who receive home visits

(iii) The percentage of children who are non-anemic or have not experienced a decline in hemoglobin levels between their six- and twelve-month blood screenings

Based on these indicators, municipalities were offered one of two alternative incentive contracts:

- Under the *Delivery-only* contract, municipal performance is evaluated exclusively on the two service delivery indicators (i and ii). Municipalities are eligible for a total potential bonus of B_d, which is divided equally across the two indicators, so that each indicator carries a potential bonus of $B_{dk} = B_d/2$.
- Under the *Delivery and Output* contract, municipal performance is evaluated on the same two service delivery indicators (i and ii) as well as the output indicator (iii). The total potential bonus B_d is held constant across contracts but is divided equally across the three indicators, such that each indicator carries a potential of $B_{dk} = B_d/3$. Critically, the *Delivery and Output* contract did not expand the total financial stakes faced by municipalities — it only reoriented them, reducing the share of the bonus tied to each service delivery indicator while introducing an additional output-based component.

5.3 Manager Motivation and Perceived Stakes

Understanding how managers perceive the incentive scheme is important for interpreting their behavioral responses to the two contracts. A natural question is why managers respond to it at all, given that the bonus accrues to the municipality rather than to them personally and arrives with a lag of roughly a year. Table 1 presents descriptive statistics from the managers' baseline survey on three dimensions relevant to this question: their beliefs about the cost of missing performance targets, their beliefs about how their own effort maps into children's health, and their underlying motivation.

[23] See, among others: Law 31953, DS 318-2023-EF, RD 0006-2024-EF/50.01, DS 132-2024-EF, and RD 0027-2024-EF/50.01.

[24] The legal norm containing the details of these indicators can be found here.

On the first, managers in rural municipalities anticipate substantial personal consequences from missing their targets, elicited as perceived probabilities at baseline: 56% believe they would be dismissed and 31% expect reassignment to another department, and they perceive these stakes as cascading down the team, assigning a 36% probability that CHWs would also face dismissal. On the second, managers see their own effort as decisive: 87% report that program success depends on what they do. They are also intrinsically motivated: 55% cite population wellbeing rather than target attainment as their strongest motivation. Altogether, these patterns indicate that managers respond to the incentive not because of the bonus itself, but because of the personal accountability the scheme creates and their intrinsic motivation to improve child health.

5.4 Randomization and Econometric Specification

I use a clustered randomized controlled trial to evaluate the impacts of the *Delivery and Output* contract on municipalities' service delivery and managerial practices. Randomization was done on the group of rural municipalities at the municipality level.[25] For evaluation purposes, I restrict the sample to rural municipalities with at least 15 children, resulting in a final group of 539 municipalities. Using a random number generator in STATA, 270 were randomly assigned to the *Delivery-only* contract, while 269 were assigned to the *Delivery and Output* contract. Figure A.1 shows the geographical distribution of the municipalities based on their involvement by the Home Visit Program and by treatment arm.

The treatment assignment was disseminated through multiple institutional channels to ensure transparency and municipal awareness. In late July 2024, the Ministry of Economy and Finance published the legal resolutions in the national newspaper and on its website. These resolutions included a list of all participating municipalities along with the indicators and performance thresholds assigned to each. The Ministry of Health complemented this announcement with in-person and virtual workshops for municipal authorities. To facilitate clearer communication, and in close coordination with the Ministry of Health, I implemented a nationwide online information campaign during September 2024. Unlike the consolidated official list (Figure A.5), this campaign delivered municipality-specific information, providing each municipality only with the indicators and thresholds relevant to its own performance evaluation. A timeline of these dissemination activities is presented in Figure 2, and additional details on the information campaign are provided in Appendix D.

For outcomes at the children level, I estimate the following specification using ordinary least squares (OLS):

$$Y_{id} = \alpha + \beta DeliveryOutput_d + \gamma Y_{id}^0 + \epsilon_{id} \tag{2}$$

[25] All the 245 urban municipalities were assigned to the *Delivery and Output* contract by government's decision.

where Y_{id} is the outcome of interest at follow-up for child i in municipality d at baseline. $DeliveryOutput_d$ is an indicator equal to 1 if municipality d was assigned to the treatment group (i.e. *Delivery and Output* contract) and 0 otherwise. The comparison group represents the status quo scenario with municipalities under the *Delivery-only* contract. Because municipalities were randomly assigned to the *Delivery and Output* and *Delivery-only* contracts, the *Delivery-only* group constitutes a valid counterfactual for the *Delivery and Output* group by design, so β identifies the causal effect of the *Delivery and Output* contract. Since the randomization was conducted at the municipality level, I cluster the standard errors by municipality. Y^0_{id} is the baseline level of the outcome variable and it is included when available to improve precision.

I estimate the same specification at two additional levels of aggregation, depending on the data source for each outcome. For outcomes from the manager survey and for the administrative measures of home visit coverage, staff and budget, the unit of observation is the municipality. Each municipality designates one manager for the Home Visit Program, so the manager-level and the municipality-level samples coincide. I estimate:

$$Y_d = \alpha + \beta DeliveryOutput_d + \gamma Y^0_d + \epsilon_d \tag{3}$$

where Y_d is the follow-up outcome for municipality d and Y^0_d is its baseline value. I surveyed managers before the intervention started, so I observe baseline levels for most outcomes and include Y^0_d when it is available.

For outcomes from the health facility survey, the unit of observation is the health facility h located in municipality d. I estimate:

$$Y_{hd} = \alpha + \beta DeliveryOutput_d + \epsilon_{hd} \tag{4}$$

I collected health facility data only at endline, so I cannot control for baseline outcomes in these regressions. Across all three levels, treatment varies only at the municipality level, so I cluster the standard errors by municipality throughout.

5.5 Incentive Awareness and Participation

Before examining behavioral responses, it is essential to confirm that municipalities correctly understood the performance criteria associated with their assigned contracts and chose to participate in the incentive scheme. Participation in the Home Visit Program was near-universal during the trial phase: 99.6% of municipalities conducted at least one home visit.[26] Regarding comprehension, approximately two months after

[26]Participation is verified using the Home Visit Program's app, through which CHWs register completed visits in real time. Two municipalities in the *Delivery-only* arm had no target children in their jurisdiction during the trial phase and were therefore unable to conduct home visits; all remaining municipalities had at least one home visit recorded.

the intervention began, 99% of municipalities correctly identified whether they were under a *Delivery-only* or a *Delivery and Output* contract (Figure 4).[27] This supports the internal validity of subsequent analyses.

5.6 Balance Checks

Table 2 presents the balance tests comparing municipalities assigned to the *Delivery-only* contract and those assigned to the *Delivery and Output* contract at baseline. Panel A describes district-level socioeconomic conditions. On average, 50% of households are connected to a public water and sewage network, and 83% of the population is covered by government-sponsored health insurance. Population (18 or more years old) has approximately 7.2 years of education. Municipalities, on average, spend an average of USD 606.7 per capita on total expenditure, and USD 16.2 per capita on health.

Panel B describes the health facilities operating in each district. Districts contain 5.6 health facilities on average. Half of these facilities are Type I-1 and participate in the Home Visit Program. These health facilities sit approximately 13.0 kilometers from the district capital on average, and have operated for 26 years.

Panel C summarizes key attributes of the Home Visit Program's workforce. Managers are, on average, 33 years old, 44% are male, and 30% have higher education. They bring an average of 1.9 years of work experience specifically in the Home Visit Program. Supervisors constitute a relatively small team, with an average of 0.6 supervisors per municipality, and 29% have a higher-education degree. Community Health Workers form the largest group, with an average of 11.6 CHWs per municipality; about 16% of them have completed higher education. Panel D describes baseline child health indicators. On average, children received 12.2 health facility visits and their caregivers collected iron supplements approximately 2.5 times.

Regarding balance across treatment arms, the only statistically significant differences arise in the gender composition of managers: municipalities under the *Delivery and Output* contract have a higher share of male managers ($p < 0.01$), and in the number of years since the Type I-1 health facility opened ($p < 0.05$).[28] Overall, fewer than 10% of the differences across all variables in the table are statistically significant, reinforcing that the random assignment generated well-balanced treatment arms at baseline.

[27] Municipal managers also demonstrated relatively strong comprehension of the definition and objective of the anemia indicator (Table A.4). Approximately 81% correctly recognized that the output indicator is a composite measure — namely, that performance improves when a child either remains non-anemic or does not experience a decline in hemoglobin between 6 and 12 months of age. In contrast, understanding of the more technical components of the indicator, such as the exact rules governing the numerator, denominator, and deductions, was substantially lower, indicating limited grasp of how performance would ultimately be evaluated.

[28] *A priori*, it is unclear whether —and in which direction— this difference could bias the estimates, as the relationship between manager gender and managerial practices is not well established. However, robustness checks that control for manager gender yield virtually identical results, indicating that this imbalance does not affect the findings.

6 Results During the Trial Phase

6.1 Resources and Home Visit Coverage

A natural first response to the output indicator would be to spend differently. Hemoglobin depends on more than home visits, so municipalities under the *Delivery and Output* contract might have funded inputs that the *Delivery-only* contract did not reward, such as higher transportation stipends for CHWs or nutrition support for children at risk of anemia. They might also have financed these inputs by shifting funds within the health budget toward maternal and child health, or by moving funds from other municipal offices toward the program. I test both possibilities using municipal budget records.

The data reject both possibilities. First, the contract did not change the size or the composition of the municipal budget (Table A.9, Panel A). Total municipal expenditure did not respond. Municipalities did not move funds across functional areas either. None of the seven areas of the municipal budget (production and communications, environment, public security, governance and social protection, health, education, and sanitation) shows a statistically significant difference across arms. Second, the contract did not change how the health budget was allocated (Table A.9, Panel B). Spending on the Home Visit Program itself did not move, whether on payments to CHWs, which fund their transportation stipends, or on other Program spending, including personnel wages. Spending on infant and maternal health outside the Program and on other health conditions was also unchanged. Municipalities thus did not respond to the output indicator with new resources.

The managers who ran the program also remained the same during the trial phase. Administrative records from the Home Visit program show that turnover among program managers and supervisors was rare in both arms (Table A.6, column 1). Between December 2023 and December 2024, 0.6% of managers and supervisors left their positions under the *Delivery-only* contract. Turnover was 1.2 p.p. higher under the *Delivery and Output* contract (s.e. 0.007), but this difference is not statistically significant.[29] These managers also kept the same workforce (Table A.8). According to managers, the number of CHWs per municipality changed by −0.28 (s.e. 1.00), relative to a *Delivery-only* mean of 11.7, which rules out increases larger than about 1.7 CHWs. The share of CHWs with higher education did not change either (1.3 p.p., s.e. 0.022), which suggests that CHW composition was also stable. The number of supervisors rose by 0.39 from a mean of 0.65, but this estimate is imprecise (s.e. 0.24) and not statistically significant, and the share of supervisors with higher education was unchanged. On the supply side, health facility staff report no change in stock-outs of supplies and equipment, or in the availability of trained personnel (Table **??**). The availability of iron supplements is

[29]The employment records do not distinguish managers from supervisors, so column 1 pools both positions. Each municipality has a single manager and, on average, 0.65 supervisors, so managers account for most of the pooled sample. Even if every departure under the *Delivery and Output* contract were a manager, manager turnover would remain around 3%.

7.9 p.p. (13%) higher in facilities located in *Delivery and Output* municipalities, an estimate significant only at the 10% level.

With the same budget, staff, and health facility resources, municipalities also delivered the same number of home visits. Before the trial, anemia rates remained high even though home-visit coverage was near-universal and municipalities routinely met the visit thresholds. Policymakers traced the problem to the quality of the visits, not their quantity. The output indicator was therefore meant to shift effort along the quality margin rather than to expand coverage.[30] Consistent with this, Table 3 shows that the *Delivery and Output* contract did not affect home visit coverage for either children or pregnant women. For children, the estimated effects on both overall coverage and target achievement are small and statistically indistinguishable from zero (0.5 and -0.4 p.p.) relative to the *Delivery-only* mean of 88% and 75%, respectively. The same pattern holds for pregnant women. Coverage remains around 94% in both arms and the likelihood of meeting the partial-bonus target is similarly unaffected.

6.2 Managerial Practices

The *Delivery and Output* contract changed how managers used these resources. Managers reorganized their work around the path from the home visit to the health facility. Table 4 reports the effects on the managerial practices described in Section 3, in four panels, one per practice.[31] I first report the practices that managers describe in the managers' survey, and then turn to the reports of health facility staff, who work alongside the program but outside the incentive scheme.

Monitoring. Managers monitored the output more closely. The first row of Panel A shows that managers under the *Delivery and Output* contract were 13.7 p.p. more likely to track the anemia indicator daily or biweekly, an 18.9% increase over the *Delivery-only* group. This is the new indicator on which part of their bonus depended. Their direct supervision of CHWs, by contrast, rose only slightly, by 4.6 p.p. (7.8%). Supervision covers any method managers use to verify visits, such as accompanying CHWs to observe visit quality first-hand, calling households, or re-visiting them.

Referral management. Referral management covers the processes managers use to identify children at risk, initiate referrals, and follow them through to treatment. Managers under the *Delivery and Output* contract

[30] Baseline coverage was already near its maximum. The program locates target families from administrative records, using the latest address reported by any public service provider who attended the child. CHWs try to trace unreachable households, but outdated addresses and recent migration leave a residual fraction unlocatable. This limits the scope for further coverage gains.

[31] Results are robust to alternative samples and levels of aggregation (Table A.7). Estimates from the managers' survey hold when the sample is restricted to municipalities with at least one surveyed health facility (Panel A), and estimates from the health facilities' survey hold when the data are aggregated to the municipality level (Panel B).

were 12.4 p.p. more likely to strengthen care for children identified in the field as at risk of anemia, a 24.4% increase (Panel B, first row).

Health facility staff corroborate these changes. They work alongside the program as part of their regular duties but are not part of the incentive scheme, so their reports serve as third-party measures that are less likely to reflect experimenter demand or managers' reporting bias. Health facility staff report that referrals were 19.5 p.p. more likely to be coordinated directly with the facility to identify open slots and arrange appointment days for referred children (83% increase). In the standard process, a mother travels to the facility, queues to register an appointment, and returns on a later date to attend it. Coordination lowers the cost families face in reaching a medical consultation. Health facility staff also report that CHWs went further outside of their duties and accompanied mothers to the health facility to ensure they do not miss their appointment (14.4 p.p.), a 55.2% increase over a *Delivery-only* mean of 0.26.[32]

Motivating CHWs. Managers also found ways to motivate volunteers they could not pay. Managers under the *Delivery and Output* contract were 17.1 p.p. more likely to reward CHWs, a 46.1% increase (Panel C). The rewards were non-monetary and not tied to performance (i.e., public recognition, training, career advancement opportunities, and in-kind goods).

Reports from health facility staff. Finally, municipalities under the *Delivery and Output* contract coordinated with health facilities to run more training sessions (Panel D). CHWs received 1.2 additional sessions from facility staff (31%); however, this estimate is significant only at the 10% level.

6.3 Use of Preventive Care

These managerial changes increased families' use of preventive care. The analysis sample consists of children who were between 1 and 9 months old during the five-month trial phase. This age window ensures two things: first, that all children in the sample were exposed to the Home Visit Program for at least one month before their first routine anemia screening; and second, that all children reached the age for that first screening during the trial phase (i.e., 6 to 8 months). This cohort is also particularly well suited for evaluating impacts on health service demand and health outcomes, as most routine preventive services — including vaccinations, growth-monitoring visits, and iron supplementation — are delivered between 1 and 9 months of age.[33]

[32] CHWs are volunteers whose transport stipend covers up to three home visits per month, and it does not cover this additional trip.

[33] According to the national immunization schedule, 12 of the 17 vaccine doses administered during the first year of life are delivered between 1 and 9 months of age (Figure A.7).

Panel A of Table 5 reports the effects on health service utilization. Children in *Delivery and Output* municipalities made 0.37 more health facility visits (3.9% relative to the *Delivery-only* mean) and received 0.28 more vaccine doses (8.8%). These gains matter for child health because routine checkups catch developmental delays and illness early, and vaccines protect against the leading infectious causes of child death (Jones et al., 2003; Black et al., 2003). Neither facility visits nor vaccines were rewarded by the contract. Once families reached the health facility, they received its standard bundle of preventive services, so the output indicator crowded in care that the contract did not target. These findings indicate that parents used preventive health services more intensively, consistent with the coordinated referrals and accompaniment previously documented.

6.4 Iron Supplementation and Child Health

Iron supplement pick-ups doubled, rising by 0.36 bottles from a *Delivery-only* mean of 0.36 (98.1%; Table 5, Panel A). Each bottle covers one month of treatment, so even children in *Delivery and Output* municipalities received less than one month of iron over the five-month trial. Clinical trials, by contrast, require about three months of daily supplementation to raise hemoglobin by 0.56 g/dL (Gera et al., 2007). Scaling this benchmark by the additional supply implies an expected gain of at most 0.07 g/dL, even with perfect adherence. At observed adherence, where only 24.4% of children reach 90% adherence (INEI, 2023a), the expected gain falls to about 0.02 g/dL.

Panel B reports the effects on the probability that a child received a blood test for anemia screening and on health outcomes. There are no effects of the *Delivery and Output* contract on blood testing. Because hemoglobin levels and anemia diagnoses are only observed for tested children, this null result is important for the interpretation of the remaining variables. The absence of an effect on blood testing suggests that the *Delivery and Output* contract did not differentially affect selection into anemia screening. Consequently, estimated impacts on hemoglobin and anemia outcomes are unlikely to be driven by changes in testing behavior and instead reflect changes in underlying health status.

As the benchmark predicts, the estimated impacts on health outcomes are small in magnitude and statistically indistinguishable from zero across all indicators (Panel B). The point estimate on hemoglobin is −0.034 g/dL (s.e. 0.069), with a 95% confidence interval from −0.17 to 0.10 g/dL. This upper bound is less than one-fifth of the 0.56 g/dL that clinical trials achieve with enforced adherence (Section 2). The interval, however, also contains the gains of 0.02 to 0.07 g/dL implied by the additional iron supply, and the minimum detectable effect of 0.20 g/dL is roughly three times the upper end of that range. The experiment therefore lacks the power to detect changes in hemoglobin of the size the increase in iron implies. The effect on anemia

prevalence is likewise statistically indistinguishable from zero, with a point estimate of 0.003 (s.e. 0.020) against a *Delivery-only* mean of 0.195, ruling out reductions larger than 3.6 p.p. The effect on weight-for-age z-scores is similarly null, ruling out improvements larger than 0.18 standard deviations. The same pattern holds among children old enough to have been potentially exposed to the Home Visit Program for at least three months (Table A.14).

Since maternal nutrition and sanitation shape what supplementation can achieve, I examine heterogeneity by baseline district-level poverty and by access to drinking water and sanitation. I find no evidence of differential treatment effects along either dimension (Tables A.10 and A.11), suggesting that the absence of average effects does not appear to be driven by variation in these baseline conditions.

7 Results During the Reversion Phase

The RCT incentive scheme was originally designed to continue for a second semester beyond December 2024. However, an unexpected administrative complication in the publication of the legal norms regulating the new incentive period prevented its renewal, and all municipalities reverted to the *Delivery-only* contract from January to May 2025. This return to the status quo was unanticipated at all levels — neither the central government nor the municipalities had foreseen it. I use the five months following the end of the trial phase to examine whether the changes in managerial practices and health service utilization induced by the *Delivery and Output* contract persisted after its removal, and whether they translated into detectable improvements in child health outcomes.

7.1 Persistence of Managerial Practices

The managerial changes outlasted the contract. Table 7 repeats the specifications of Table 4 for the five months after the *Delivery and Output* contract ended. Managers did not revert once the output indicator disappeared from their contract. Panel A shows they remained 13.4 p.p. more likely to track the anemia indicator, nearly the same effect as the 13.7 p.p. estimated during the trial phase, and their supervision of CHWs shows no effect. Panel B shows the same pattern in referral management. Managers were still 11.0 p.p. more likely to strengthen care for children at risk. Health facility staff report that referrals remained 10.3 p.p. more likely to be coordinated with the facility, and that CHWs continued to accompany mothers (11.3 p.p.). Panel C shows managers still rewarding CHWs more (13.4 p.p.). Only the gain in CHW training fades: the estimate falls to 0.01 sessions, compared with 1.2 sessions (significant at the 10% level) during the trial phase.

Resources remained unchanged as well. Staffing and municipal expenditure show no effects in this period

either (Tables A.15 and A.16), with one exception: municipalities previously under the *Delivery and Output* contract had 0.32 more supervisors, an estimate significant only at the 10% level. Health facility staff report no change in supplies or stock-outs, but a 13.3 p.p. (38%) higher availability of trained personnel, also significant only at the 10% level (Table **??**).

7.2 Children Exposed During the Trial Phase

I first follow the children studied in Section 6.3 into the five months after the trial phase ended, when every municipality operated under the *Delivery-only* contract (Table 6). The utilization gains persist. Children in *Delivery and Output* municipalities received 0.53 more health facility visits (5.5%) and 0.42 more iron pick-ups, though the vaccine effect disappears. Health outcomes remain null. The hemoglobin estimate is 0.047 g/dL (s.e. 0.058), the anemia estimate is −0.009 (s.e. 0.021), and the weight-for-age estimate is 0.017 (s.e. 0.045). Fewer children were screened during this period, so the hemoglobin bound widens to 0.16 g/dL, still well below the 0.56 g/dL that supervised trials achieve. Five additional months of follow-up therefore leave the health null intact.

7.3 A New Cohort of Children

These managers also served a new cohort of younger children. Table 8 repeats the specifications of Table 5 for children born too late to be reached during the trial phase, whose families never experienced the *Delivery and Output* contract. The health service utilization gains did not fade. Children in municipalities previously assigned to the *Delivery and Output* contract attended 0.90 more health facility visits (10.2%). Iron pick-ups and vaccine doses show no differences across arms in this period. These families never experienced the *Delivery and Output* contract, so this persistence cannot reflect habits that households formed during the trial. It points instead to lasting changes in how managers run the program.

8 Conclusion

This paper examines whether a central government should reward local governments for service delivery, for the outputs these services aim to improve, or both. Using a clustered randomized controlled trial across 539 rural municipalities in Peru, I compare a *Delivery-only* contract with a *Delivery and Output* contract that adds an output indicator tied to child anemia. Municipalities assigned to the *Delivery and Output* contract monitored the anemia indicator more closely, provided more discretionary rewards to Community Health Workers, took more active steps to manage high-risk cases, and strengthened coordination with health facilities

— all without any increase in fiscal expenditure or staffing. This suggests that realigning managerial attention, rather than expanding resources, was the operative mechanism.

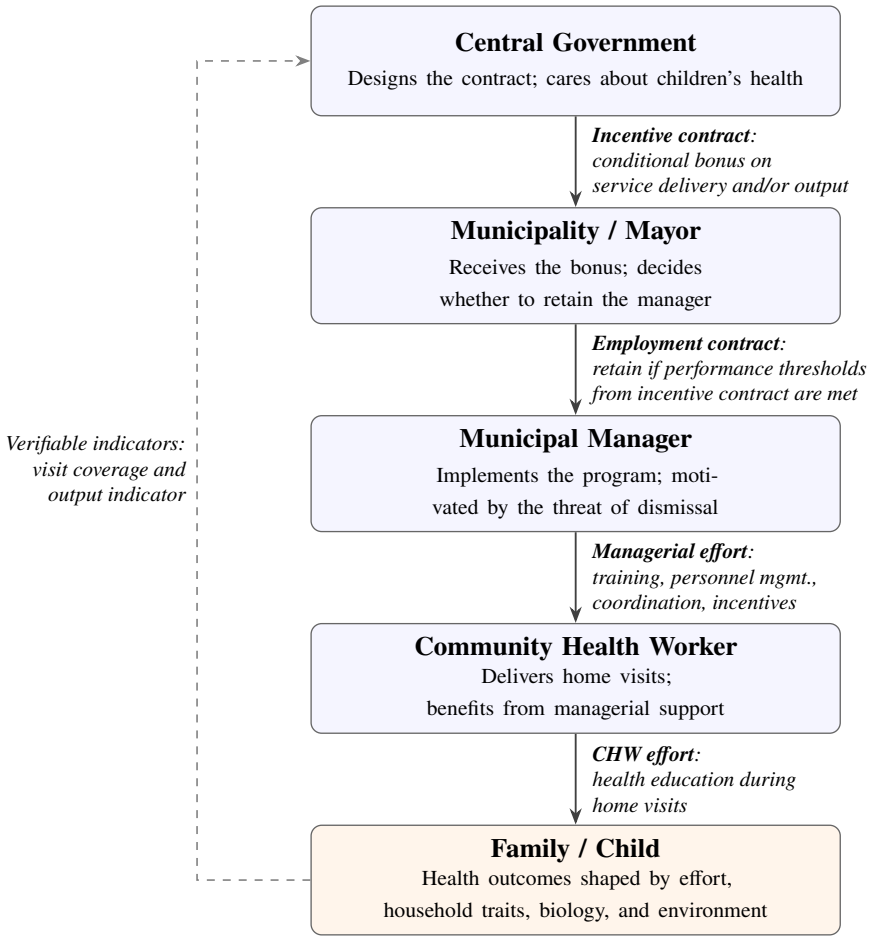
These managerial responses translated into measurable improvements in health service utilization. Children in *Delivery and Output* municipalities attended more health facility visits, collected iron supplements at higher rates, and received more vaccine doses during the intervention period. The breadth of these effects across multiple service dimensions points to a systematic improvement in community health outreach, consistent with CHWs becoming more effective at encouraging families to engage with the health system.

An unanticipated administrative delay in the publication of the legal norms governing the new incentive period caused all municipalities to revert to the status quo *Delivery-only* contract from January to May 2025, providing an opportunity to examine the durability of the effects of the *Delivery and Output* contract. The managerial changes induced during the trial phase proved largely persistent: municipalities previously exposed to the *Delivery and Output* contract continued to monitor anemia indicators more closely, support high-risk children more proactively, and reward CHWs at higher rates even after the contract ended. Persistence was uneven across practices, however — relationship-based behaviors such as accompanying mothers to health facilities and coordinating referrals with health facility staff remained strong, while more resource-intensive efforts such as additional CHW training dissipated once the incentive was removed. These durable managerial changes continued to translate into higher health service utilization, with children in previously treated municipalities attending more facility visits in the reversion phase, including a new cohort of children whose families never experienced the *Delivery and Output* contract.

Despite higher health service utilization both during and after the intervention, the study produced no detectable improvements in child health outcomes over the full observation window. Anemia prevalence, hemoglobin levels, and weight-for-age z-scores remained statistically indistinguishable from zero across all specifications. The contract doubled iron pick-ups, but from a low base: children received less than one month of iron over the five-month trial phase, which implies an expected hemoglobin gain of at most 0.07 g/dL. That is well below the minimum detectable effect of 0.20 g/dL, so the experiment lacks the power to detect health effects of the size the increase in iron implies. While these findings suggest that contract design is a viable lever for improving implementation quality — and that its effects on managerial behavior and service utilization can outlast the contract itself — they also caution against expecting short-run health returns from interventions targeting conditions shaped by cumulative and slow-moving biological processes.

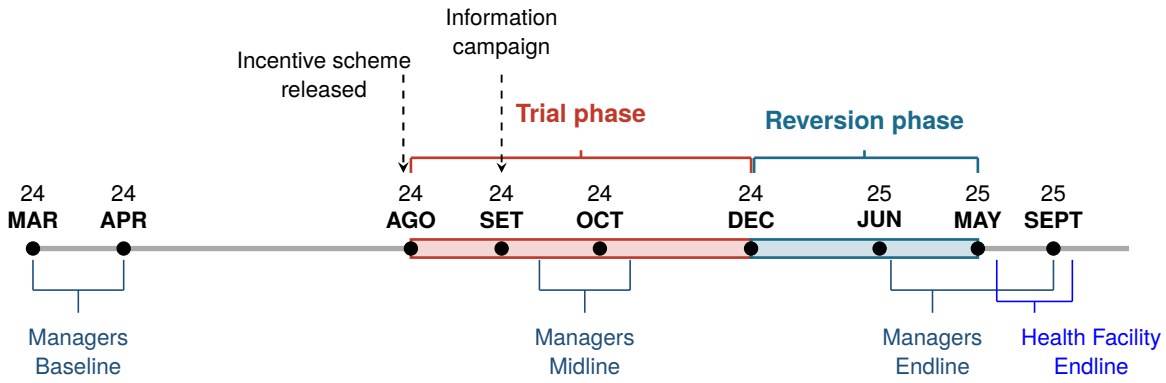
Tables and Figures

Figure 1: Institutional structure and flow of incentives



Notes: Solid arrows indicate contractual and hierarchical relations from principal to agent at each layer. The dashed arrow represents the verifiable indicators that the Central Government uses to evaluate municipal performance and disburse conditional transfers.

Figure 2: Research Timeline



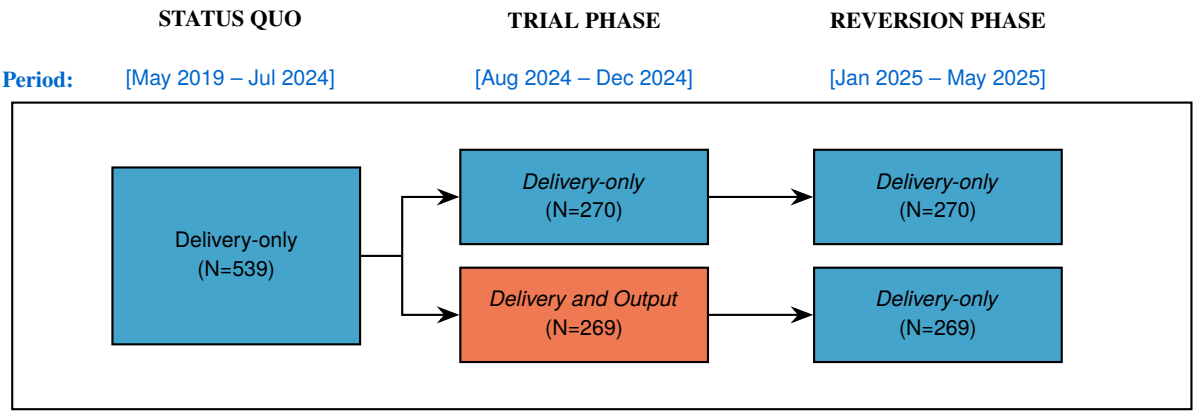
Notes: The figure summarizes the research timeline, including the three rounds of surveys to managers, one round of data collection to health facilities, the release of the incentive scheme in August 2024, the information campaign, and the randomized controlled trial (RCT) implementation period from August to December 2024. Unexpectedly, all municipalities reverted to the *Delivery-only* contract from January to May 2025 (reversion phase). This period is used to assess the persistence of results.

Table 1: Managers' Beliefs, Motivation, and Perceived Stakes

	Total (%)	Urban (%)	Rural (%)
Number of districts	877	245	632
Manager's primary motivation is population wellbeing	58.24	66.94	54.85
Manager's effort is key to program success	86.31	83.81	87.29
Home Visit is the most effective strategy to prevent anemia	47.37	55.92	44.04
Perceived consequences of missing performance targets			
Manager dismissed	59.11	66.26	56.32
Manager reassigned to another department	30.06	26.85	31.31
Supervisors dismissed	54.25	65.70	49.79
CHWs dismissed	37.48	41.49	35.92
Perceived key barrier to anemia recovery			
Biology	45.88	40.82	47.85
Economy	30.89	31.84	30.52
Culture	12.13	15.10	10.97
Other barrier	11.10	12.24	10.65

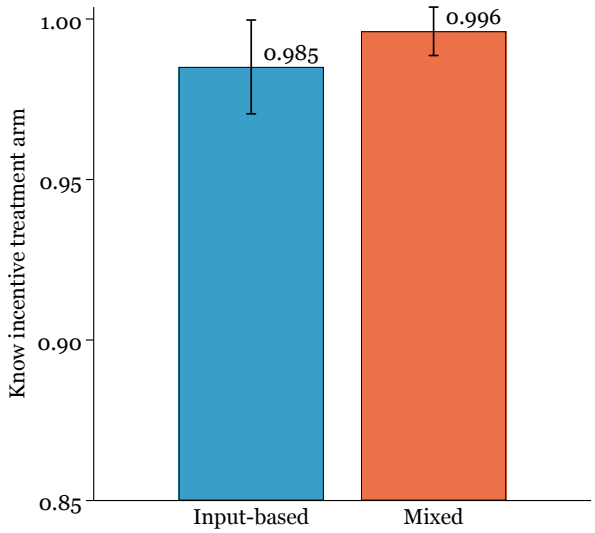
Notes: Manager's primary motivation is population wellbeing is a dummy equal to 1 if the manager reports that contributing to population wellbeing is their primary motivation for implementing the program. *Manager's effort is key to program success* is a dummy equal to 1 if the manager reports that program success depends on what they do. *Home Visit is the most effective strategy to prevent anemia* is a dummy equal to 1 if the manager indicates that home visits is the most effective strategy to prevent anemia over the alternative strategies offered. *Manager dismissed*, *Manager reassigned*, *Supervisors dismissed*, and *CHWs dismissed* are the expected probabilities that the respective position would face that consequence should the municipality fail to meet all performance indicators. *Perceived key barrier to anemia recovery* reports the share of managers selecting each option; the categories are mutually exclusive. Data come from the managers' baseline survey.

Figure 3: Experimental Design



Notes: This figure illustrates the experimental design. In the status quo (May 2019 to July 2024), all municipalities used to be incentivized only by the provision of service delivery (i.e., home visit coverage). During the trial phase (August to December 2024), municipalities were randomly assigned to either continue with the same scheme (N=270) or to switch to the *Delivery and Output* contract during the trial phase (N=269). In December 2024, an unanticipated administrative delay at the central government prevented the renewal of the *Delivery and Output* contract. Therefore, all municipalities reverted to the *Delivery-only* contract between January and May 2025 (reversion phase).

Figure 4: Knowledge of Treatment Status



Notes: This graph shows the proportion of managers who correctly identified the indicators on which their municipality would be evaluated under the incentive scheme. The information is based on the Midline Survey to Managers.

Table 2: Balance at Baseline

	Delivery-only incentive mean	Delivery & Output incentive mean	Diff (SE)	p-value of Diff
Number of Districts	**270**	**269**	**-**	**-**
Number of Health facilities	**1,507**	**1,494**	**-**	**-**
Number of Children	**7,303**	**6,501**	**-**	**-**
Panel A: District Characteristics				
Years of education (adults 18+)	7.25	7.21	-0.04 (0.07)	0.56
Total expenditure (USD per capita)	556.05	657.61	101.56 (66.15)	0.13
Expenditure in health (USD per capita)	17.17	15.28	-1.90 (5.80)	0.74
Connected to public water and sewage network	0.479	0.521	0.042 (0.044)	0.339
Poverty rate	0.528	0.487	-0.042 (0.044)	0.340
Panel B: Health Facility Characteristics				
Number of HF	5.581	5.554	-0.028 (0.441)	0.950
Health Facility Type				
Type I-1: No physician post	0.486	0.509	0.022 (0.027)	0.404
Type I-2: Physician post	0.341	0.324	-0.017 (0.027)	0.537
Type I-3: Polyclinic	0.148	0.139	-0.009 (0.019)	0.652
Type I-4: Inpatient facility	0.024	0.028	0.003 (0.008)	0.666
Distance from Type I-1 health facility to capital (km)	13.340	12.700	-0.640 (1.970)	0.745
Years since Type I-1 health facility opened	29.589	22.398	-7.191 (3.093)	0.021**
Panel C: Home Visit Program Characteristics				
Managers' Characteristics				
Age	33.18	33.26	0.08 (0.70)	0.91
Dummy if male	0.38	0.50	0.12 (0.04)	0.01***
Has higher education	0.30	0.29	-0.01 (0.04)	0.87
Years of experience in the HV Program	1.81	1.94	0.12 (0.15)	0.40
Supervisors' Characteristics				
Number of supervisors	0.65	0.57	-0.08 (0.14)	0.58
Has higher education	0.28	0.30	0.02 (0.04)	0.69
Community Health Workers' Characteristics				
Number of CHW	11.08	12.05	0.97 (1.13)	0.39
Has higher education	0.16	0.15	-0.01 (0.02)	0.79
Panel D: Children characteristics				
Number of HF visits	11.85	12.61	-0.62 (0.40)	0.12
Number of iron pick-ups	2.38	2.56	-0.13 (0.10)	0.18
Number of vaccine doses	5.46	5.70	-0.14 (0.12)	0.23

Notes: Panel A reports district-level socioeconomic characteristics. Panel B reports characteristics of the health facilities located in each district, including facility type, distance to the district capital, and years of operation. Panel C summarizes characteristics of managers, supervisors, and Community Health Workers involved in the Home Visit Program. Panel D presents baseline child health and service-use indicators. All variables are measured at baseline. Differences come from a regression of each variable on the treatment indicator, with standard errors clustered at the municipality level reported in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Table 3: Effects on home visit coverage

	Delivery & Output	Delivery-only mean	Effect (%)	N
Children: visit coverage	0.005	0.880	0.6	539
	(0.009)			
Children: met partial-bonus threshold	-0.004	0.748	-0.5	539
	(0.021)			
Pregnant women: visit coverage	-0.008	0.941	-0.9	539
	(0.015)			
Pregnant women: met partial-bonus threshold	-0.004	0.878	-0.5	539
	(0.028)			
Output indicator	0.011	0.811	1.4	539
	(0.011)			

Notes: Estimates come from municipality-level regressions. Standard errors clustered at the municipality level are in parentheses. *Visit coverage* is the share of eligible children or pregnant women who received timely home visits; *met partial-bonus threshold* equals 1 if the municipality reached at least the partial-bonus threshold for that indicator. The children's specifications control for the baseline value of the dependent variable; the pregnant women's specifications do not, as no baseline is available. No other controls are included. Statistical significance: $*p < 0.1$, $**p < 0.05$, $***p < 0.01$ (unadjusted); $\star p < 0.1$, $\star\star p < 0.05$, $\star\star\star p < 0.01$ (MHT-adjusted). All outcomes belong to a single MHT family.

Table 4: Effects on Managerial Practices During the Trial Phase

	Delivery & Output	Delivery-only mean	Effect (%)	N
Panel A. Monitoring				
Monitored anemia KPI indicator	$0.137^{***}_{\star\star\star}$	0.723	18.949	442
	(0.038)			
Monitored CHWs daily or biweekly	$0.046_{\star}$	0.589	7.810	442
	(0.036)			
Panel B. Referral management				
Improved care for children at risk of anemia	$0.124^{***}_{\star\star\star}$	0.509	24.361	442
	(0.047)			
Referral is coordinated with health facility	$0.195^{***}_{\star\star\star}$	0.235	82.979	397
	(0.052)			
CHWs take mothers to health facility	$0.144^{**}_{\star\star}$	0.261	55.172	397
	(0.058)			
Panel C. Motivation of CHWs				
CHW was incentivized	$0.171^{***}_{\star\star\star}$	0.371	46.092	442
	(0.047)			
Panel D. Training CHWs				
Number of CHW training sessions	$1.194^{*}_{\star\star}$	3.797	31.446	397
	(0.697)			

Notes: Estimates come from municipality-level and health facility-level regressions. Standard errors clustered at the municipality level in parentheses. Outcomes in Panels A and C and in the first row of Panel B come from the managers' survey and are measured at the municipality level (N = 442); the remaining rows come from the health facilities' survey and are measured at the facility level (N = 397). *Effect (%)* reports the point estimate as a percentage of the *Delivery-only* mean. Only variables in Panel A control for the baseline value of the dependent variable; the remaining specifications do not, as no baseline is available. No other controls are included. Statistical significance: $*p < 0.1$, $**p < 0.05$, $***p < 0.01$ (unadjusted); $\star p < 0.1$, $\star\star p < 0.05$, $\star\star\star p < 0.01$ (MHT-adjusted). All outcomes belong to a single MHT family.

Table 5: Effects on Health Service Utilization and Health Outcomes during the Trial Phase
(Children exposed for at least 1 month during the trial phase)

	Delivery & Output	Delivery-only mean	Effect (%)	*N*
Panel A. Health service utilization				
Number of HF visits	$0.367^{**}_{\star}$	9.514	3.9	13,804
	(0.172)			
Number of iron pick-ups	$0.355^{**}_{\star}$	0.362	98.1	13,804
	(0.174)			
Number of vaccine doses	$0.283^{*}_{\star}$	3.215	8.8	13,804
	(0.165)			
Panel B. Health outcomes				
Weight-for-Age Z-score (WAZ)	0.053	0.439	12.1	12,505
	(0.066)			
Took blood test	0.020	0.857	2.3	13,804
	(0.012)			
Anemic	0.003	0.195	1.5	11,573
	(0.020)			
Hemoglobin (g/dL)	-0.034	12.600	-0.3	11,573
	(0.069)			

Notes: This table reports the effects of the *Delivery and Output* contract on health service utilization and child health outcomes for the children who were exposed at least 1 month to the intervention during the trial phase and before the blood testing. Exposure duration is determined by birth cohort. Each row is a separate regression estimated at the child level. Panel A reports the total number of health facility (HF) visits, iron supplement pick-ups, and vaccine doses received during the study period. According to the national immunization schedule, children aged 2–12 months are scheduled to receive up to 17 routine vaccine doses. In Panel B, *Took blood test* indicates whether the child received a blood test for anemia screening; *Hemoglobin (g/dL)* is the hemoglobin concentration measured during routine anemia screening; *Anemic* equals one if the child is classified as anemic based on altitude-adjusted hemoglobin thresholds. Effect (%) reports the point estimate as a percentage of the *Delivery-only* mean. All estimates control for the outcome at baseline, except *Took blood test*. Cohort fixed effects are included throughout, and the *Anemic* regression additionally includes linear and quadratic controls for altitude. Standard errors clustered at the municipality level are in parentheses. Statistical significance: $*$ $p < 0.1$, $**$ $p < 0.05$, $***$ $p < 0.01$ (unadjusted); $\star$ $p < 0.1$, $\star\star$ $p < 0.05$, $\star\star\star$ $p < 0.01$ (MHT-adjusted). When correcting the p-values for multiple hypothesis testing, I group the outcomes in two MHT families, one per panel.

Table 6: Effects on Health Service Utilization and Health Outcomes during the Reversion Phase
(Children exposed for at least 1 month during the trial phase)

	Delivery & Output	Delivery-only mean	Effect (%)	*N*
Panel A. Health service utilization				
Number of HF visits	0.526*	9.514	5.5	13,804
	(0.290)			
Number of iron pick-ups	0.419*	0.362	115.7	13,804
	(0.238)			
Number of vaccines doses	0.070	3.215	2.2	13,804
	(0.045)			
Panel B. Health outcomes				
Weight-for-Age Z-score (WAZ)	0.017	0.425	4.0	12,952
	(0.045)			
Took blood test	0.020	0.624	3.2	13,804
	(0.011)			
Anemic	-0.009	0.265	-3.4	8,493
	(0.021)			
Hemoglobin (g/dL)	0.047	12.266	0.4	8,493
	(0.058)			

Notes: This table reports the effects of the *Delivery and Output* contract on health service utilization and child health outcomes for the children who were exposed at least 1 month to the intervention before the blood testing. Exposure duration is determined by birth cohort. Each row is a separate regression estimated at the child level. Panel A reports the total number of health facility (HF) visits, iron supplement pick-ups, and vaccine doses received during the study period. According to the national immunization schedule, children aged 2–12 months are scheduled to receive up to 17 routine vaccine doses. In Panel B, *Took blood test* indicates whether the child received a blood test for anemia screening; *Hemoglobin (g/dL)* is the hemoglobin concentration measured during routine anemia screening; *Anemic* equals one if the child is classified as anemic based on altitude-adjusted hemoglobin thresholds. *Effect (%)* reports the point estimate as a percentage of the *Delivery-only* mean. Estimates control for the outcome at baseline, except *Took blood test*. Cohort fixed effects are included throughout, and the *Anemic* regression additionally includes linear and quadratic controls for altitude. Standard errors clustered at the municipality level are in parentheses. Statistical significance: $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$ (unadjusted); $\star \ p < 0.1$, $\star\star \ p < 0.05$, $\star\star\star \ p < 0.01$ (MHT-adjusted). When correcting the p-values for multiple hypothesis testing, I group the outcomes in two MHT families, one per panel.

Table 7: Effects on Managerial Practices During the Reversion Phase

	Delivery & Output	Delivery-only mean	Effect (%)	N
Panel A. Monitoring				
Monitored anemia KPI indicator	$0.134^{***}_{\star\star}$	0.722	18.6	352
	(0.042)			
Monitored CHWs daily or biweekly	0.003	0.550	0.5	352
	(0.044)			
Panel B. Referral management				
Improved care for children at risk of anemia	$0.110^{**}_{\star\star}$	0.606	18.2	352
	(0.050)			
Referral with HF is coordinated	$0.103^{*}_{\star}$	0.205	50.2	237
	(0.056)			
CHWs take mothers to HF	$0.113^{**}_{\star\star}$	0.154	73.4	237
	(0.048)			
Panel C. Motivation of CHWs				
CHW was incentivized	$0.134^{**}_{\star\star}$	0.372	36.0	352
	(0.053)			
Panel D. Training				
Number of CHW training sessions	0.010	4.923	0.2	237
	(0.714)			

Notes: Estimates come from municipality-level and health facility-level regressions. Standard errors clustered at the municipality level in parentheses. Outcomes in Panels A and C and in the first row of Panel B come from the managers' survey and are measured at the municipality level (N = 352); the remaining rows come from the health facilities' survey and are measured at the facility level (N = 237). *Effect (%)* reports the point estimate as a percentage of the *Delivery-only* mean. Only variables in Panel A control for the baseline value of the dependent variable; the remaining specifications do not, as no baseline is available. No other controls are included. Statistical significance: $*p < 0.1$, $**p < 0.05$, $***p < 0.01$ (unadjusted); $\star p < 0.1$, $\star\star p < 0.05$, $\star\star\star p < 0.01$ (MHT-adjusted). All outcomes belong to a single MHT family.

Table 8: Effects on Health Service Utilization and Health Outcomes during the Reversion Phase
(Children exposed for at least 1 month during the reversion phase)

	Delivery & Output	Delivery-only mean	Effect (%)	*N*
Number of HF visits	$0.897^{***}_{\star\star\star}$	8.752	10.2	12,797
	(0.260)			
Number of iron pick-ups	0.437	6.886	6.3	12,797
	(0.272)			
Number of vaccine doses	0.038	7.113	0.5	12,797
	(0.150)			

Notes: This table reports the effects of the *Delivery and Output* contract on health service utilization for the children who were exposed at least 1 month to the Home Visit Program during the reversion phase and before the blood testing. Exposure duration is determined by birth cohort. Each row is a separate regression estimated at the child level. The table reports the total number of health facility (HF) visits, iron supplement pick-ups, and vaccine doses received during the study period. According to the national immunization schedule, children aged 2–12 months are scheduled to receive up to 17 routine vaccine doses. *Effect (%)* reports the point estimate as a percentage of the *Delivery-only* mean. Estimates control for the outcome at baseline. Cohort fixed effects are included throughout.Standard errors clustered at the municipality level are in parentheses. Statistical significance: $*$ $p < 0.1$, $**$ $p < 0.05$, $***$ $p < 0.01$ (unadjusted); $\star$ $p < 0.1$, $\star\star$ $p < 0.05$, $\star\star\star$ $p < 0.01$ (MHT-adjusted). When correcting the p-values for multiple hypothesis testing, I group the outcomes in two MHT families, one per panel.

Appendix

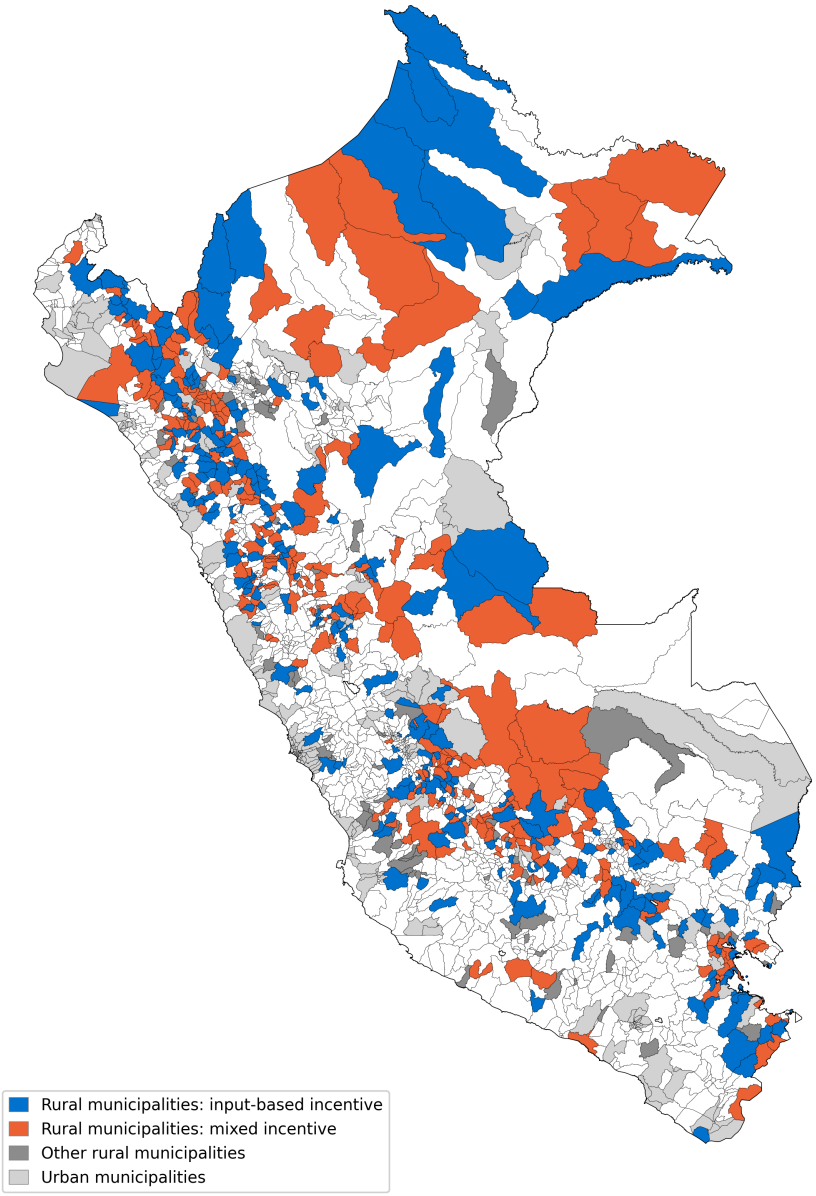
A. Appendix – Additional Tables and Figures

Table A.1: Municipality classification (2024)

Municipality level	**Type of Municipalities**			**Number of municipalities**	**HV Program**
Provincial	Belong to a main city		A	74	Yes
	Does not belong to a main city		B	122	
District	Belong to Metropolitan Lima		C	42	Yes
	Belong to other main cities		D	129	Yes
	Does not belong to a main city	Urban population ≤ 70%	E	378	
		35% ≤ Urban population ≤ 70%	F	509	
		Urban population ≤ **35%**	G	637	Yes
			Total	1,891	

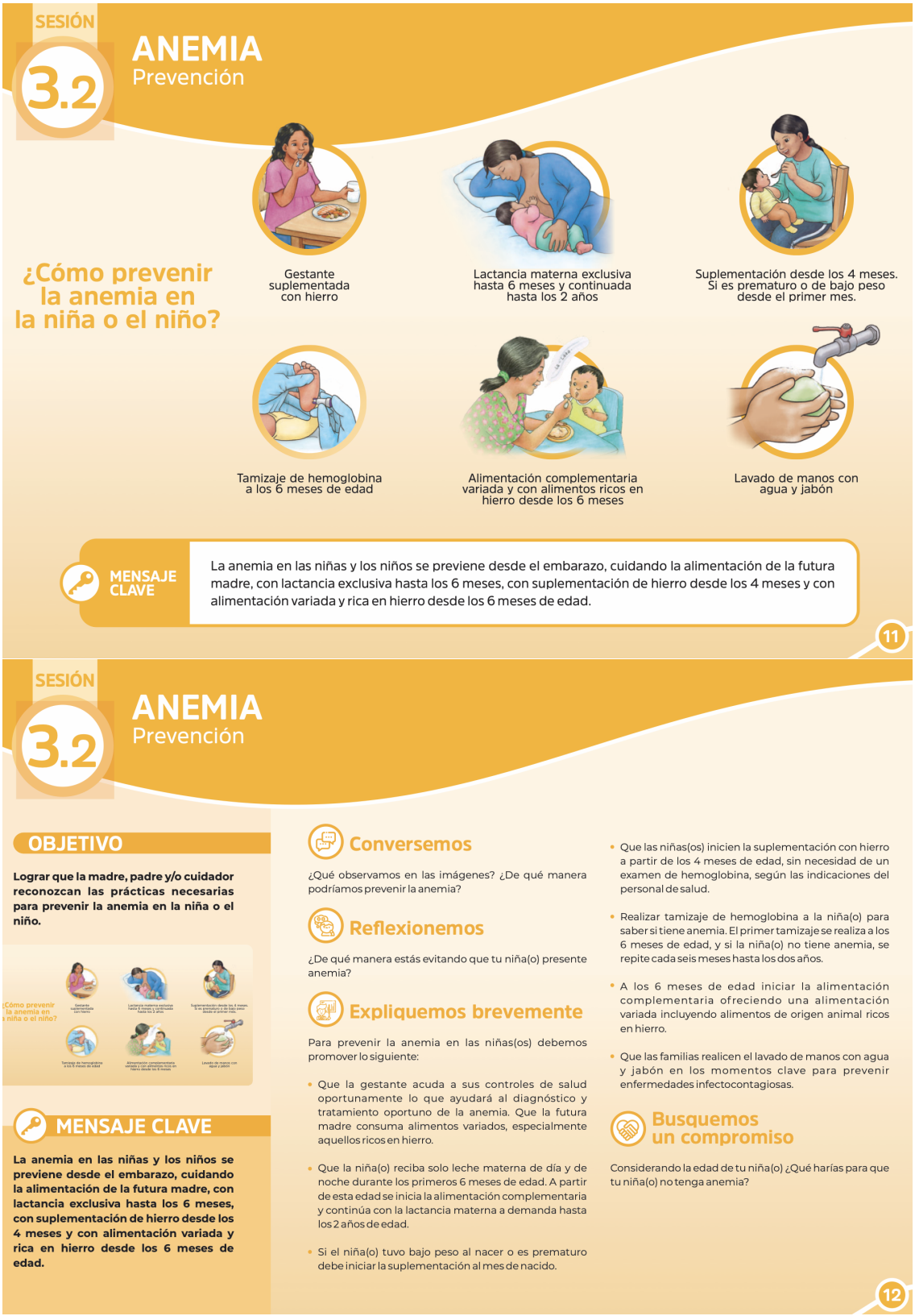
Notes: The table reports the municipality classification used by the Municipal Incentive Program in 2024. Municipalities are classified by level (provincial or district), by whether they belong to a main city, and by their share of urban population. Main cities are Metropolitan Lima and the remaining principal urban centers. The last column indicates whether the Home Visit Program operates in each type. The Program covers types A, C, and D, which are the urban municipalities of main cities, and type G, the most rural municipalities. *Source:* Ministry of Economy and Finance.

Figure A.1: Geographic Distribution of Municipalities



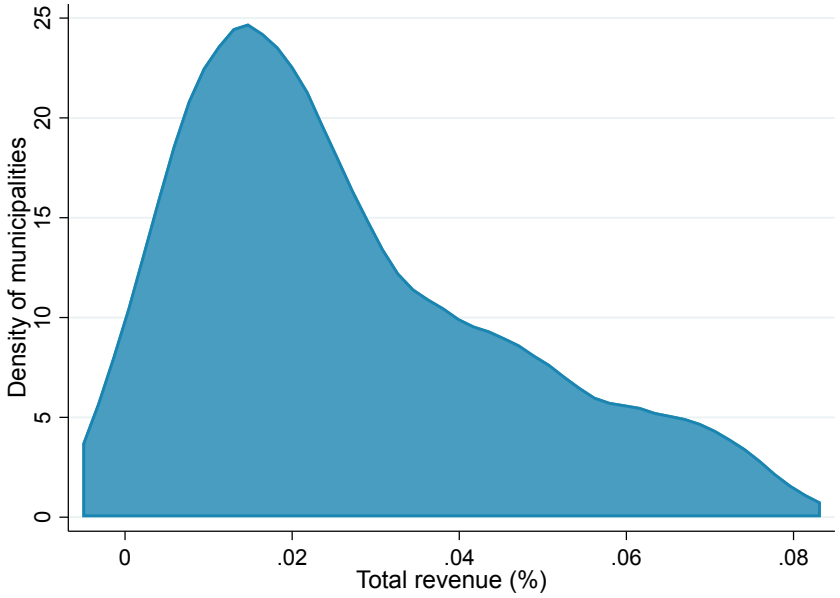
Notes: The figure shows the geographic distribution of Peruvian districts by Home Visit Program status and treatment arm. Blue districts are rural municipalities in the RCT assigned to the *Delivery-only* contract ($N = 270$). Orange districts are rural municipalities in the RCT assigned to the *Delivery and Output* contract ($N = 269$). Dark gray districts are rural municipalities that implement the Program but were excluded from the RCT because they had fewer than 15 eligible children. Light gray are urban municipalities, which implement the Program but were assigned to the *Delivery and Output* contract by government decision and were therefore not randomized. Districts in white do not implement the Home Visit Program. Program coverage and treatment assignment were established by Supreme Decree No. 072-2019-EF and RD 0027-2024-EF/50.01

Figure A.2: Flipchart example of one session



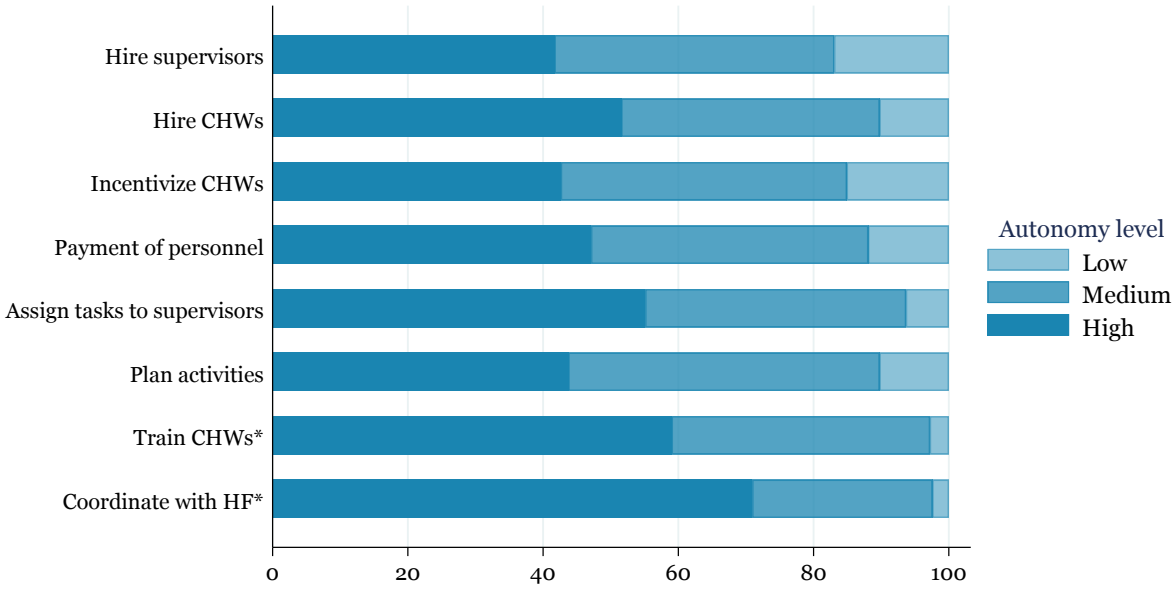
Notes: This content corresponds to Session 3.2 from the 2024 Home Visit Program's flipchart (pages 11 and 12). The entire flipchart can be found here.

Figure A.3: Ratio of potential incentives to total budget



Notes: The ratio of potential incentives to the total municipal budget is defined as the total amount a municipality could receive from the Municipal Incentive Program for perfect compliance in 2024 divided by its total budget in 2024. The Home Visit Program is implemented in 882 municipalities. For visualization purposes, the figure excludes the upper 10% of the distribution; municipalities with ratios above 200 are therefore not shown.

Figure A.4: Level of Autonomy of Managers and Supervisors in Implementing the Program.



Notes: This figure shows self-reported levels of autonomy (low, medium, high) across key human-resource, planning, and coordination functions related to CHWs. For each function, managers were asked to indicate the degree of autonomy they and the supervisors exercise using three predefined categories: low, medium, or high autonomy. *Hire supervisors/CHWs* refers to authority to recruit and select staff. *Incentivize CHWs* captures autonomy to set incentive structures. *Payment of personnel* reflects control over payment modalities. *Assign tasks to supervisors* refers to allocating responsibilities within the supervisory team. *Plan activities* captures autonomy in defining activities and budgeting. *Train CHWs* indicates authority over training content. *Coordinate with HF* reflects decision-making autonomy in coordination with local health facilities.

Figure A.5: Assigned Performance Thresholds by Municipality and Indicator

METAS DEL PROGRAMA DE INCENTIVOS A LA MEJORA DE LA GESTIÓN MUNICIPAL (PI) QUE DEBEN CUMPLIR LAS MUNICIPALIDADES EN EL AÑO 2024 - TRAMO II, COMPROMISO 1

Compromiso 1: "Mejora del estado nutricional y de salud de las gestantes y niños hasta los 12 meses de edad para la prevención de anemia"

N°	UBIGEO	DEPARTAMENTO	PROVINCIA	DISTRITO	Indicador 1.1		Indicador 1.2		Indicador 1.3	
					Umbral mínimo	Meta máxima	Umbral mínimo	Meta máxima	Umbral mínimo	Meta máxima
1	010101	AMAZONAS	CHACHAPOYAS	CHACHAPOYAS	29%	43%	85%	92%	87%	92%
2	010103	AMAZONAS	CHACHAPOYAS	BALSAS	-	-	85%	85%	87%	92%
3	010105	AMAZONAS	CHACHAPOYAS	CHILIQUIN	-	-	85%	92%	87%	92%
4	010106	AMAZONAS	CHACHAPOYAS	CHUQUIBAMBA	25%	39%	85%	85%	87%	92%
5	010114	AMAZONAS	CHACHAPOYAS	MOLINOPAMPA	-	-	85%	92%	87%	92%
6	010117	AMAZONAS	CHACHAPOYAS	QUINJALCA	-	-	85%	85%	87%	92%
7	010120	AMAZONAS	CHACHAPOYAS	SOLOCO	-	-	85%	92%	87%	92%
8	010201	AMAZONAS	BAGUA	BAGUA	30%	44%	85%	92%	87%	92%
9	010202	AMAZONAS	BAGUA	ARAMANGO	-	-	85%	92%	87%	92%
10	010205	AMAZONAS	BAGUA	IMAZA	-	-	85%	92%	87%	92%
11	010304	AMAZONAS	BONGARA	COROSHA	-	-	85%	92%	87%	92%
12	010308	AMAZONAS	BONGARA	RECTA	-	-	85%	92%	87%	92%
13	010311	AMAZONAS	BONGARA	VALERA	-	-	85%	92%	87%	92%
14	010402	AMAZONAS	CONDORCANQUI	EL CENEPA	-	-	85%	92%	87%	92%
15	010403	AMAZONAS	CONDORCANQUI	RIO SANTIAGO	-	-	85%	92%	87%	92%
16	010503	AMAZONAS	LUYA	COCABAMBA	-	-	85%	85%	87%	92%
17	010507	AMAZONAS	LUYA	LONGUITA	-	-	85%	92%	87%	92%
18	010513	AMAZONAS	LUYA	OCUMAL	-	-	85%	92%	87%	92%
19	010514	AMAZONAS	LUYA	PISUQUIA	-	-	85%	85%	87%	92%
20	010515	AMAZONAS	LUYA	PROVIDENCIA	50%	64%	85%	92%	87%	92%
21	010518	AMAZONAS	LUYA	SAN JERONIMO	-	-	85%	92%	87%	92%
22	010519	AMAZONAS	LUYA	SAN JUAN DE LOPECANCHA	-	-	85%	85%	87%	92%
23	010520	AMAZONAS	LUYA	SANTA CATALINA	-	-	85%	92%	87%	92%
24	010521	AMAZONAS	LUYA	SANTO TOMAS	-	-	85%	92%	87%	92%
25	010602	AMAZONAS	RODRIGUEZ DE MENDOZA	CHIRIMOTO	-	-	85%	85%	87%	92%
26	010604	AMAZONAS	RODRIGUEZ DE MENDOZA	HUAMBO	-	-	85%	92%	87%	92%
27	010605	AMAZONAS	RODRIGUEZ DE MENDOZA	LIMABAMBA	-	-	85%	92%	87%	92%
28	010609	AMAZONAS	RODRIGUEZ DE MENDOZA	OMIA	14%	27%	85%	92%	87%	92%
29	010610	AMAZONAS	RODRIGUEZ DE MENDOZA	SANTA ROSA	-	-	85%	92%	87%	92%

Notes: This figure reproduces an excerpt from the official legal norm (RD 0027-2024-EF/50.01) published by the Ministry of Economy and Finance in July 2024, which formalized the incentive scheme for the period August–December 2024. The table lists the performance floors (*umbral mínimo*) and ceilings (*meta máxima*) assigned to each participating municipality for each of the three indicators under the Home Visit Program *Commitment 1*: Indicator 1.1 (share of children under 12 months receiving home visits), Indicator 1.2 (share of pregnant women receiving home visits), and Indicator 1.3 (share of children who are non-anemic or have maintained or increased their hemoglobin levels between the six- and twelve-month screenings). The full document is available here.

Table A.2: Baseline Balance for the 448 Municipalities Re-Surveyed at Endline

	Delivery-only incentive mean	Delivery & Output incentive mean	Diff (SE)	p-value of Diff
Number of Districts	**226**	**222**	**-**	**-**
Number of Children	**6,376**	**5,558**	**-**	**-**
Number of Health facilities	**1,272**	**1,263**	**-**	**-**
Panel A: District Characteristics				
Population's years of education	7.24	7.20	-0.04 (0.08)	0.63
Total expenditure (USD per capita)	576.06	652.04	75.98 (77.78)	0.33
Expenditure in health (USD per capita)	17.13	16.73	-0.40 (6.60)	0.95
Connected to public water and sewage network	0.473	0.495	0.022 (0.048)	0.640
Poverty rate	0.527	0.491	-0.036 (0.048)	0.449
Panel B: Health Facility (HF) Characteristics				
Number of HF	5.652	5.720	0.068 (0.503)	0.892
Health Facility Type				
Type I-1: No physician post	0.476	0.492	0.016 (0.030)	0.602
Type I-2: Physician post	0.351	0.335	-0.016 (0.030)	0.592
Type I-3: Polyclinic	0.147	0.144	-0.003 (0.021)	0.892
Type I-4: Inpatient facility	0.026	0.030	0.004 (0.009)	0.656
Distance from Type I-1 HF to capital (km)	12.958	13.266	0.308 (2.315)	0.894
Years since HF opened	32.839	28.137	-4.702 (2.101)	0.026
Panel C: Home Visit Program Characteristics				
Managers' Characteristics				
Age	33.49	33.23	-0.26 (0.79)	0.74
Dummy if male	0.39	0.49	0.10 (0.05)	0.03**
Has higher education	0.31	0.29	-0.02 (0.04)	0.66
Total years of experience	8.33	9.22	0.90 (0.90)	0.32
Years of experience in the HV Program	1.89	1.97	0.08 (0.16)	0.63
Supervisors' Characteristics				
Number of supervisors	0.65	0.56	-0.09 (0.15)	0.56
Has higher education	0.29	0.29	0.00 (0.04)	0.98
Community Health Workers' Characteristics				
Number of CHW	11.67	12.46	0.79 (1.27)	0.54
Has higher education	0.17	0.14	-0.02 (0.03)	0.39
Panel D: Children characteristics				
Number of HF visits	11.84	12.44	-0.40 (0.43)	0.35
Number of iron pick-ups	2.39	2.48	-0.04 (0.09)	0.70
Number of vaccine doses	5.47	5.72	-0.10 (0.13)	0.42

Notes: Panel A reports district-level socioeconomic characteristics. Panel B reports characteristics of the health facilities located in each district, including facility type, distance to the district capital, and years of operation. Panel C summarizes characteristics of managers, supervisors, and Community Health Workers involved in the Home Visit Program. Panel D presents baseline child health and service-use indicators. All variables are measured at baseline. Differences come from a regression of each variable on the treatment indicator, with standard errors clustered at the municipality level reported in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Table A.3: Balance Table - Primary Care Health Facilities

	(1) N	(2) Delivery-only incentive mean	(3) Delivery & Output incentive mean	(4) Diff (SE)	(5) p-value of Diff
District with ≥ 1 surveyed HF	539	0.604	0.613	0.010	0.818
				(0.042)	
Number of surveys per district	539	1.641	2.093	0.452	0.079
				(0.257)	
Proportion of surveyed HF	539	0.415	0.424	0.009	0.833
				(0.043)	
Health Facility Characteristics					
Distance to district capital (km)	330	8.883	27.252	18.369	0.328
				(18.736)	
Age (years)	290	26.817	25.129	-1.687	0.162
				(1.203)	
Health Facility Type:					
Type I-1: No physician post	330	0.387	0.429	0.041	0.313
				(0.041)	
Type I-2: Physician post	330	0.270	0.251	-0.019	0.630
				(0.040)	
Type I-3: Polyclinic	330	0.207	0.142	-0.065	0.054
				(0.034)	
Type I-4: Inpatient facility	330	0.049	0.044	-0.006	0.784
				(0.021)	

Note: Primary care health facilities range from basic health posts without physicians (Type I-1) to facilities with medical doctors (Type I-2), comprehensive health centers and polyclinics (Type I-3), and centers offering short-term inpatient care (Type I-4).

Table A.4: Understanding of the output indicator

	Delivery-only Mean (SD) (1)	Delivery and Output Mean (SD) (2)	Diff (1)-(2)
It is a composite indicator	0.810	0.814	0.913
	(0.393)	(0.390)	
Objective	0.656	0.686	0.468
	(0.476)	(0.465)	
Deductions to denominator	0.306	0.341	0.390
	(0.462)	(0.475)	
Deductions to numerator	0.164	0.257	0.009
	(0.371)	(0.438)	

Notes: Composite indicator is a dummy variable equal to 1 if the supervisor understands that the output indicator is a composite measure. *Objective* is a dummy variable equal to 1 if the supervisor knows the performance objective for the anemia indicator. *Deductions to denominator* is a dummy variable equal to 1 if the supervisor knows the rules governing the denominator of the Child Home Visit Indicator. *Deductions to numerator* is a dummy variable equal to 1 if the supervisor validates cases where a child in the denominator does not contribute to the numerator for the Child Home Visit Indicator. These statistics only consider rural municipalities. The information is drawn from the midline manager survey.

Table A.5: Municipality Attrition at Endline

	Attrited (1 = not re-surveyed)
Delivery and Output	0.027
	(0.032)
Delivery-only mean	0.156
Observations	539

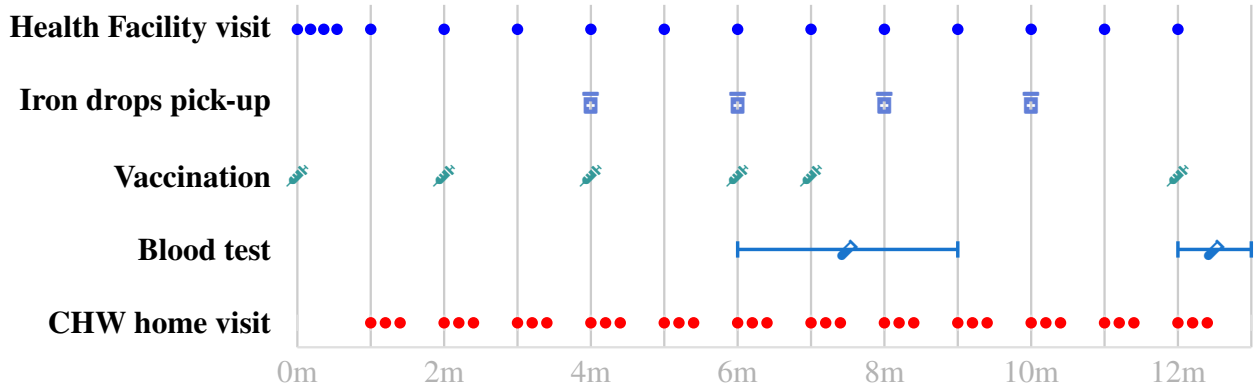
Notes: The dependent variable is an indicator =1 if the municipality could not be re-surveyed at endline. Regression coefficients are based on an OLS model. Standard errors are clustered at the municipality level in parentheses. Statistical significance is indicated by $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A.6: Personnel Turnover by Contract

	(1) Manager/supervisor	(2) Health facility staff
Delivery and Output	0.012	0.000
	(0.007)	(0.002)
Delivery-only mean	0.006	0.002
Magnitude of the effect (%)	200.0	0.3
Observations	934	1,749

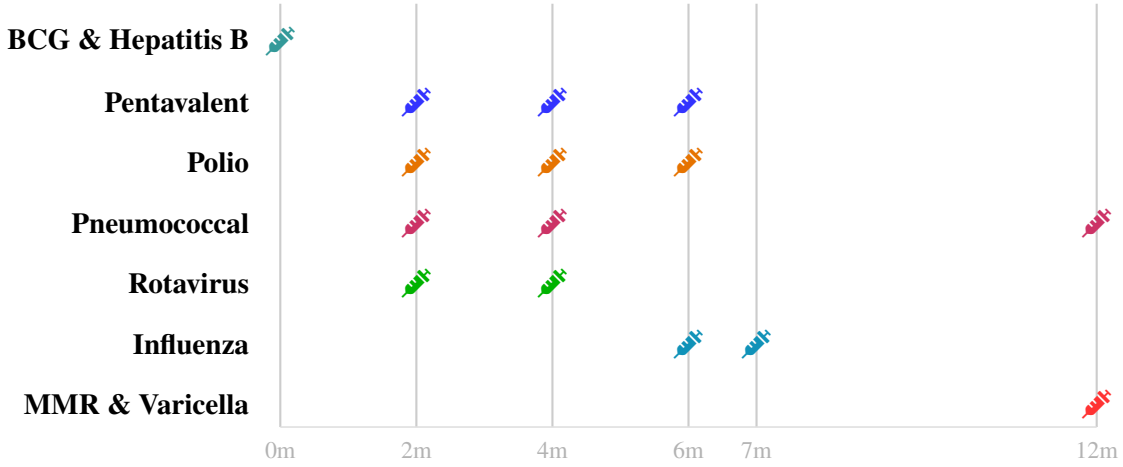
Notes: This table assesses whether assignment to the *Delivery and Output* contract affected personnel turnover during the trial phase. The unit of observation is the individual and includes municipal personnel (managers and supervisors) as well as health facility staff responsible for delivering anemia-related services to the Home Visit Program's target population. Turnover is measured using administrative employment records and is coded as one if an individual appears as employed in December 2023 but not in December 2024. Estimates are based on an OLS regression, with standard errors clustered at the municipality level. Statistical significance is indicated by $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$.

Figure A.6: Schedule of Health Facility Services by child's age (in months)



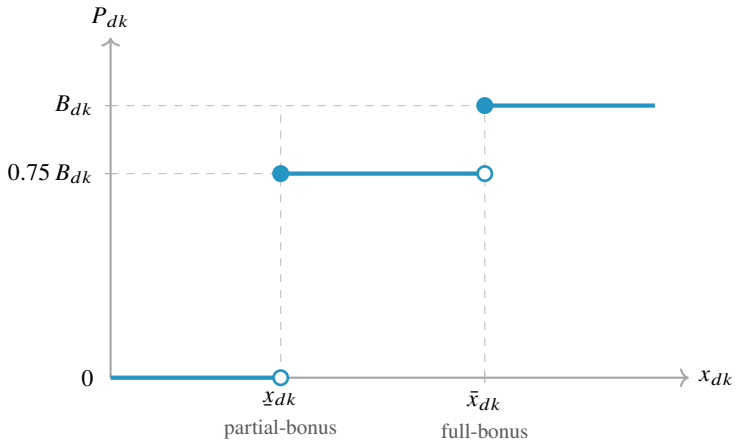
Notes: This figure illustrates the timing of routine health-facility services and community health worker (CHW) home visits during a child's first year of life. Children attend growth-monitoring and counseling visits at health facilities four times during the first month of life and monthly thereafter through month 11. Vaccinations are scheduled at specific ages in accordance with national guidelines (at birth and at 2, 4, 6, 7, and 12 months). Anemia screening through blood tests is conducted at two points: between 6 and 9 months of age and again between 12 and 13 months. Iron supplementation is provided preventively at month 4 and reactively at months 6, 8, and 10, conditional on anemia status. CHW home visits last approximately 20 minutes and are conducted up to three times per month between months 1 and 11 (up to 33 visits in total). These visits are designed to complement facility-based care by reinforcing adherence to preventive services, providing counseling, and promoting timely utilization of health services.

Figure A.7: Vaccination Schedule



Notes: This figure presents the national vaccination schedule established by national guidelines for children from birth through 12 months of age in Peru. At birth, children receive BCG and Hepatitis B vaccines. During early infancy, Pentavalent, Polio, Pneumococcal, and Rotavirus vaccines are administered at 2 and 4 months, with additional doses of Pentavalent and Polio at 6 months. Influenza vaccines are administered at 6 and 7 months. At 12 months of age, children receive the Measles–Mumps–Rubella (MMR) and Varicella vaccines, as well as the final dose of the Pneumococcal vaccine.

Figure A.8: Bonus Payment Schedule for a Single Indicator



Notes: The figure plots the payment function in Equation (1). It shows the bonus P_{dk} that municipality d receives for indicator k as a function of its performance x_{dk}. A municipality earns nothing if performance falls below the partial-bonus threshold $\underline{x}_{dk}$, 75 percent of the indicator bonus B_{dk} if performance reaches that threshold but stays below the full-bonus threshold $\bar{x}_{dk}$, and the full bonus B_{dk} once performance reaches $\bar{x}_{dk}$. Performance above $\bar{x}_{dk}$ earns no additional payment. Filled circles mark included endpoints and open circles mark excluded endpoints. Both thresholds were set before randomization and vary by municipality.

Table A.7: Managerial practices: alternative samples and levels of aggregation

	Delivery & Output	Delivery-only mean	Effect (%)	*N*
Panel A. Managers' survey, districts with a surveyed health facility				
Monitored anemia KPI indicator	0.240***	0.726	33.1	171
	(0.058)			
Monitored CHWs daily or biweekly	0.065	0.476	13.7	171
	(0.061)			
Improved care for children at risk of anemia	0.100	0.440	22.7	171
	(0.077)			
CHW was incentivized	0.126*	0.345	36.5	171
	(0.075)			
Panel B. Health facilities' survey, aggregated to the district level				
Referral is coordinated with health facility	0.185***	0.270	68.5	210
	(0.056)			
CHWs take mothers to health facility	0.129**	0.262	49.2	210
	(0.055)			
Number of CHW training sessions	1.739**	3.669	47.4	210
	(0.716)			

Notes: Panel A restricts the managers' survey to municipalities in which at least one health facility completed the endline survey. Panel B aggregates the health facilities' survey to the municipality level. Standard errors clustered at the municipality level in parentheses. *Effect (%)* reports the point estimate as a percentage of the *Delivery-only* mean. Only the anemia-monitoring specification controls for the baseline value of the dependent variable; the remaining specifications do not. No other controls are included. Statistical significance: $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table A.8: Effects on Staff

	Delivery & Output	Delivery-only mean	Effect (%)	*N*
Panel A. Community health workers (CHWs)				
Total CHW	-0.279	11.674	-2.4	442
	(1.003)			
Share with higher education	0.013	0.166	7.8	442
	(0.022)			
Panel B. Supervisors				
Total supervisors	0.388	0.652	59.5	442
	(0.236)			
Share with higher education	0.040	0.289	13.8	442
	(0.044)			

Notes: This table reports the effects of the *Delivery and Output* contract on staffing and personnel management. Each row is a separate regression. Panel A reports effects on community health workers (CHWs) and Panel B on supervisory staff; both use the managers' survey and are estimated at the municipality level ($N = 442$). *Total CHWs* is the number of CHWs in the municipality. *Share with higher education* is the proportion of CHWs (Panel A) or supervisors (Panel B) with a higher-education degree. *Effect (%)* reports the point estimate as a percentage of the *Delivery-only* mean. All estimates control for the outcome baseline. No controls are included. Standard errors clustered at the municipality level are in parentheses. Statistical significance: $* p < 0.1$, $** p < 0.05$, $*** p < 0.01$ (unadjusted); $\star\ p < 0.1$, $\star\star\ p < 0.05$, $\star\star\star\ p < 0.01$ (MHT-adjusted). All outcomes belong to a single MHT family.

Table A.9: Municipal Expenditure (in log)

	Delivery & Output	Delivery-only mean	Effect (%)	N
Panel A. Aggregate expenditure				
Total municipal expenditure	0.047	3,205.100	0.001	539
	(0.047)			
Total health expenditure	0.014	109.916	0.013	539
	(0.086)			
Panel B. Health expenditure by category				
HV Program: payments to CHWs	0.004	3.673	0.109	539
	(0.046)			
HV Program: other spending	0.042	7.799	0.539	539
	(0.054)			
Infant and maternal health outside the Program	−0.054	28.916	−0.187	539
	(0.104)			
Other health conditions	5.457	69.528	7.849	539
	(14.823)			

Notes: This table reports treatment effects on municipal expenditure (in logs) by spending category during the trial phase (August to December 2024). Each row is a separate regression, estimated at the municipality level ($N = 539$). Panel A reports total municipal expenditure and total health expenditure. Panel B disaggregates total health expenditure into four categories. The first two cover the Home Visit (HV) Program, separating payments to CHWs from all other Program spending. The remaining two cover spending outside the HV Program, separating infant and maternal health from other health conditions (i.e., tuberculosis, HIV/AIDS, vector-borne and zoonotic diseases, non-communicable diseases, and cancer prevention), which also includes administrative expenses. *Effect (%)* reports the point estimate as a percentage of the *Delivery-only* mean. *Delivery-only* means reflect annual expenditure in 2023 and are reported in thousands of USD. Standard errors clustered at the municipality level are in parentheses. Statistical significance: $*$ $p < 0.1$, $**$ $p < 0.05$, $***$ $p < 0.01$ (unadjusted); $\star$ $p < 0.1$, $\star\star$ $p < 0.05$, $\star\star\star$ $p < 0.01$ (MHT-adjusted). All outcomes belong to a single MHT family.

Table A.10: Effects on Health Outcomes by Poverty Index

	Hemoglobin (g/dL) (1)	Anemic (2)	Weight-for-Age Z-score (WAZ) (3)
Delivery and Output	0.026	-0.006	-0.096
	(0.125)	(0.032)	(0.119)
Delivery and Output × poverty	-0.104	0.016	0.255*
	(0.143)	(0.041)	(0.134)
poverty	0.106	-0.029	-0.350***
	(0.125)	(0.031)	(0.096)
P-value	0.294	0.716	0.011
Mean control (low poverty ratio)	12.326	0.201	0.349
Mean control (high poverty ratio)	12.817	0.193	0.349
N. Clusters	528	528	528
Observations	11,573	11,573	12,505

Notes: This table reports the effects of the *Delivery and Output* contract on child health outcomes. Column (1) reports the hemoglobin concentration measured during routine anemia screening. Column (2) reports an indicator equal to one if the child is classified as anemic based on altitude-adjusted hemoglobin thresholds, and zero otherwise. Column (3) reports the child's weight-for-age z-score (WAZ), constructed using WHO growth standards. All estimates do not control for the outcome baseline. *Poverty* is an indicator equal to one if the district's poverty index —measured using the 2018 district-level poverty map— is above the median, and zero otherwise. Cohorts fixed effects are included. Additionally, Columns (1) and (2) include linear and quadratic controls for altitude. Standard errors clustered at the municipality level are reported in parentheses. Statistical significance is indicated by $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$ for p-values.

Table A.11: Effects on Health Outcomes by Water and Sanitation Quality

	Hemoglobin (g/dL) (1)	Anemic (2)	Weight-for-Age Z-score (WAZ) (3)
Delivery and Output	-0.086	-0.000	-0.007
	(0.083)	(0.031)	(0.096)
Delivery and Output × water	0.127	0.012	0.135
	(0.149)	(0.039)	(0.126)
water	-0.094	-0.025	-0.031
	(0.127)	(0.029)	(0.092)
P-value	0.736	0.639	0.118
Mean control (low water ratio)	12.511	0.204	0.312
Mean control (high water ratio)	12.707	0.187	0.399
N. Clusters	528	528	528
Observations	11,573	11,573	12,505

Notes: This table reports the effects of the *Delivery and Output* contract on child health outcomes. Column (1) reports the hemoglobin concentration measured during routine anemia screening. Column (2) reports an indicator equal to one if the child is classified as anemic based on altitude-adjusted hemoglobin thresholds, and zero otherwise. Column (3) reports the child's weight-for-age z-score (WAZ), constructed using WHO growth standards. All estimates do not control for the outcome baseline. *Water* is an indicator equal to one if the district's share of households with piped water and sewage connected to the public network —measured using the 2017 Population and Housing Census— is above the median (p50), and zero otherwise. Cohorts fixed effects are included. Additionally, Columns (1) and (2) include linear and quadratic controls for altitude. Standard errors clustered at the municipality level are reported in parentheses. Statistical significance is indicated by $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$ for p-values.

Table A.12: Effects on Health Outcomes by Malaria index

	Hemoglobin (g/dL) (1)	Anemic (2)	Weight-for-Age Z-score (WAZ) (3)
Delivery and Output	-0.031	-0.001	0.064
	(0.070)	(0.020)	(0.064)
Delivery and Output × malaria	0.037	0.051	-0.097
	(0.299)	(0.082)	(0.292)
malaria	-0.281	0.040	-0.268*
	(0.239)	(0.039)	(0.140)
P-value	0.983	0.526	0.907
Mean control (low malaria ratio)	12.696	0.194	0.346
Mean control (high malaria ratio)	10.972	0.246	0.397
N. Clusters	528	528	528
Observations	11,573	11,573	12,505

Notes: This table reports the effects of the *Delivery and Output* contract on child health outcomes. Column (1) reports the hemoglobin concentration measured during routine anemia screening. Column (2) reports an indicator equal to one if the child is classified as anemic based on altitude-adjusted hemoglobin thresholds, and zero otherwise. Column (3) reports the child's weight-for-age z-score (WAZ), constructed using WHO growth standards. All estimates do not control for the outcome baseline. *Malaria* is an indicator equal to one if there is at least one reported malaria case in the district and zero otherwise, based on data from the National Center for Epidemiology, Prevention and Disease Control. Cohorts fixed effects are included. Additionally, Columns (1) and (2) include linear and quadratic controls for altitude. Standard errors clustered at the municipality level are reported in parentheses. Statistical significance is indicated by $* \ p < 0.1$, $** \ p < 0.05$, $* * * \ p < 0.01$ for p-values.

Table A.13: Effects on Health Outcomes by Anemia index

	Hemoglobin (g/dL) (1)	Anemic (2)	Weight-for-Age Z-score (WAZ) (3)
Delivery and Output	0.053	-0.011	-0.023
	(0.044)	(0.015)	(0.039)
Delivery and Output × anemia	-0.021	-0.003	-0.053
	(0.094)	(0.032)	(0.057)
anemia	-0.572***	0.211***	0.024
	(0.065)	(0.021)	(0.038)
P-value	0.719	0.645	0.266
Mean control (low anemia ratio)	12.490	0.207	0.230
Mean control (high anemia ratio)	11.730	0.416	0.153
N. Clusters	521	521	524
Observations	12,222	12,222	38,292

Notes: This table reports the effects of the *Delivery and Output* contract on child health outcomes. Column (1) reports the hemoglobin concentration measured during routine anemia screening. Column (2) reports an indicator equal to one if the child is classified as anemic based on altitude-adjusted hemoglobin thresholds, and zero otherwise. Column (3) reports the child's weight-for-age z-score (WAZ), constructed using WHO growth standards. All estimates do not control for the outcome baseline. *Anemia* is an indicator equal to one if a child aged 1–12 months is classified as anemic in 2023, and zero otherwise, based on records from the Public Healthcare Information System (HIS). Cohorts fixed effects are included. Additionally, Columns (1) and (2) include linear and quadratic controls for altitude. Standard errors clustered at the municipality level are reported in parentheses. Statistical significance is indicated by $*$ $p < 0.1$, $**$ $p < 0.05$, $***$ $p < 0.01$ for p-values.

Table A.14: Effects on Health Service Utilization and Health Outcomes during the Trial Phase
(Children exposed for at least 3 month during the trial phase)

	Delivery & Output	Delivery-only mean	Effect (%)	*N*
Panel A. Health service utilization				
Number of HF visits	0.417**	7.825	5.3	6,191
	(0.192)			
Number of iron pick-ups	0.289*	0.072	401.4	6,191
	(0.166)			
Number of vaccine doses	0.358	1.389	25.8	6,191
	(0.230)			
Panel B. Health outcomes				
Weight-for-Age Z-score (WAZ)	0.145	0.421	34.4	5,615
	(0.091)			
Took blood test	0.012	0.804	1.5	6,131
	(0.015)			
Anemic	0.007	0.188	3.7	4,798
	(0.023)			
Hemoglobin (g/dL)	-0.097	12.714	-0.8	4,798
	(0.068)			

Notes: This table reports the effects of the *Delivery and Output* contract on health service utilization and child health outcomes for the children who were exposed at least 3 month to the intervention during the trial phase and before the blood testing. Exposure duration is determined by birth cohort. Each row is a separate regression estimated at the child level. Panel A reports the total number of health facility (HF) visits, iron supplement pick-ups, and vaccine doses received during the study period. According to the national immunization schedule, children aged 2–12 months are scheduled to receive up to 17 routine vaccine doses. In Panel B, *Took blood test* indicates whether the child received a blood test for anemia screening; *Hemoglobin (g/dL)* is the hemoglobin concentration measured during routine anemia screening; *Anemic* equals one if the child is classified as anemic based on altitude-adjusted hemoglobin thresholds. Effect (%) reports the point estimate as a percentage of the *Delivery-only* mean. All estimates control for the outcome at baseline, except *Took blood test*. Cohort fixed effects are included throughout, and the *Anemic* regression additionally includes linear and quadratic controls for altitude. Standard errors clustered at the municipality level are in parentheses. Statistical significance: $*$ $p < 0.1$, $**$ $p < 0.05$, $***$ $p < 0.01$ (unadjusted); $\star$ $p < 0.1$, $\star\star$ $p < 0.05$, $\star\star\star$ $p < 0.01$ (MHT-adjusted). When correcting the p-values for multiple hypothesis testing, I group the outcomes in two MHT families, one per panel.

Table A.15: Longer-Run Effects on Staff

	Delivery & Output	Delivery-only mean	Effect (%)	N
Panel A. Community health workers (CHWs)				
Total CHW	-0.040	12.078	-0.3	352
	(1.027)			
Share with higher education	0.010	0.156	6.4	352
	(0.029)			
Panel B. Supervisors				
Total supervisors	0.316*	0.739	42.8	352
	(0.165)			
Share with higher education	0.035	0.293	11.9	352
	(0.050)			

Notes: This table reports the effects of the *Delivery and Output* contract on staffing and personnel management. Each row is a separate regression. Panel A reports effects on community health workers (CHWs) and Panel B on supervisory staff; both use the managers' survey and are estimated at the municipality level ($N = 352$). *Total CHWs* is the number of CHWs in the municipality. *Share with higher education* is the proportion of CHWs (Panel A) or supervisors (Panel B) with a higher-education degree. *Effect (%)* reports the point estimate as a percentage of the *Delivery-only* mean. All estimates control for the outcome baseline. No controls are included. Standard errors clustered at the municipality level are in parentheses. Statistical significance: $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$ (unadjusted); $\star \ p < 0.1$, $\star\star \ p < 0.05$, $\star\star\star \ p < 0.01$ (MHT-adjusted). All outcomes belong to a single MHT family.

Table A.16: Longer-Run Effects on Municipal Expenditure (in log)

	Delivery & Output	Delivery-only mean	Effect (%)	N
Panel A. Aggregate expenditure				
Total municipal expenditure	-0.005	3,205.100	-0.000	539
	(0.054)			
Total health expenditure	−0.074	109.916	−0.067	539
	(0.088)			
Panel B. Health expenditure by category				
HV Program: payments to CHWs	−0.003	3.673	−0.082	539
	(0.043)			
HV Program: other spending	0.057	7.799	0.731	539
	(0.046)			
Infant and maternal health outside the Program	0.040	28.916	0.138	539
	(0.077)			
Other health conditions	−15.325	69.528	-22.041	539
	(14.208)			

Notes: This table reports treatment effects on municipal expenditure (in logs) by spending subcategory during the reversion phase (January to May 2025). Each row is a separate regression, estimated at the municipality level (N = 539). Panel A reports total municipal expenditure and total health expenditure. Panel B disaggregates total health expenditure into four categories. The first two cover the Home Visit (HV) Program, separating payments to CHWs from all other Program spending. The remaining two cover spending outside the HV Program, separating infant and maternal health from other health conditions (i.e., tuberculosis, HIV/AIDS, vector-borne and zoonotic diseases, non-communicable diseases, and cancer prevention), which also includes administrative expenses. *Effect (%)* reports the point estimate as a percentage of the *Delivery-only* mean. *Delivery-only* means reflect annual expenditure in 2023 and are reported in thousands of USD. Standard errors clustered at the municipality level are in parentheses. Statistical significance: $*$ $p < 0.1$, $**$ $p < 0.05$, $***$ $p < 0.01$ (unadjusted); $\star$ $p < 0.1$, $\star\star$ $p < 0.05$, $\star\star\star$ $p < 0.01$ (MHT-adjusted). All outcomes belong to a single MHT family.

Table A.17: Effects on Health Service Utilization and Health Outcomes during the Reversion Phase
(Children exposed for at least 3 month during the trial phase)

	Delivery & Output	Delivery-only mean	Effect (%)	*N*
Panel A. Health service utilization				
Number of HF visits	0.562	7.825	7.2	6,191
	(0.360)			
Number of iron pick-ups	0.358	0.072	497.2	6,191
	(0.268)			
Number of vaccines doses	0.085	1.389	6.1	6,191
	(0.065)			
Panel B. Health outcomes				
Weight-for-Age Z-score (WAZ)	0.031	0.412	7.5	5,825
	(0.049)			
Took blood test	0.005	0.647	0.8	6,131
	(0.015)			
Anemic	-0.006	0.273	-2.2	3,895
	(0.026)			
Hemoglobin (g/dL)	0.033	12.288	0.3	3,895
	(0.068)			

Notes: This table reports the effects of the *Delivery and Output* contract on health service utilization and child health outcomes for the children who were exposed at least 3 month to the intervention before the blood testing. Exposure duration is determined by birth cohort. Each row is a separate regression estimated at the child level. Panel A reports the total number of health facility (HF) visits, iron supplement pick-ups, and vaccine doses received during the study period. According to the national immunization schedule, children aged 2–12 months are scheduled to receive up to 17 routine vaccine doses. In Panel B, *Took blood test* indicates whether the child received a blood test for anemia screening; *Hemoglobin (log)* is the log of hemoglobin concentration measured during routine anemia screening; *Anemic* equals one if the child is classified as anemic based on altitude-adjusted hemoglobin thresholds. *Effect (%)* reports the point estimate as a percentage of the *Delivery-only* mean. Estimates control for the outcome at baseline, except *Took blood test*. Cohort fixed effects are included throughout, and the *Anemic* regression additionally includes linear and quadratic controls for altitude. Standard errors clustered at the municipality level are in parentheses. Statistical significance: $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$ (unadjusted); $\star \ p < 0.1$, $\star\star \ p < 0.05$, $\star\star\star \ p < 0.01$ (MHT-adjusted). When correcting the p-values for multiple hypothesis testing, I group the outcomes in two MHT families, one per panel.

B. Pre-trial evaluation of the Home Visit Program under the *Delivery-only* contract (2023)

Before the trial, I evaluated the Home Visit Program under the *Delivery-only* contract in 2023. This appendix reports that evaluation. Its results informed the *Delivery and Output* contract that I designed with the Ministry of Economy and Finance and the Ministry of Health for the RCT.

BI Design and data

In 2023, the program ran from July to December. The central government paid municipalities for delivering home visits, regardless of child health outcomes. Eligibility followed the municipal classification of MEF (Table B.1). Among municipalities outside main cities, type G municipalities have a rural population share above 65% in the 2017 Census and joined the program. Type F municipalities have a rural share between 30% and 65% and did not. The government applied this rule without exception. All 614 municipalities above the threshold in the sample are type G, and all 485 below it are type F (Figure B.1, Panel A). The rule therefore yields a sharp regression discontinuity (RD) design, with the rural share as the running variable.

Table B.1: Municipality classification

Municipality level	**Type of Municipalities**			**Number of municipalities**	**HV Program**
Provincial	Belong to a main city		A	74	Yes
	Does not belong to a main city		B	122	
District	Belong to Metropolitan Lima		C	42	Yes
	Belong to other main cities		D	129	Yes
	Does not belong to a main city	Rural population < 30%	E	378	
		30% ≤ Rural population ≤ **65%**	F	509	
		Rural population > **65%**	G	637	Yes
			Total	1,891	

Notes: The table reports the municipality classification used by the Municipal Incentive Program in 2024. Municipalities are classified by level (provincial or district), by whether they belong to a main city, and by the rural share of their population in the 2017 National Population and Housing Census. Main cities are Metropolitan Lima and the remaining principal urban centers. The last column indicates whether the Home Visit Program operates in each type. The Program covers types A, C, and D, which are the urban municipalities of main cities, and type G, the most rural municipalities. *Source:* Ministry of Economy and Finance.

The sample includes children who were 3 to 5 months old in at least one month of the program. I keep children who were at least 170 days old by December 31, 2023, so that their 6-month anemia screening falls within the data. The sample covers 44,045 children in 1,099 municipalities. Outcomes come from the Public Healthcare Information System for July to December 2023 (Section 4.1). I estimate

$$Y_{im} = \delta_0 + \delta_1 D_m + f(R_m) + D_m \cdot g(R_m) + \mathbf{X}'_{im}\boldsymbol{\gamma} + \varepsilon_{im}, \tag{5}$$

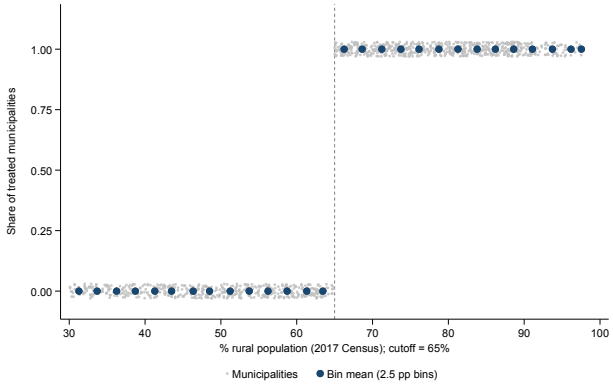
where Y_{im} is the outcome of child i in municipality m, R_m is the rural share minus 0.65, and $D_m = \mathbf{1}\{R_m > 0\}$ indicates eligibility. The linear functions f and g let the slope differ on each side of the threshold. The vector $\mathbf{X}_{im}$ includes sex, age on December 31, 2023, and months in the target age range. The coefficient δ_1 measures the effect of the program under the *Delivery-only* contract relative to no program. I use a triangular kernel and the MSE-optimal bandwidth of Calonico et al. (2020) for each outcome, and I report robust bias-corrected inference (Calonico et al., 2014). I cluster standard errors by municipality, the level at which the running variable varies.

BII Validity of the design

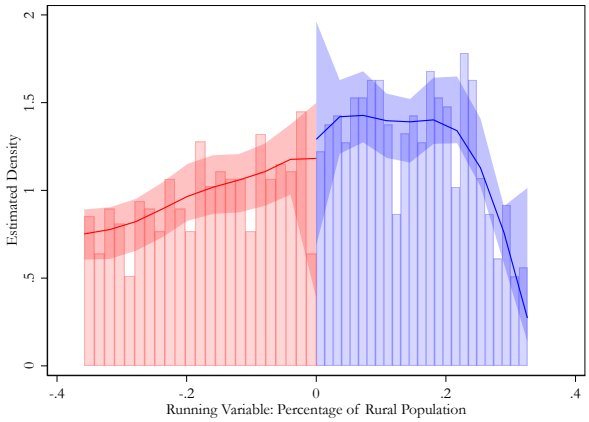
The design requires that municipalities just below the threshold be comparable to those just above it. Table **??** compares municipalities within 10 percentage points of the threshold on population characteristics from the 2017 Census and on municipal capacity. None of the 16 characteristics shows a statistically significant discontinuity. The design also requires that municipalities cannot manipulate their rural share. The National Institute of Statistics (INEI), an agency independent of the program, measured this share in the 2017 Census, two years before the program started. The classification has relied on that census ever since. Consistent with this, the density test of Cattaneo et al. (2020) finds no discontinuity in the number of municipalities at the threshold ($p = 0.381$; Figure B.1, Panel B).

Figure B.1: First stage and density of the running variable

Panel A: First stage



Panel B: Density of municipalities



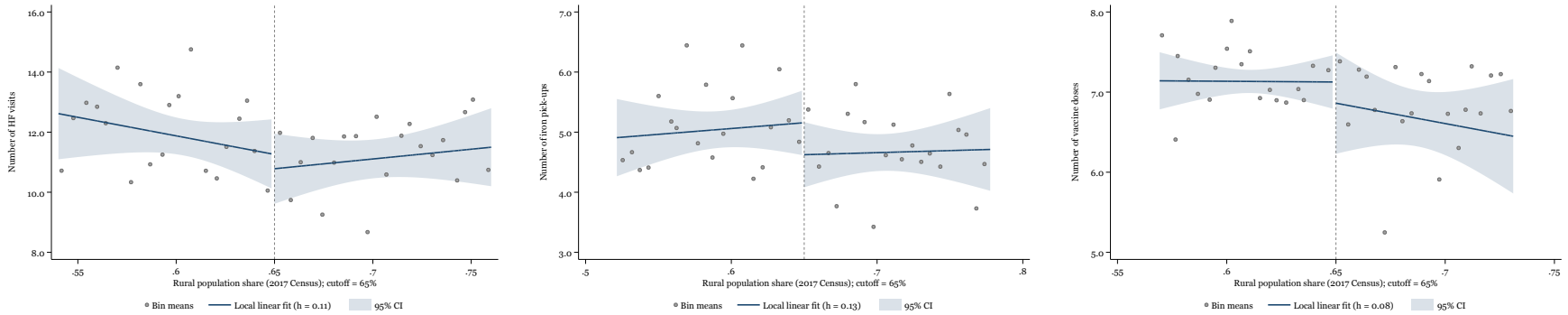
Notes: Panel A plots the share of municipalities that participated in the Home Visit Program against their rural population share in the 2017 Census. Gray dots show individual municipalities, and navy dots show means within 2.5 percentage-point bins. Panel B shows the density test of Cattaneo et al. (2020) for the rural share, centered at the threshold. The dashed line in Panel A and the zero in Panel B mark the 65% eligibility threshold.

BIII Results

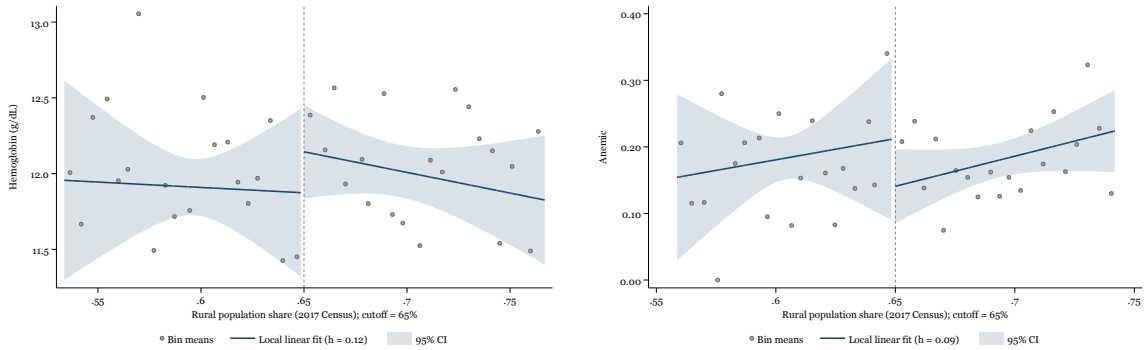
Figure B.2 plots each outcome against the running variable. The program had no detectable effect on health service utilization. The estimates for health facility visits (−0.41, s.e. 0.99), iron pick-ups (−0.60 bottles, s.e. 0.46), and vaccine doses (−0.02, s.e. 0.40) are statistically indistinguishable from zero. They are also small relative to the means below the threshold of 11.8 visits, 5.0 bottles, and 7.1 doses. The program also had no detectable effect on child health at the 6-month screening. Among tested children, the estimates for anemia (−9.0 percentage points, s.e. 8.3) and hemoglobin (0.35 g/dL, s.e. 0.40) are also statistically indistinguishable from zero. The evidence shows that the *Delivery-only* contract failed at the first step of the chain. It paid municipalities for the home visits they delivered, but those visits did not translate into more health facility visits, iron pick-ups, or vaccine doses. Without more iron pick-ups, the program had no channel through which to raise hemoglobin. Because payments depended only on visits, the contract gave municipal managers no incentive to make sure that families followed through. This evidence motivated a change in the incentive scheme.

Figure B.2: Effects of the Home Visit Program under the Delivery-only contract (2023)

Panel A: Health service utilization



Panel B: Health outcomes at the 6-month screening



Notes: Each panel plots an outcome against the rural population share of the municipality in the 2017 Census. The dashed line marks the 65% eligibility threshold, and municipalities to its right participated in the program. Dots show bin means. Lines show local linear fits estimated separately on each side of the threshold, with a triangular kernel and the mean squared error optimal bandwidth of each outcome. Shaded areas show 95% confidence intervals, with standard errors clustered by municipality. The fits do not include covariates. The sample includes children aged 3 to 5 months in at least one month between July and December 2023. Panel A measures outcomes over July to December 2023. Panel B measures anemia and hemoglobin at the first test between 170 and 269 days of age, for tested children only.

D. Appendix — Information Campaign about the Incentive Scheme

The Ministry of Health held a [virtual/in-person] training session for municipalities to explain the operational rules of the incentive scheme for the period from August to December 2024. This training took place between XXX and XXX, 2024. In addition, a virtual information campaign was launched on September 9th, 2024. By October 9th, 2024, all municipalities had accessed the campaign, though municipal managers continued to review the materials until October 20th, 2024.

The information campaign included both personalized and general content. Initially, once the municipality identified itself, it would receive details about the type of contract applicable for the period from August to December 2024 (Figure C.1). Municipalities in the *Delivery-only* arm were informed that their contract would be based on two service delivery indicators: i) the percentage of children who receive home visits, and ii) the percentage of pregnant women who receive home visits (Figure C.1a). The contract for the *Delivery and Output* arm, on the other hand, incorporated three indicators: i) the percentage of children who receive home visits, ii) the percentage of pregnant women who receive home visits, and iii) the percentage of children who are not anemic or whose hemoglobin does not fall between the 6- and 12-month screening tests (Figure C.1b).

Additionally, for each of the indicators, the corresponding partial-bonus threshold ($\underline{x}_i$) and full-bonus threshold ($\bar{x}_i$) were provided. The campaign also explained the percentage contribution of each indicator i to the total payment P. In particular, it was indicated that the fulfillment of each indicator i contributed equally toward the total payment P (i.e., a 50-50% split in the *Delivery-only* contract, and 33.3-33.3-33.3% split for the *Delivery and Output* contract). Municipalities were also provided with a straightforward explanation of the payment method (Equation 1):

- If they did not meet $\underline{x}_i$ for the indicator i, no payment would be made for that indicator
- If they met at least $\underline{x}_i$ but less than $\bar{x}_i$ for the indicator i, they would receive a partial payment for that indicator
- If they reached or exceeded $\bar{x}_i$, they would receive the full potential payment for that indicator i

Finally, all municipalities, regardless of their treatment status, received infographics for the three indicators (Figure C.2).[34] These infographics explained details on the estimation of the indicators (including clear examples), as well as other rules on the implementation of the Home Visit Program.

The information campaign was delivered via Qualtrics, enabling the tracking of key metrics such as views, participation, and time spent on materials. In rural municipalities, an average of one participant per

[34] All the infographics can be found here.

Figure C.1: Example of contracts – Indicators and Thresholds

(a) *Delivery-only* contract

Para el Tramo II, la **MUNICIPALIDAD DE EL CENEPA** será evaluada en base a una serie de indicadores. Cada uno de estos indicadores aporta un porcentaje específico al logro del Monto Máximo correspondiente al Compromiso 1. Además, cada indicador cuenta con sus propias vallas: un Umbral Mínimo y una Meta Máxima.
A continuación, se presentan los indicadores para tu municipalidad junto con sus respectivas vallas y contribuciones:

Indicador	Umbral Mínimo	Meta Máxima	Contribución del Indicador
Indicador de producto 1.2: Porcentaje de niñas y niños hasta los 12 meses de edad que reciben visitas domiciliarias oportunas y consecutivas	85%	92%	50%
Indicador de producto 1.3: Porcentaje de gestantes que reciben visitas domiciliarias oportunas y completas por actor social a partir del II trimestre del embarazo	87%	92%	50%

Fuente: Anexo B1 de la RD Nº 0027-2024-EF/50.01 y Anexo del Decreto Supremo N.° 132-2024-EF.

Nota importante:

Para cada indicador podrían darse 3 situaciones:

Situación	Resultado
No superas el umbral mínimo	No recibes ningún incentivo
Superas el umbral mínimo, pero no alcanzas la meta máxima	Recibes un incentivo parcial
Superas la meta máxima	Obtienes el incentivo completo

(b) *Delivery and Output* contract

Para el Tramo II, la **MUNICIPALIDAD DE LIMA** será evaluada en base a una serie de indicadores. Cada uno de estos indicadores aporta un porcentaje específico al logro del Monto Máximo correspondiente al Compromiso 1. Además, cada indicador cuenta con sus propias vallas: un Umbral Mínimo y una Meta Máxima.
A continuación, se presentan los indicadores para tu municipalidad junto con sus respectivas vallas y contribuciones:

Indicadorx	Umbral Mínimo	Meta Máxima	Contribución del Indicador
Indicador de resultado 1.1: Porcentaje de niñas y niños de 6 y 12 meses sin anemia y/o con incremento de hemoglobina	34%	48%	33.3%
Indicador de producto 1.2: Porcentaje de niñas y niños hasta los 12 meses de edad que reciben visitas domiciliarias oportunas y consecutivas	85%	85%	33.3%
Indicador de producto 1.3: Porcentaje de gestantes que reciben visitas domiciliarias oportunas y completas por actor social a partir del II trimestre del embarazo	87%	92%	33.3%

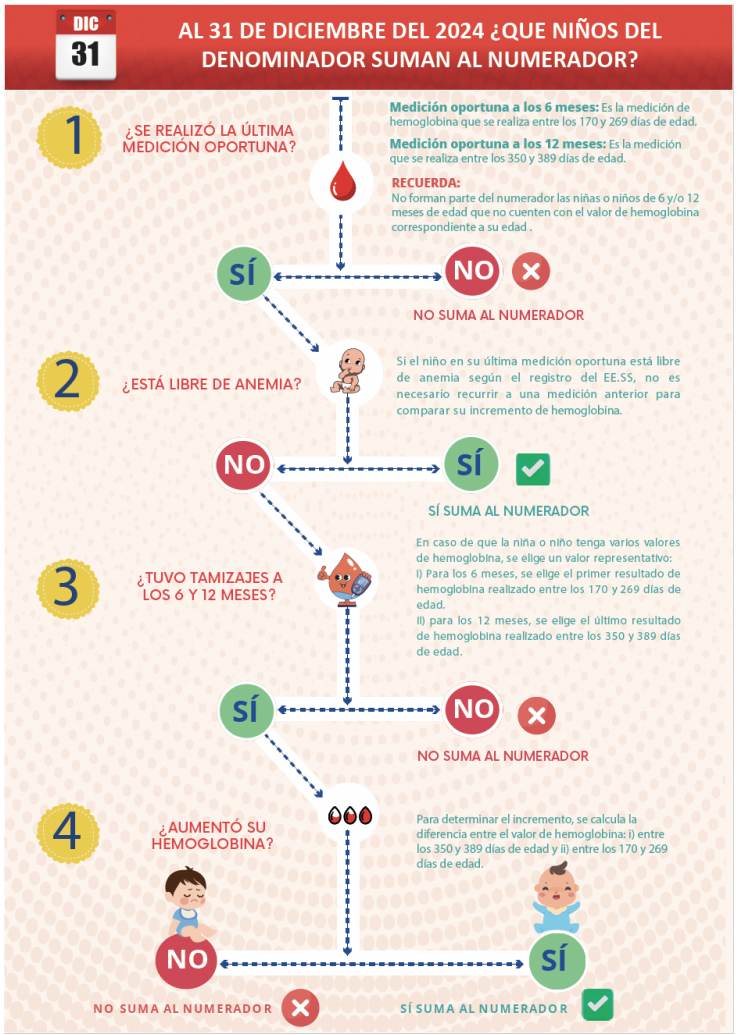
Fuente: Anexo B1 de la RD Nº 0027-2024-EF/50.01 y Anexo del Decreto Supremo N.° 132-2024-EF.

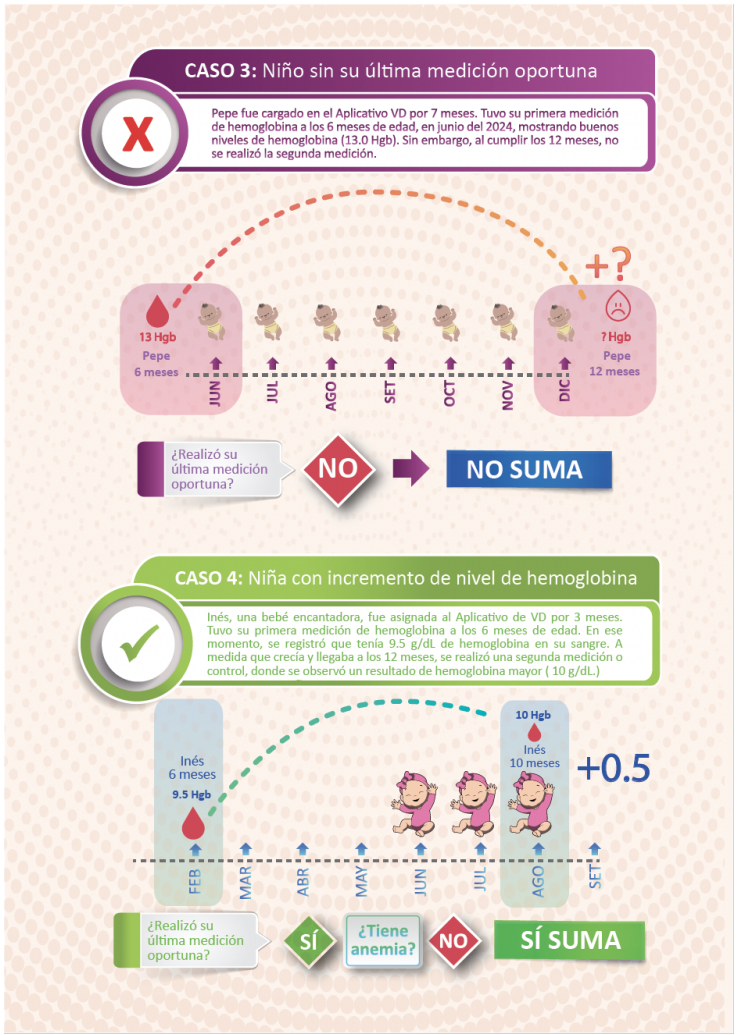
Nota importante:

Para cada indicador podrían darse 3 situaciones:

Situación	Resultado
No superas el umbral mínimo	No recibes ningún incentivo
Superas el umbral mínimo, pero no alcanzas la meta máxima	Recibes un incentivo parcial
Superas la meta máxima	Obtienes el incentivo completo

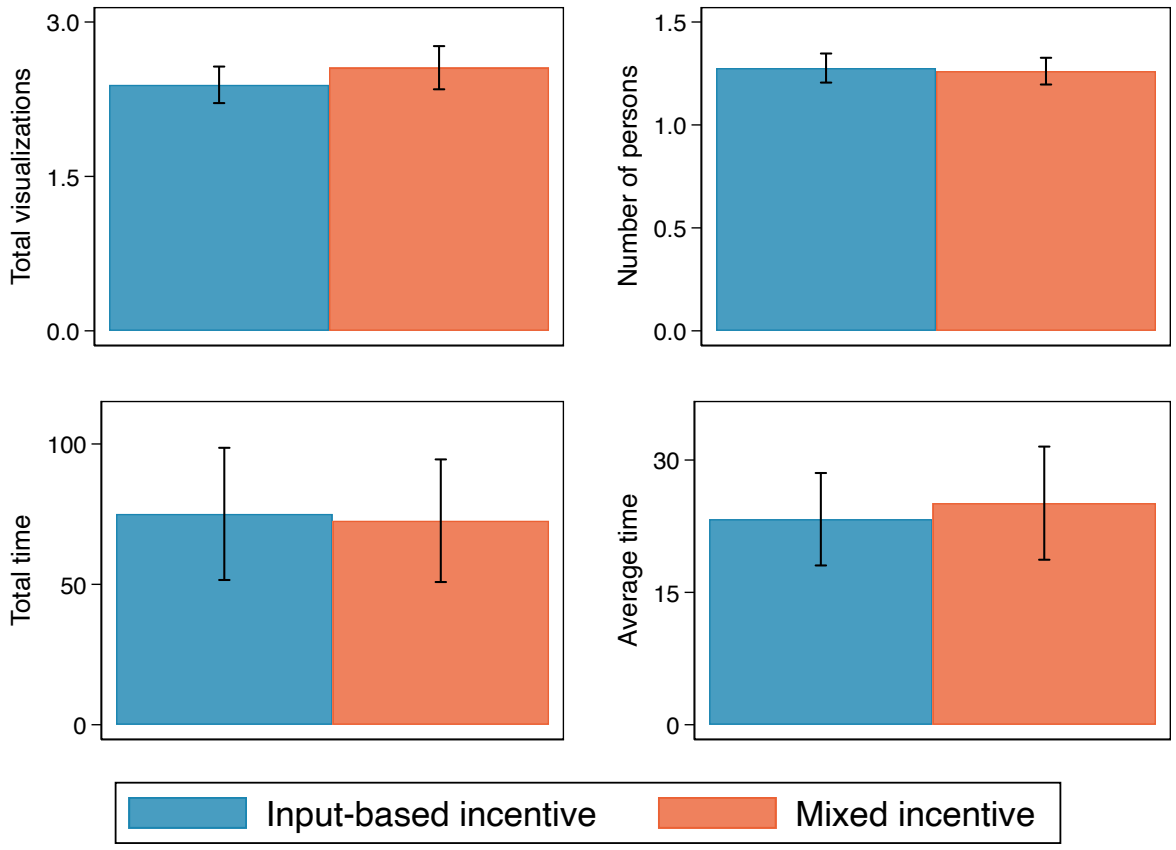
Figure C.2: Examples of infographics





municipality engaged with the content. These municipalities recorded an average of 2.5 views and a total review time of 74.1 minutes, averaging 24.1 minutes per participant (Figure C.3).

Figure C.3: Statistics from Information Campaign



References

Alves, Gabriel et al. (2025). *Nearby solutions: The role of regional and local governments in Latin America and the Caribbean*. https://scioteca.caf.com/handle/123456789/2432.

Ashraf, Nava, Oriana Bandiera, and B. Kelsey Jack (2014). "No Margin, No Mission? A Field Experiment on Incentives for Public Service Delivery". In: *Journal of Public Economics* 120, pp. 1–17.

Ashraf, Nava, Oriana Bandiera, and Scott S. Lee (2016). "Do Financial Incentives Crowd Out Intrinsic Motivation for Public Service Delivery? Evidence from a Field Experiment in Zambia". In: *American Economic Review* 106.5, pp. 357–361.

Ashraf, Nava et al. (2020). "Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services". In: *American Economic Review* 110.5, pp. 1355–1394.

Bandiera, Oriana, Erika Deserranno, and Imran Rasul (2022). "Incentives for Public Service Delivery in Developing Countries". In: *Annual Review of Economics* 14, pp. 325–354.

Bandiera, Oriana et al. (2021). *Social Incentives, Delivery Agents, and the Effectiveness of Development Interventions*. NBER Working Paper 29231. National Bureau of Economic Research.

Basinga, Paulin et al. (2011). "Effect on Maternal and Child Health Services in Rwanda of Payment to Primary Health-Care Providers for Performance: An Impact Evaluation". In: *The Lancet* 377.9775, pp. 1421–1428.

Besley, Timothy and Maitreesh Ghatak (2005). "Competition and Incentives with Motivated Agents". In: *American Economic Review* 95.3, pp. 616–636. DOI: 10.1257/0002828054201413.

Björkman Nyqvist, Martina et al. (2019). "Reducing Child Mortality in the Last Mile: A Randomized Social Entrepreneurship Intervention in Uganda". In: *American Economic Journal: Applied Economics* 11.3, pp. 155–192.

Black, Robert E., Saul S. Morris, and Jennifer Bryce (2003). "Where and Why Are 10 Million Children Dying Every Year?" In: *The Lancet* 361.9376, pp. 2226–2234.

Bloom, Nicholas and John Van Reenen (2007). "Measuring and Explaining Management Practices Across Firms and Countries". In: *Quarterly Journal of Economics* 122.4, pp. 1351–1408.

Calonico, Sebastian, Matias D Cattaneo, and Max H Farrell (2020). "Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs". In: *The Econometrics Journal* 23.2, pp. 192–210.

Calonico, Sebastian, Matias D. Cattaneo, and Rocío Titiunik (2014). "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs". In: *Econometrica* 82.6, pp. 2295–2326.

Cattaneo, Matias D, Michael Jansson, and Xinwei Ma (2020). "Simple local polynomial density estimators". In: *Journal of the American Statistical Association* 115.531, pp. 1449–1455.

Celhay, Pablo A. et al. (2019). "Long-Run Effects of Temporary Incentives on Medical Care Productivity". In: *American Economic Journal: Applied Economics* 11.3, pp. 92–127.

Cerami, Carla (2017). "Iron Nutriture of the Fetus, Neonate, Infant, and Child". In: *Annals of Nutrition and Metabolism* 71.Suppl. 3, pp. 8–14. DOI: 10.1159/000481447.

Chong, Alberto et al. (2016). "Iron Deficiency and Schooling Attainment in Peru". In: *American Economic Journal: Applied Economics* 8.4, pp. 222–255.

Deserranno, Erika (2019). "Financial Incentives as Signals: Experimental Evidence from the Recruitment of Village Promoters in Uganda". In: *American Economic Journal: Applied Economics* 11.1, pp. 277–317. DOI: 10.1257/app.20170670.

Duflo, Esther, Rema Hanna, and Stephen P. Ryan (2012). "Incentives Work: Getting Teachers to Come to School". In: *American Economic Review* 102.4, pp. 1241–78. DOI: 10.1257/aer.102.4.1241.

Dunsch, Felipe A. et al. (2023). "Management, supervision, and healthcare: A field experiment". In: *Journal of Economics & Management Strategy* 32.3, pp. 583–606. DOI: https://doi.org/10.1111/jems.12471.

Fenizia, Alessandra (2022). "Managers and Productivity in the Public Sector". In: *Econometrica* 90.3, pp. 1063–1084. DOI: https://doi.org/10.3982/ECTA19244.

Fracchia, Mattia, Teresa Molina-Millán, and Pedro C. Vicente (2023). "Motivating volunteer health workers in an African capital city". In: *Journal of Development Economics* 163, p. 103096. DOI: https://doi.org/10.1016/j.jdeveco.2023.103096.

Frison, Sarah et al. (2016). "Anthropometric indices and measures to assess change in the nutritional status of a population: a systematic literature review". In: *BMC Nutrition* 2, p. 76. DOI: 10.1186/s40795-016-0104-4.

Fryer, Roland G. (2011). "Teacher Incentives and Student Achievement: Evidence from New York City Public Schools". In: *Journal of Labor Economics* 29.3, pp. 639–681.

Galasso, Emanuela et al. (2019). "Effects of nutritional supplementation and home visiting on growth and development in young children in Madagascar: a cluster-randomised controlled trial". In: *The Lancet Global Health* 7.9, e1257–e1268.

Gera, Tarun et al. (Apr. 2007). "Effect of Iron Supplementation on Haemoglobin Response in Children: Systematic Review of Randomised Controlled Trials". In: *Journal of Pediatric Gastroenterology and Nutrition* 44.4, pp. 468–486. DOI: 10.1097/01.mpg.0000243440.85452.38.

Gertler, Paul and Christel Vermeersch (2012). "Using Performance Incentives to Improve Health Outcomes". In: *Policy Research Working Paper* 6100. URL: http://hdl.handle.net/10986/9316.

Grantham-McGregor, Sally et al. (2020). "Group sessions or home visits for early childhood development in India: a cluster RCT". In: *Pediatrics* 146.6, e2020002725.

Holmström, Bengt and Paul Milgrom (1991). "Multitask Principal–Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design". In: *Journal of Law, Economics, and Organization* 7.Special Issue, pp. 24–52. DOI: 10.1093/jleo/7.special_issue.24.

INEI (2023a). *Perú: Encuesta Demográfica y de Salud Familiar - ENDES 2023*. Tech. rep. Lima, Peru: INEI.

— (2023b). *Perú: Encuesta Demográfica y de Salud Familiar 2023, Nacional y Departamental*. Accedido el 21 de mayo de 2025. URL: https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1950/libro.pdf.

Janmohamed, Amynah et al. (2020). "The effects of community home visit and peer group nutrition intervention delivery platforms on nutrition outcomes in low and middle-income countries: a systematic review and meta-analysis". In: *Nutrients* 12.2, p. 440.

Jones, Gareth et al. (2003). "How Many Child Deaths Can We Prevent This Year?" In: *The Lancet* 362.9377, pp. 65–71.

Kothari, Monica T et al. (2019). "Exploring associations between water, sanitation, and anemia through 47 nationally representative demographic and health surveys". In: *Annals of the New York Academy of Sciences* 1450.1, pp. 249–267.

Larsen, David A et al. (2017). "An individual-level meta-analysis assessing the impact of community-level sanitation access on child stunting, anemia, and diarrhea: Evidence from DHS and MICS surveys". In: *PLoS neglected tropical diseases* 11.6, e0005591.

Lavy, Victor (2002). "Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement". In: *Journal of Political Economy* 110.6, pp. 1286–1317.

Luo, Renfu et al. (2020). "The Impact of Incentive Contracts on School Administrator Behavior and Student Health in China". In: *Journal of Development Economics* 146, p. 102503.

MIDIS (2018). *Plan Multisectorial de Lucha Contra la Anemia*. Tech. rep. Ministerio de Desarrollo e Inclusión Social.

Miller, G. and K.S. Babiarz (2014). "Pay-for-Performance Incentives in Low- and Middle-Income Country Health Programs". In: *Encyclopedia of Health Economics*. Ed. by Anthony J. Culyer. San Diego: Elsevier, pp. 457–466. ISBN: 978-0-12-375679-4. DOI: https://doi.org/10.1016/B978-0-12-375678-7.00126-7.

Miller, Grant et al. (2020). "Can Bureaucrats Really Be Paid Like CEOs? Substitution Between Incentives and Resources Among School Administrators in China". In: *Journal of Development Economics* 145, p. 102462.

Mithra, Prasanna et al. (2020). "Interventions for Addressing Anemia Among Children and Adolescents: An Overview of Systematic Reviews". In: *Frontiers in Public Health* 8, p. 632. DOI: 10.3389/fpubh.2020.00632.

Mohanan, Manoj et al. (2021). "Different Strokes for Different Folks? Experimental Evidence on the Effectiveness of Input and Output Incentive Contracts for Health Care Providers with Varying Skills". In: *American Economic Journal: Applied Economics* 13.4, pp. 34–69. DOI: 10.1257/app.20190220.

Muralidharan, Karthik and Abhijeet Singh (Nov. 2020). *Improving Public Sector Management at Scale? Experimental Evidence on School Governance India*. NBER Working Paper 28129. National Bureau of Economic Research. DOI: 10.3386/w28129. URL: https://www.nber.org/papers/w28129.

Muralidharan, Karthik and Venkatesh Sundararaman (2011). "Teacher Performance Pay: Experimental Evidence from India". In: *Journal of Political Economy* 119.1, pp. 39–77. DOI: 10.1086/659655.

OECD (2022). *2022 Synthesis Report: World Observatory on Subnational Government Finance and Investment*. Paris: OECD Publishing. DOI: 10.1787/b80a8cdb-en. URL: https://doi.org/10.1787/b80a8cdb-en.

Olken, Benjamin A., Junko Onishi, and Susan Wong (2014). "Should Aid Reward Performance? Evidence from a Field Experiment on Health and Education in Indonesia". In: *American Economic Journal: Applied Economics* 6.4, pp. 1–34. DOI: 10.1257/app.6.4.1.

Otero, Cristóbal, Pablo Muñoz, and Gonzalo Ramos (2025). *Managers and Productivity in Public Hospitals*. Tech. rep. Working Paper.

Pasricha, Sant-Rayn et al. (2021). "Iron deficiency". In: *The Lancet* 397.10270, pp. 233–248.

Pimenta, Carlos et al. (2023). *El presupuesto por resultados en América Latina y el Caribe: hallazgos de la encuesta 2023*. Tech. rep. DOI: https://doi.org/10.18235/0005487.

Rasul, Imran and Daniel Rogger (2018). "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service". In: *Economic Journal* 128.608, pp. 413–446.

Rasul, Imran, Daniel Rogger, and Martin J. Williams (2021). "Management, Organizational Performance and Task Clarity: Evidence from the Ghana Civil Service". In: *Journal of Law, Economics, and Organization* 37.2, pp. 433–477.

Stansert Katzen, Linnea et al. (2020). “Home visits by community health workers in rural South Africa have a limited, but important impact on maternal and child health in the first two years of life”. In: *BMC health services research* 20.1, p. 594.

WHO (2001). *Iron Deficiency Anaemia: Assessment, Prevention and Control. A Guide for Programme Managers*. Geneva: World Health Organization.

— (2006). *Guidelines on food fortification with micronutrients*. World Health Organization.

— (2011). *Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity*. Tech. rep. WHO/NMH/NHD/MNM/11.1. Geneva: World Health Organization.

— (2016). *Guideline: Daily Iron Supplementation in Infants and Children*. Geneva: World Health Organization. URL: https://iris.who.int/bitstream/handle/10665/204712/9789241549523_eng.pdf.

— (2023). *Accelerating Anemia Reduction: WHO Global Framework for Action*. Tech. rep. Geneva, Switzerland: World Health Organization.

— (2025). *Anemia*. Publicado el 10 de febrero de 2025. URL: https://www.who.int/news-room/fact-sheets/detail/anaemia.

Wu, Meng-Jia, Kelly Zhao, and Francisca Fils-Aime (2022). “Response rates of online surveys in published research: A meta-analysis”. In: *Computers in Human Behavior Reports* 7, p. 100206. DOI: https://doi.org/10.1016/j.chbr.2022.100206.

Yousafzai, Aisha K et al. (2014). “Effect of integrated responsive stimulation and nutrition interventions in the Lady Health Worker programme in Pakistan on child development, growth, and health outcomes: a cluster-randomised factorial effectiveness trial”. In: *The Lancet* 384.9950, pp. 1282–1293.

Zhou, Jin et al. (2023). “Comparing china REACH and the Jamaica home visiting program”. In: *Pediatrics* 151.Supplement 2, e2023060221I.