EDI WORKING PAPER SERIES

FINANCIAL INCENTIVES IN MULTI-LAYERED ORGANIZATIONS: AN EXPERIMENT IN THE PUBLIC SECTOR

Erika Deserranno

Stefano Caria

Philipp Kastrau

Gianmarco León-Ciliotta

September 2021

Abstract

A classic problem faced by organizations is to decide how to distribute incentives among their different layers. By means of a field experiment with a large public-health organization in Sierra Leone and a structural model, we show that financial incentives maximize output when they are equally shared between a frontline worker and her supervisor. The impact of this intervention on completed health visits is 61% larger than the impact of incentive schemes that target exclusively the worker or the supervisor. Also, the shared incentives uniquely improve overall health-service provision and health outcomes. We use these experimental results to structurally estimate a model of service provision and find that shared incentives are effective because worker and supervisor effort are strong strategic complements in a setting where the supervisor cannot redistribute the incentive costlessly through a side transfer to the worker. Finally, through the use of counterfactual model experiments, we highlight the importance of effort complementarities across the different layers of an organization for optimal policy design.

JEL Codes: O15, O55, I15, J31, M52
Keywords: Financial incentives, multi-layered organizations, effort complementarities, contractual frictions

*Deserranno: Northwestern University, E-mail: erika.deserranno@kellogg.northwestern.edu. Caria: Warwick University, E-mail: stefano.caria@warwick.ac.uk. Kastrau: Universitat Pompeu Fabra & Barcelona School of Economics & IPEG, E-mail: philipp.kastrau@upf.edu. León-Ciliotta: Universitat Pompeu Fabra & Barcelona School of Economics & IPEG, E-mail: gianmarco.leon@upf.edu. This project would have not been possible without the support from the staff at the Sierra Leonean Ministry of Health and Sanitation, the CHW Hub and the International Rescue Committee. We thank Michael Best, Ernesto Dal Bó, Rachel Glennester, Tim Fedderson, Frederico Finan, Jonas Hjort, Anne Karing, Niccolo Meriggi, Alessandro Tarozzi for insightful suggestions. We also thank participants at the BREAD Development Conference, NBER Development Summer Institute, RIDGE Political Economy, SIOE, the Stockholm School of Economics, IIES and Oxford University for helpful comments. Luz Azlor, Andre Cazor Katz, Margaux Jutant and Raquel Lorenzo Vidal provided outstanding research assistance. Financial support for this project was provided by UK aid from the UK government (through the Economic Development and Institutions initiative), the Rockefeller Foundation (through the IGC), and J-PAL's Governance Initiative. León-Ciliotta thanks the Spanish Ministry of Economy and Competitiveness (through the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2011-0075) and grant ECO2014-55555-P. IRB has been approved by the Government of Sierra Leone, Universitat Pompeu Fabra (Parc de Salut MAR: 2018/7834/I) and Northwestern University (ID: STU00207110). Study pre-registration: AEARCTR-0003167. There are no conflicts of interest. Authors are listed in a random order (AEA random author order archive code 69q7bqNU79Fz).

About Economic Development & Institutions

Institutions matter for growth and inclusive development. But despite increasing awareness of the importance of institutions on economic outcomes, there is little evidence on how positive institutional change can be achieved. The Economic Development and Institutions – EDI – research programme aims to fill this knowledge gap by working with some of the finest economic thinkers and social scientists across the globe.

The programme was launched in 2015 and will run until 2022. It is made up of four parallel research activities: path-finding papers, institutional diagnostic, coordinated randomised control trials, and case studies. The programme is funded with UK aid from the UK government. For more information seehttp://edi.opml.co.uk.

1 Introduction

Large organizations are typically divided in multiple, hierarchical layers (Wilson 1989). For example, it is common for a selected group of experienced workers to take on managerial tasks, while frontline employees focus on direct service provision or client interactions. The efforts of workers in these various layers of the organization jointly contribute to the production of final output: without good management, frontline workers are often ineffective, and similarly, the efforts of managers can only pay off if frontline workers are motivated to do their job. This gives rise to a problem of optimal incentive provision which has been surprisingly neglected in empirical research: How should incentives be divided among the different layers of a vertical organization?[1]

We show that financial incentives perform best when they are shared between workers at different layers of the organization. In an experiment with a large public-health organization, we introduce a piece-rate incentive and randomize how the payment is shared between frontline workers and their supervisor. We find that sharing incentives equally between the two agents generates an increase in health visits – the main output of the organization – that is 17% larger than the one achieved when offering the full incentive either to the worker or to the supervisor. We then structurally estimate a model of service provision and show that shared incentives are particularly effective due to (i) the strong complementarity in worker and supervisor effort, and (ii) large contractual frictions, which limit the redistribution of the incentive through side payments. These features, which are likely to occur in many other organizations, have important implication for optimal policy design, which we explore through counterfactual experiments.

The program we study is a large community-based health program designed to improve health-service provision in the aftermath of the 2014 Ebola outbreak in Sierra Leone, with a focus on the local provision of pre- and post-natal care. Community-health services play a crucial role in reducing the burden of common diseases and child mortality (Nyqvist et al. 2019; Deserranno, Nansamba, and Qian 2020). Furthermore, robust community health campaigns play a key role in the early detection and suppression of epidemics (Ord 2020; Christensen et al. 2021). Access to primary health care is still a major issue in rural areas of developing countries and the expansion of community health worker programs is an important part of the global public health strategy working towards universal health care access (Campbell et al. 2013). Finding ways of optimizing the performance of community health workers is hence a first order policy priority.

Our experimental design enables us to provide causal evidence on how the number of

[1] Seminal theoretical contributions studying this problem include Tirole (1986, 1992); Gibbons (1996); Holmström (2017).

household visits provided by the program responds to the introduction of three different incentive schemes. The program is structured around Peripheral Health Units (PHUs), each composed of an average of 8 health workers, who provide health services to their communities, and one supervisor, who trains and advises the health workers in the PHU, and who accompanies them on household visits. The role of the supervisors in our setting is thus not limited to "monitoring" the workers: they are "enablers" who play a crucial role in the health workers' ability to perform their tasks by providing them with the necessary skills and by building trust towards the health worker in the community.[2]

We introduce a new performance-based incentive scheme in a random sub-sample of PHUs, which pays 2,000 Sierra Leone Leones (SLL) per household visit performed by the health worker and reported through an SMS reporting platform. The recipient of the incentive is varied experimentally across the treated PHUs: (i) the incentive is either paid only to the health workers (bottom-tier incentives), (ii) only to the supervisor (top-tier incentives), or (iii) is shared equally between the health worker and the supervisor. This experimental design allows us to assess which split of the incentives achieves highest output (visits). We measure output by interviewing a random subsample of households in their community and asking them the number and the quality of the visits received by the health worker. Due to potential misreporting, we do not rely on the number of visits reported by the health worker.

To guide our empirical analysis, we propose a simple model of service provision that illustrates the trade offs involved in the choice of how to divide incentives across layers. In the model, a supervisor and a worker interact over two time periods. In the first period, the supervisor chooses how much effort to invest in training and advising the worker, and offers her a side payment conditional on the amount of services delivered at the end of the game. In the second period, the worker chooses how much effort to exert to provide the service. A key intuition is that the optimal share of the incentive to be offered to each agent depends on: (i) the complementarity of worker and supervisor effort (which raises the returns to each agent's effort), and (ii) the presence of contractual frictions (which increase the cost of offering side payments). In our empirical setting, both of these features are likely to be present. The return to worker effort plausibly increases in the amount of training provided by the supervisor, and, vice-versa, the return to supervisor training increases in the effort exerted by the worker. Furthermore, a number factors limit agents' ability to offer side payments, and hence reduce the scope

[2] Community members may initially have doubts about the expertise of the health worker — who is typically known by the community as a farmer or a shopkeeper. The supervisor plays a key role in legitimizing the health worker position in the eyes of the community and thus reinforces the demand for the health services.

for Coasian bargaining.[3]

Our model predicts that, when efforts of the worker and the supervisor do not complement each other, it is optimal to offer the entire incentive to the worker. When efforts are instead complements and the supervisor can redistribute the payment costlessly through a side transfer to the worker, there may be multiple optimal schemes, including one that offers the entire incentive to the supervisor. Importantly, when we have both complementarities and frictions, one-sided schemes do not maximize effort and hence the optimal policy is to share the incentive between both layers of the organization.

In the first part of the paper, we present the causal effects of our treatments on the number of visits carried out by the health workers, as reported by the households. Our central empirical finding is that the shared incentives treatment maximizes the number of visits. Workers in the control group without performance-based incentives (status quo) carried out 5.3 visits per household in the six months prior to our endline survey. This number significantly increases to 7.1 visits (a 40% increase over the control condition) when the incentive is only offered either to the worker or to the supervisor and to 8.7 visits (a 63% increase over the control condition) when the incentive is shared between the worker and supervisor. Overall, the shared incentives generate an increase in health visits that is 61% larger than the increase caused by either of the "one-sided" incentives treatments. We rule out concerns related to quantity-quality trade-offs. The observed increase in the quantity of household visits provided in the shared incentives treatment is not compensated by a reduction in visit length, nor by changes in targeting: i.e., workers in the shared incentives treatment are equally likely to target poor and deserving households as in the other treatment arms.

The large positive impact of the shared incentives treatment on household visits translates into better access to pre- and post-natal care and lower disease incidence. Pregnant or expecting women are more likely to report having received at least four pre-natal visits from any provider and having delivered in a health facility (rather than at home). Households also report fewer instances of fever among children below the age of five. In contrast, the interventions that targeted the full incentive either to the worker or to the supervisor do not have systematic significant impacts on either service provision or on disease incidence.

We find that the shared incentives outperform both one-sided incentives not only in terms of final output but also in terms of cost-effectiveness: they lead to more visits at

[3]These include the limited observability and predictability of worker effort (Duflo, Hanna, and Ryan 2012), the difficulty of making binding commitments (Casaburi and Macchiavello 2019), social norms on the inappropriateness of side payments or institutional rules that limit managerial autonomy (Banerjee et al. 2020; Bandiera et al. 2021), and flypaper effects whereby payments are expected to stay in the layer of the organization to which they are originally allocated (Hines and Thaler 1995).

the same or lower cost. Among the two one-sided incentives, the supervisor incentives are the most cost-effective: they cost less while achieving the same output. These results are driven by the fact that the total incentive payout is a function of the number of visits *reported* by the health worker. While visits are under-reported in all treatments, plausibly due to the reporting costs which we discuss in Section 2, under-reporting reduces with the share of the incentive offered to the worker. Over-reporting is instead minimal across all workers. This is because we back-checked a share of the SMS reports and warned workers that their names would be reported to the local health authority if over-reporting was detected.[4]

In the second part of the paper, we study the mechanisms explaining the large boost in output generated by shared incentives. In line with our theoretical model, we show that both effort complementarity and contractual frictions play an important role. Three key results point to the presence of large effort complementarities. First, the increase in supervisor effort, as measured by the frequency of worker training and advising, is of almost the exact same magnitude in both the supervisor and the shared incentives treatments. This could seem surprising, since the *direct* incentive offered to the supervisor is lower in the shared incentives treatment. However, as predicted by our model, shared incentives compensate for this by providing a strong boost to worker effort, which raises the return to training and hence *indirectly* incentivizes the supervisor to exert more effort. Second, shared incentives generate a larger increase in visits and supervisor effort when effort complementarity is plausibly higher due to the low level of experience of the worker (an inexperienced worker relies more strongly on supervisor training). Third, we carry out a formal mediation analysis where we compute the impact of the interventions holding supervisor effort fixed at different levels. This analysis enables us to identify the portion of the treatment effect that comes from the worker effort alone. Consistent with our hypothesis, we find that the treatment effect due to worker effort increases with the level of supervisor effort.

Two key results indicate that contractual frictions play a role in our setting. First, in line with our model, we find evidence of net positive transfers from the supervisor to the worker. However, the total amount of the transfers is only a small proportion of the total incentive paid and, crucially, it fails to equalize output across treatments. This points to the likely presence of contractual frictions: in absence of these frictions, positive side payments would lead the supervisor and shared incentive treatments to generate the same number of visits. Second, consistent with the presence of contractual frictions, we also find that side payments are virtually zero for supervisors whose are unable to rank the performance of their health workers' performance correctly. These are likely to be supervisors who can only imperfectly observe worker effort and output,

[4] Refer to Section 2 for details.

and thus find it hard to enforce side contracts with the workers.

We present several pieces of evidence which are inconsistent with two alternative explanations of our results. First, we consider non-linearities in the utility function. Shared incentives could be highly effective in the absence of effort complementarities if, for both agents, the marginal utility generated by the incentive declines rapidly after 1,000 SLL (the size of the incentives paid in the shared incentives scheme) or the marginal cost of effort increases steeply after the level of effort generated by a 1,000 SLL incentive. Such sharp non-linearities, however, seem implausible for a sample of low-income individuals with plenty of opportunities to increase effort. Indeed, when we analyze non-parametrically the relationship between treatment effects and proxies of utility (wealth) and costs (distance between the worker and her patients, or between the supervisor and the worker), we do not observe any sharp discontinuities.

Second, we consider aversion to pay inequality. One-sided incentive treatments could be ineffective due to a negative morale effect or if they are perceived as unfair (Breza, Kaur, and Shamdasani 2018; Deserranno, Kastrau, and León-Ciliotta 2021). Our empirical design minimizes this mechanism as health workers are not made aware of the presence of supervisor incentives (if any) and only few seem to have learned it from the supervisors. Moreover, we do not find differential treatment effects by a measure of workers' inequality aversion and we document that, contrary to the prediction of the inequality aversion model, supervisor effort increases (albeit insignificantly) when the incentive is only offered to the worker.

In the third part of the paper, we leverage the experimental variation to structurally estimate our model of service provision and perform different counterfactual simulations. For the estimation, we use moments capturing household visits and supervisor effort in the three treatment conditions and in the control group. The model is able to match these moments with great precision. Most importantly, it reproduces the key finding that the shared incentives treatment maximizes household visits. Further, parameter estimates are highly robust to the introduction of moments that capture side payments.

The estimated model parameters confirm that our results are driven by a strong complementarity of effort. In particular, we estimate that the marginal return to worker effort is up to 36% higher due to the complementarity with supervisor effort. Second, we find that, in the absence of the intervention, supervisors have weak incentives to provide effort. Third, we estimate strong contractual frictions, though this result relies on weaker identification. The combination of these factors determines a fairly inefficient level of service provision in the control group.

We derive three lessons on optimal policy based on the structural model. First, given the estimated parameters, we calculate that the optimal policy would offer 54% of the value of the incentive to the worker, and 46% to the supervisor — a split that is very close to that offered in our shared incentives treatment. Second, we study how the optimal policy changes for different levels of effort complementarity. We find that the optimal allocation of the incentive is quite sensitive to the exact value of this parameter, which emphasizes the importance of re-calibrating the policy in new contexts. Third, the strong complementarity determines a large positive external effect of individual effort, which the agents fail to internalize. This, in turn, makes interventions that tie incentives to joint output substantially more effective than interventions that incentivize effort directly, even in settings where effort is perfectly observable. This result has broad implications for optimal pay structure in organizations where workers at different layers complement each other in the production of output.

This paper contributes to four strands of the literature. First, we contribute to the literature on the optimal allocation of incentives in organizations with multiple tiers, which has been mostly theoretical (Tirole 1986; Milgrom and Roberts 1988). Empirically, the evidence is scarce. Indeed, most empirical papers study the effect of raising incentives in one specific layer of the organization, while holding incentives in the other layer fixed. These include papers focusing on the bottom layer — e.g., frontline workers (Glewwe, Ilias, and Kremer 2010; Muralidharan and Sundararaman 2011; Duflo, Hanna, and Ryan 2012; Ashraf, Bandiera, and Jack 2014), sales associates (e.g., Lazear, 2000) — and papers focusing on the top layer — e.g., high-level public sector officials (Rasul and Rogger 2018; Luo et al. 2019), private sector CEOs/managers (Bandiera, Barankay, and Rasul 2007a; Bertrand 2009; Frydman and Jenter 2010) — with Behrman et al. (2015) as an exception.[5] Our paper contributes to this empirical literature by studying the allocation of incentives across *different* layers of an organization, while holding the total payment per unit of output fixed. This is important because many organizations have limited resources, and hence raising incentives in one layer often implies reducing it in the other layer. To fully assess the effectiveness of incentives, one thus needs to consider the organization as a whole rather than focusing on one layer only.

Second, we add to the literature analyzing the role of supervisors and middle managers in hierarchical organizations. This literature — which spans seminal theoretical contributions (e.g., Tirole 1986, 1992) and a number of recent empirical papers (Dodge et al. 2018; Cilliers et al. 2018; Bandiera et al. 2021; Dal Bó et al. 2021) — explores the implications of the information advantage that supervisors have relative to the principal. For example, it studies how to optimally delegate authority and how to avoid

[5] Behrman et al. (2015) evaluate the effectiveness of three alternative performance incentive schemes on mathematics tests scores in Mexican schools: (1) individual incentives for students only, (2) individual incentives for teachers only, and (3) individual and group incentives for students, teachers, and administrators. Program impact estimates reveal the largest average effects for (3). The paper cannot assess whether this is because of complementarities across layers or because of the different incentives structure (e.g., individual vs. group).

harmful collusion between workers and supervisors, but it typically ignores the productive role of supervisors, and therefore any effort complementarities. In our experiment, we explicitly minimize the scope for collusion through frequent back-checks of worker reports. This enables us to shed light on how the top layer of the hierarchy enables the frontline layer to be productive, and on the implications of this complementarity for the design of incentives.

Third, we advance the literature on effort complementarities in organizations. Seminal theoretical work by Alchian and Demsetz (1972); Itoh (1991); Ray, Baland, and Dagnelie (2007) has suggested that effort complementarities are a key rationale for the existence of organizations, and has reflected on the implications of complementarities for incentive design. Empirically, a number of papers have demonstrated that group incentives that reward joint output are effective in increasing output in "horizontal" teams — composed of workers from the *same* layer of the organization — even if at the potential cost of increasing free-riding (Muralidharan and Sundararaman 2011; Babcock et al. 2015; Friebel et al. 2017). A key distinction between these papers and ours is that we focus on "vertical" teams composed of workers in *different* layers of the organization which have different roles and skills, and which potentially require different levels of effort. This is an important distinction because incentives for joint output are often constrained to be symmetric across workers in the same horizontal team due to fairness concerns (Breza, Kaur, and Shamdasani 2018), while they can be asymmetric in vertical teams.

Fourth, we advance the literature on contractual frictions. While most of the literature has focused on contractual frictions *across organizations or firms* (Coase 1937; Gibbons 2005; Lafontaine and Slade 2007; Bubb, Kaur, and Mullainathan 2018), our paper highlights the importance of such frictions *within an organization*. More specifically, we show that contracting difficulties between different layers of the same vertical organization – e.g., due to the poor observability of subordinates' behavior or the difficulty of making binding commitments – limit the scope for Coasian bargaining across layers and impact the optimal allocation of incentives. While Williamson (1973) emphasizes how contractual frictions can lead to failures internal to the organization, Jackson and Wilkie (2005) give a more nuanced discussion of the efficiency implications of contractual frictions. In particular, when these frictions prevent agents from engaging in behaviors that are contrary to the objectives of the organization, they can actually be efficiency enhancing. Our structural results give empirical support on this point, since they show that, absent the intervention, supervisors are poorly motivated to generate visits. Hence, they will use side payments to maximize an objective function that differs substantially from the objective function of the organization.

The paper is structured as follows. Section 2 discusses the context and research design. Section 3 presents a simple model of service delivery with effort complementarity across layers. Section 4 shows how our incentive treatments affect output and reporting. Section 5 explores the mechanisms underlying our main results. Section 6 presents the structure model and performs a number of relevant counterfactual policies. Section 7 concludes.

2 Context and Research Design

2.1 The Community Health Program

Sierra Leone is one of the poorest countries in the world, with the third-highest maternal mortality rate and the fourth-highest child mortality rate in 2017 (World Health Organization 2017). Such elevated mortality rates have been attributed to the slow post-civil war recovery, the 2014 Ebola outbreak, and a critical shortage of health workers, together with limited access to health facilities throughout the country (World Health Organization 2016). In order to strengthen the provision of primary health care, Sierra Leone's Ministry of Health and Sanitation (MoHS) created a national Community Health Program in 2017. The program is organized around Peripheral Health Units (PHUs), small health facilities staffed with doctors (when available), nurses, and midwives. Each PHU has typically a catchment area of seven to ten villages with one community health worker per village and one supervisor, for a total of approximately 15,000 health workers and 1,500 supervisors nationwide. We will refer to health workers as the "bottom layer" of the organization and to supervisors as the "top layer."

Health workers (bottom layer) – The role of the health workers is to provide a basic and polyvalent package of healthcare services in their community. They do so by making home visits to expecting mothers or mothers who recently gave birth, during which they provide natal-related services: (i) health education (e.g., about the benefits of a hospital delivery); (ii) timely pre- and post-natal check-ups, and (iii) accompany women for birth to the health facility. They also conduct visits to households with young children in which they: (i) educate them on how to prevent and recognize symptoms of malaria, diarrhea, and pneumonia, (ii) treat non-severe cases of malaria and diarrhea, (iii) screen for danger signs and refer for further treatment at health facility when necessary.

Health workers are hired locally and typically have no experience in the health sector prior to joining the program. They work part-time and carry out other daily occupations such as farming, petty trading, or small shopkeeping. They are paid a fixed monthly allowance of 150,000 SLL ($17.5) by the MoHS.[6]

Supervisors (top layer) – The role of the supervisors is to train and advise the seven to ten health workers located in their PHU. They do so in three ways. First, they organize monthly one-day "general trainings" at the local health facility which cover key health topics, such as diagnosing, treating and recognizing danger signs for referral to health facilities. All health workers in the PHU are asked to attend these monthly trainings. Second, supervisors organize "one-to-one meetings" with health workers in their respective villages. These visits are more personalized, and allow the supervisors to further advise the workers and address any issues not covered in the general trainings. Third, supervisors provide "in-the-field supervision" by accompanying health workers on household visits. Supervisors do not provide services themselves to the households, but rather provide health workers with concrete feedback on how to improve service delivery. Supervisor's presence during these household visits also helps build trust towards the health worker in the community and reinforces the demand for her services. This is particularly important, since community members may initially have doubts about the expertise of the health worker — who is typically known by the community as a farmer or shopkeeper — and the supervisor can play a key role in legitimizing their position in the eyes of the community.

The role of the supervisors in our setting is thus not limited to "monitoring" the workers. Supervisors are "enablers" who play a crucial role in the health workers' ability to perform their tasks, akin to middle managers in private sector organizations.[7] Note that the supervisors are not in charge of hiring or firing the health workers. Any personnel decision in our context are made by the head of the PHU. Also, note that a substantial share of the support that the supervisor offers to the worker is personalized. Thus, there is limited scope for economies of scale in supervisor effort.

Most supervisors worked as a health worker in the past, are familiar with the health procedures and organizational practices, and have already acquired health knowledge. They work part-time and are paid a fixed monthly allowance of 250,000 SLL ($29.2) by the MoHS.[8]

[6] In our sample, health workers report dedicating 18 hours per week to their health worker job and 22 hours to other jobs from which they earn another 127,000 SLL ($14.85) per month. The hourly rate from the health worker job is thus comparable to their outside option. We use the January 2019 exchange rate: 1 USD = 8,550 SLL (Sierra Leonean Leones).

[7] This is in contrast with Dodge et al. 2018; Callen et al. 2020; Muralidharan et al. 2021; Bandiera et al. 2021; Dal Bó et al. 2021 in which the supervisor's role is a monitoring one.

[8] In our sample, they report working 11 hours per week as a supervsior and dedicate 21 hours to other jobs from which they earn another 156,000 SLL ($28.1) per month.

Complementarities across layers – As in many organizations, the context we analyze is one where there is scope for strategic complementarities between workers and supervisors. The success of the Community Health Program relies on health workers being willing *and* able to provide health services in their community, and thus having a supervisor that supports them with adequate training and advising, and that builds trust towards the health worker in the community. Without this, health workers are unlikely to correctly diagnose and treat diseases, or to provide adequate pre- and post-natal services to pregnant women. This in turn reduces households' trust in the health worker and likely hampers the demand for her services, ultimately affecting health outcomes. Similarly, the impact of the training that the supervisor offers to a given worker depends on the effort that this worker exerts on the job. If they are not motivated and make few attempts to contact households, the additional knowledge and skills transmitted by the supervisor are unlikely to have a large return. In contrast, if they are hard-working and visit the community regularly, the training that they receive will have a high impact. In sum, it is natural to expect that the effort of the one of the two agents raises the returns to the effort of the other agent. We will quantify these strategic complementarities in Section 5.

2.2 Research Design

Our experiment takes place in 372 PHUs across six districts of Sierra Leone (Bo, Kenema, Bombali, Tonkolili, Kambia, Western Area Rural). The 372 PHUs were randomly assigned to one in four groups of equal size, in which we vary the presence of a new incentive scheme and whether the payments are one-sided (paid only to the worker or only to the supervisor) or two-sided (paid to both the worker and the supervisor).

The new scheme pays an incentive of 2,000 SLL ($0.23) for each reported household visit.[9] We describe the reporting system in the next subsection. In the *worker incentives treatment (T_{worker})*, the incentive of 2,000 SLL is paid entirely to the health worker who provides the visit. In the *supervisor incentives treatment (T_{supv})*, the incentive of 2,000 SLL is paid entirely to the supervisor of the health worker who provides the visit. In the *shared incentives treatment (T_{shared})*, the incentive is equally shared between the health worker and the supervisor (1,000 SLL each). In the *control* group (status quo), the incentive is paid neither to the health worker nor to the supervisor.

[9]The incentives are capped to 60,000 SLL per month per health worker. The list of incentivized households visits (which are also listed in their job description) are: (i) prenatal visits to a pregnant woman, (ii) accompanying a pregnant woman to the PHU for child birth, (iii) postnatal visits within 1 month of birth, (iv) child health checkup visits (for children 1-15 months), (v) visits in which a disease is diagnosed and the patient is either treated or referring to the health facility, (vi) follow-up visits of sick patients, (vii) routine household visits (e.g., providing health education on how to prevent diseases).

Three features of the randomization are of note. First, the randomization was performed at the PHU level. This limits spillovers across treatments as staff interactions are minimal across PHUs. Second, the randomization was stratified by district, average distance between the residence of the supervisor and the health workers in the PHU, and by the number of health workers in the PHU. These variables are important predictors of our main outcome variables and will be controlled for in all our specifications. Third, a random sub-sample of the health workers in our study experienced a change in the promotion process six months after the start of the new incentive scheme, which is the focus of Deserranno, Kastrau, and León-Ciliotta (2021). In Appendix B, we describe the change in the promotion system in details and show that: (a) the results hold if we restrict the analysis to the sample of health workers who did not experience any change in the promotion system, (b) the treatment effects are orthogonal to whether the health worker experienced a change in the promotion system or not.

Structure of the incentive scheme – The new performance-based incentives were effective from May 2018 to August 2019, and were paid by a reputable external organization independent from the government. To boost the credibility of the new incentive scheme, workers were paid on a monthly basis through mobile money and without any delay.

The performance incentives were calculated on a monthly basis based on an SMS reporting system which consists of three steps. (i) Each time a household visit is provided, the health worker is asked to send an SMS to a toll-free number indicating the date of the service, the name and phone number of the patient, and a one letter code corresponding to the service type. If the SMS does not include all the required information, the system returns an error message.[10] All health workers of our study (including those in the control group) are asked to report their visits. (ii) The SMS information is automatically uploaded to a server from which the performance incentives are calculated on a monthly basis and are paid without delay.[11] (iii) The SMS information is continuously back-checked by a team of 34 monitors who contact a random 25% of households each week either by phone or in-person (unannounced visits), and ask them to confirm the date and the type of the household visit.

In the month before the incentives started being rolled out (April/May 2018), all health workers and supervisors of our study received a training on the new reporting system in their respective PHU, and were made aware of the back-checking.[12] Impor-

[10]When the patient is a child, the name and phone number reported are those of the primary care giver. When the household does not have a phone, the phone number of a neighbor is reported.

[11]The SMS is uploaded to the server only if it is sent from the health worker's registered phone number. In practice, this implies that supervisors are unable to report services for their workers.

[12]To keep things as comparable as possible across groups, the training and the subsequent back-

tantly, any incentivized health worker and supervisor was warned that she would not be eligible for any further incentive payment and would be reported to the MoHS if over-reporting was detected. We will later show that this threat was credible and effectively shuts down over-reporting in our context.

Our design however does not eliminate under-reporting. Even though the SMS reporting tool is free to use, reporting is inherently costly, as it is usually the case. First, reporting takes time as this requires gathering information on the patients' name, their phone number and the code corresponding to the service type. Second, a non-trivial share of households refuse to share their phone numbers for privacy reasons. Third, mobile phone coverage is often unreliable and unpredictable in rural areas of Sierra Leone, thus forcing health workers to send the SMS hours after a visit takes place and from other locations. This leads to under-reporting if the worker forgets to send the SMS after the visit and/or forgets to record the information needed for the SMS to be valid. Because health workers are not able to perfectly predict ex-ante issues about the phone network or about households not sharing their phone numbers (and since workers likely have other, intrinsic incentives to exert effort), they sometimes carry out household visits that they later do not report.

We will later show that under-reporting is present in our context, especially in the control group and in T_{supv} where health workers were not directly incentivized for reporting but also in T_{worker} and T_{shared} where workers are directly incentivized for reporting. The latter is quite surprising as the incentive scheme is relatively generous in our setting: a health worker who provides (and reports) one visit every other day can earn up to 13% of her monthly fixed allowance in T_{worker} (2,000*10/150,000) and up to 7% of her monthly fixed allowance in T_{shared} (1,000*10/150,000). The presence of under-reporting in T_{worker} and T_{shared} thus indicates that workers may "leave money on the table." This finding has been documented in various contexts including in Sierra Leonean health facilities. Karing (2021) shows for example that local health facility in Sierra Leone under-report vaccination entries, despite the presence of financial incentives. She shows that this is likely explained by the fact that filling out registers is an effortful task for clinic staff and by nurses often not having the time to record immunizations on the spot and forgetting to do so afterwards.

Due to under-reporting, our main measure of health worker performance is not be based on the number of visits reported by SMS by the health worker, but is rather based on the number of household visits reported by the households.

Finally, note that the potential earnings supervisors can generate from T_{supv} are potentially much larger than the earnings of the health workers in T_{worker} as each of them is in charge of an average of 8 workers. The extent to which these payments

checking were held fixed across all health workers, including those that did not receive any incentive.

incentivize their effort will however depend on (i) the extent to which effort of the supervisor and the worker are complement in the production function of household visits and (ii) the extent to which health workers report their visits even if not directly incentivized to do so. We discuss both of these later in the paper.

Transparency of the incentive scheme and side payments – To mirror most workplace environments where supervisors/managers know the pay of the subordinates but subordinates are not informed about their superior's compensation (Cullen and Perez-Truglia 2019, 2021; Deserranno, Kastrau, and León-Ciliotta 2021), we informed all supervisors of our study about the worker incentives but did not inform the workers about the supervisor incentives. In other words, the presence and the level of the incentive paid to the worker in T_{worker} and T_{shared} was *publicly* disclosed to both workers and supervisors in the PHU, while the incentive paid to the supervisor in T_{supv} and T_{shared} was *privately* disclosed to the supervisor. As we will discuss later in the paper, this limits the presence of negative morale concerns of pay inequality in our context because the only way workers could have learned about incentives paid to the supervisors is if the supervisors chose to share that information with them, and few did so.

While supervisors are made aware of the level of the incentives paid to the workers, they are not told about workers' actual earnings from the incentive scheme. In other words, we did not share with supervisors information on the number of SMS reports sent by the worker each month. We did so (i) to preserve the lack of supervisor's perfect observability of worker behavior, which is common in rural settings, and (ii) to avoid introducing differential observability of worker effort across treatments.[13] The fact that supervisor are not aware of worker actual earnings further minimizes the possibility that the supervisor and the worker collude to report visits that have not actually been carried out.[14]

Importantly, we made clear to all supervisors that they could share all or part of their incentive with workers at their discretion. In fact, given the size of the incentives potentially earned by the supervisors, using side payments is a potentially effective way to translate these incentives into output. As formally explained in our theoretical model in Section 3, the size of the side payments will ultimately depend on their cost. Contracting costs can be the consequence of a number of features in our context. First, supervisors have a limited ability to precisely observe the worker's level of effort and

[13]Reporting is more tightly correlated with actual worker effort in T_{worker} and T_{shared} than in the other treatments groups. As such, providing information on the number of SMS reports to supervisors would have caused a larger increase in supervisor observability of worker effort (and thus a larger reduction in contractual frictions) in these two treatments.

[14]This is in stark contrast with a number of recent empirical papers which focus on collusion between supervisors and workers as a key outcome variable (e.g., Cilliers et al. 2018; Bandiera et al. 2021).

reporting behavior, since production is decentralized and each supervisor is responsible for several workers (also, as explained above, we did not inform supervisors of the number of reports filed by each worker). This makes it hard for the supervisor to assess whether workers exert the level of effort that was requested from them in exchange for a side payment. Second, making binding commitments may be difficult because side contracting is inherently informal and the worker would have limited means to punish the supervisor for defaulting on a side payment (e.g., the worker's threat to reduce future effort would not be credible, since the organization may punish the worker for such low effort). Given this difficulty, the supervisor may need to compensate the worker for the perceived risk of default (Bubb, Kaur, and Mullainathan 2018). Finally, there may be social norms or psychological factors that limits redistributions within the boundaries of the same organization (Hines and Thaler 1995). In line with the presence of contractual frictions, in Section 5, we show that the frequency and intensity of side payments is limited in our context: less than a fifth of supervisors transfer money to workers, especially when they have difficulty observing workers' effort. While side payments may also theoretically occur from the bottom to the top layer of the organization — i.e, workers transferring incentives to the supervisor under her request or to motivate her to work more — we observe no such transfers. As in many organizations, the context we analyze is thus one where transfers flow from the top to the bottom of the organization, and not vice-versa.[15]

Key features of the incentive scheme – A key feature of our design is that the unit cost of the incentive is identical across treatments, in the sense that each household visit reported by the health worker always costs 2,000 SLL to the organization, regardless of whether the money is given to the worker, to the supervisor, or is shared between them. As such, we distinguish ourselves from papers which analyze the effect of raising the incentives in one layer of the organization while holding incentives in the other layer constant.[16] This is an important distinction because any organization that is resource constrained must decide how to allocate pay across the different layers, and raising incentives for the lower-tier may entail reducing incentives at the upper-tier (and

[15] This is in contrast with the setting analyzed in Weaver 2021, where applicants regularly pay to obtain a job. One crucial difference is that in Weaver 2021, as in many other papers that study bribes, those who pay the bribe and those who receive it are not part of the same organization. There are also some cases (e.g., Lameke et al. 2020), where bribes are paid within an organization, but these seem to occur less frequently.

[16] These papers can be divided in two groups: those analyzing the effect of raising the incentives for the bottom layer of the organization while holding incentives at the top constant (e.g., Basinga et al. 2011; Muralidharan and Sundararaman 2011; Deserranno 2019) and those analyzing the effect of raising the incentives at the top while holding incentives at the bottom constant (e.g., Bandiera, Barankay, and Rasul 2007b).

vice-versa).[17]

Another key feature of our incentive scheme is that it rewards both layers of the organization based on (reported) output, rather than based on direct measures of their effort. Given the difficulty to observe and precisely measure effort, these types of scheme are common in private, public and non-profit organizations. In the financial sector, for example, a large fraction of the pay of financial analysts is variable and proportional to the amount of money they raise, while the head of the unit is typically paid a bonus proportional to the amount of money raised in the entire unit. In the retail sector, the commissions earned by both managers and frontline salesmen are a function of total revenues. In most micro-finance or agriculture extension programs, frontline workers are rewarded for the number of clients who take up the financial/agriculture product in their village, while their supervisors are rewarded for the total number of clients in the district. Finally, in the higher education sector, professors are often paid (or promoted) based on the number or quality of their publications, while universities in a number of countries are funded based on the publication record of their staff.

2.3 Data and Balance Checks

2.3.1 Data Sources

Our study leverages three main sources of data:

Staff surveys – All 372 supervisors and 2,970 health workers in the 372 PHUs were surveyed at baseline (in April-May 2018) and at endline (fifteen to sixteen months after the implementation of the treatments, in July-September 2019). They were surveyed on their demographic background, their health knowledge, and their job (number of years of experience, number of hours dedicated to the job). We also have access to village-level information (e.g., distance to the health facility, mobile network coverage) collected from a leaflet that is given to each health worker by the PHU.

[17] Theoretically, the set of possible splits an organization can select from is wider than the three splits in our design (100%-0%, 50%-50% or 0%-100%). An organization could for instance decide to give 25% of the incentive to the worker and 75% to the supervisor (or vice-versa). Due to the limited sample size of the experiment, we couldn't test the efficiency of all possible splits. We chose the 50%-50% split because informal discussions we had with supervisors (outside of our experimental areas) and government officials indicated that such split was the most natural in our setting. More precisely, we asked these figures how they would split an incentive of 2,000 SLL between supervisors and workers such that the number of visits provided in the PHU is maximized. 63% of the respondents answered that the supervisor should be assigned half of the incentive (1,000 SLL), 8% answered that they should be assigned 60% of the incentive (1,200 SLL), 21% answered that they should be assigned 3/4 of the incentive (1,500 SLL), and the remaining 8% answered some other split. In line with this, our structural model will confirm that the optimal split is indeed very close to the 50%-50% one.

Household surveys – A random sample of three eligible households ($\sim$7%) per village were surveyed at endline (in July-September 2019). The respondent of the survey is the female household head, who is the most knowledgeable about health topics. Each respondent was asked questions on the number of visits received by the health workers and the quality of these visits, trust in the health worker, disease incidence among young children, access to pre- and post-natal care.

Administrative data – Throughout the duration of our experiment, we have access to two sources of administrative data. First, we observe the number of SMS reports sent by each health worker, along with the incentive payments. Second, the MoHS provided us with information on the number of health services/patients treated by each local health facility at the monthly level (e.g., number of institutional births at the facility, number of children fully immunized at the facility, number of fever/malaria/diarrhea cases treated at the facility). Each PHU is composed of one health facility.

2.3.2 Summary Statistics and Balance Checks

Table 1 reports summary statistics and balance checks for the health workers' and the supervisors' characteristics (Panel A and B, respectively). Panel A shows that 71% of the health workers are male, 70% have completed primary education and 8% have completed secondary school. On average, health workers are 37 years old, are responsible for 106 households each, and report working 18 hours per week. Panel B shows that supervisors are more likely to be men than the health workers (92%) and are more likely to have completed secondary school (25%). They dedicate fewer hours per week to their job (11 hours per week), and are responsible for an average of 8 health workers each.

To perform the balance checks, we regress each baseline staff characteristic on a dummy variable for each of the 3 treatments, controlling for stratification variables and clustering standard errors at the PHU level. Column (11) reports the p-value from a joint F-test of the equality of all treatment groups. Health workers' and supervisors' characteristics are well balanced across treatments except for the age of the health worker. In Table A.1 (Panels A and B), we report the p-value for each pairwise treatment comparison. Out of 90 pairwise comparisons, 13 are statistically significant.

Table A.2 presents summary statistics on village-level characteristics collected before the onset of the experiment (Panel A) and on household-level characteristics collapsed by village (Panel B). Given the absence of a baseline household survey, we asked households in our endline survey a set of retrospective questions (i.e., age, household composition, occupation at the time baseline) and report those in Panel B. We also report summary statistics on variables that are unlikely to have varied over time (e.g., distance from the health worker house or the PHU, education). Panel C presents summary statistics for health services provided by the local health facility (one per PHU) in the month before the start of the experiment.

The average village has 214 households and is 6.6 km away from the closest health facility. 77% of the villages have an accessible road to the health facility. Phone network is available in all villages but is reliable everywhere only in 41% of the villages. In the rest of the villages, network is reliable only in specific locations. Household respondents are less educated than health workers and supervisors, with only 25% having completed primary school. Household members are also less wealthy, as measured by a wealth score from 0 to 8 that counts the number of items owned on a list of household items (e.g., clothes, pair of shoes, cooking pots). On average, a household owns 1 out of the 8 items while workers and supervisors own 2.5 and 3 items respectively. Most (87%) households live within 30 minutes of the health worker's house. These characteristics are mostly balanced across treatments.

3 Model

We propose a simple model of service provision that features both contractual frictions and a positive complementarity between worker and supervisor effort. The model illustrates how the combination of contractual frictions and complementarities makes one-sided incentive schemes sub-optimal.

For simplicity, we consider the case of a single frontline worker (player 1) and a single supervisor (player 2). The worker's task is to visit households and offer them health services. The supervisor's task is to make it easier for the worker to deliver this service, as explained in Section 2 (e.g., by training, advising, and monitoring the worker). The players interact over two periods. In the first period, the supervisor chooses a level of effort e_2, and offers to pay the worker a side payment of $s \in [0, \infty)$ for every visit that the worker completes. In the second period, the worker observes the effort choice of the supervisor and the side payment they offer, and then chooses effort e_1. This sequential structure reflects the hierarchical nature of the relationship as well as the fact that much of the support offered to the worker (e.g., training) is given in advance of the worker's choice of effort.

Offering side payments is costly. We model this by assuming that a side payment of s costs to the supervisor zs, with $z \geqq 1$. z is a reduced form parameter that captures any barrier to the offer of a side payment (e.g., the poor observability of worker effort, social norms, stickiness of payments), or the difficulty of making binding commitments (e.g., the supervisor may need to compensate the worker for the perceived risk of default). These contractual frictions limit the scope for Coasian bargaining.

Household visits y are produced as a result of both worker and supervisor effort. We capture this with the following output function:

$$y = \alpha e_1 + \gamma e_1 e_2 \tag{1}$$

with α weakly positive. Importantly, when $\gamma > 0$, efforts are strategic complements: the higher the effort of one player, the larger the return to the effort of the other player. Also, this functional form captures the intuition that, when $e_1 = 0$, the supervisor cannot generate any visit no matter how much effort she spends training and advising the worker.

Both players maximize a private payoff that is given by the benefit that the player gets from the visits completed by the worker minus the cost of effort. We assume that each player i gets a benefit of b_i for every completed visit. This captures the combination of intrinsic and extrinsic motives that subjects may have to exert effort in the absence of performance-based incentives (e.g., there may be a threat of losing the job or social status that decreases in y).[18] Additionally, the worker gets a monetary payment of pm per visit in the three treatments, where $p \in [0, 1]$ is the share of the output incentive assigned to the worker, i.e., in the *worker incentives treatment*, $p = 1$; in the *shared incentives treatment*, $p = .5$; and in the *supervisor incentives treatment*, $p = 0$. The supervisor, on the other hand, is paid an incentive of $(1 - p)m$ per visit completed by the worker. Further, the worker also receives a transfer from the supervisor of s per visit, and the supervisor pays an amount zs per visit in order to make this transfer. Finally, both agents bear a convex cost of effort: $c(e_i) = c_i e_i^2$. In sum, the payoffs of the worker and of the supervisor are given by:

$$\pi_1 = (b_1 + pm + s) * y(e_1, e_2) - c(e_1) \tag{2}$$

$$\pi_2 = (b_2 + (1 - p)m - zs) * y(e_1, e_2) - c(e_2). \tag{3}$$

To solve the model, we first find the optimal level of worker effort in the second period, conditional on supervisor effort and side payments. We then derive the optimal effort and side payment chosen by the supervisor in the first period, taking into account the worker's optimal response in the second period. To obtain our main analytical results, we simplify the problem and assume that $b_1 = b_2 = 0$, $c_1 = c_2 = c$, $m = 1$ and $\alpha = 1$. This enables us to illustrate the core features of the model, which are determined by the production function, the possibility of side payments, and the sequential interaction, while setting aside additional considerations that emerge when

[18] In our empirical setting, both agents also receive a fixed wage. However, given the linear utility specification, the introduction of this additional term will not affect our conclusions in any way.

costs or benefits are asymmetric. We will relax these assumptions when we take the model to the data in Section 6. In the simplified setting, the optimal side payment is given by:

$$s^* = \begin{cases} \frac{1-p(1+z)}{2z} & p \leqq \frac{1}{1+z} \\ 0 & p > \frac{1}{1+z} \end{cases} \tag{4}$$

This formula shows that the optimal side payment decreases when contractual frictions and the incentive offered to the worker increase. If contractual frictions are large and the worker receives a large share of the incentive ($p > \frac{1}{1+z}$), the supervisor will not make any side payments. We derive optimal efforts for these two cases — positive side payments ($p \leq \frac{1}{1+z}$) and zero side payments ($p > \frac{1}{1+z}$) — and present the complete mathematical analysis of the model in Appendix C. As expected, the efforts of both players increase in the strength of the complementarity. Further, due to the complementarity, agents' effort do not necessarily increase monotonically in the share of the incentive that is offered to them. This is evident, for example from the optimal efforts when $p > \frac{1}{1+z}$:

$$e_1^* = \frac{pc}{2c^2 - \gamma^2 p(1-p)}. \tag{5}$$

$$e_2^* = \frac{\gamma p(1-p)}{2c^2 - \gamma^2 p(1-p)}. \tag{6}$$

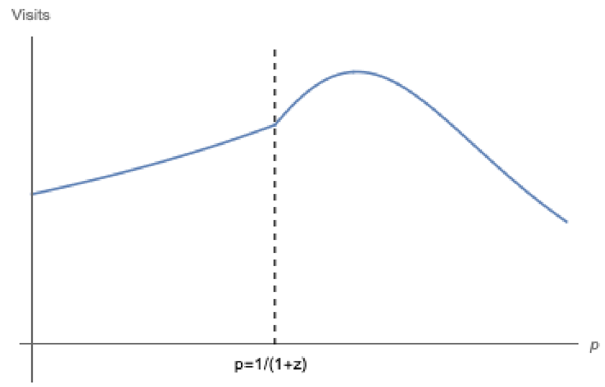
We can use the model to illustrate how the optimal incentive scheme depends on contractual frictions and complementarities in effort. In particular, we consider a policy maker that wants to find the level of p that maximizes visits. In what follows, we will call incentive schemes that only incentivize one player ($p = 1$ or $p = 0$) *"one-sided,"* and schemes that incentivize both players ($0 < p < 1$) *"two-sided."* Also, we will refer to incentive schemes that weakly maximize visits as "optimal". Importantly, we restrict attention to values of γ and c such that $\gamma^2 < 8c^2$. This condition limits the relative size of the complementarity, guaranteeing positive optimal efforts (as we show in Appendix C). We can prove the following result.

Result 1. *When effort complementarity is lower than a threshold level t, there is a unique optimal incentive scheme, which is one-sided: $p^* = 1$. When effort complementarity is larger than t, there is always a two-sided scheme which is optimal: $p^* \epsilon (0, 1)$. If there are contractual frictions, this optimal two-sided scheme is the unique optimal scheme. If there are no contractual frictions, $p = 0$ may also be optimal.*

This result is established in two steps, which are discussed in detail in Appendix

C and summarized here. First, when complementarities are low ($\gamma < t$), supervisor effort has only a limited effect on the worker's ability to carry out household visits. In this case, it is straightforward to show that household visits are maximized by offering the entire incentive to the worker. Second, when complementarities are large ($\gamma \geq t$), supervisor effort becomes central to the optimal incentive decision. If contracting is costly ($z > 1$), incentive schemes that concentrate most of the rewards on one subject are not effective, since the drop in productivity that comes from the low effort of one subject more than offsets the monetary incentive offered to the other subject. Instead, efforts are maximized by intermediate values of p. Thus, the optimal incentive scheme is two-sided, as we show in Figure 1.[19] More precisely, the optimal incentive is either $p^* = \frac{1}{1+z}$ (which is the optimal incentive in the interval $[0, \frac{1}{1+z}]$) or $p^* = \frac{1}{6} + \frac{\sqrt{\gamma^2+24c^2}}{6\gamma}$ (which is the optimal incentive in the interval $(\frac{1}{1+z}, 1]$). On the other hand, if there are no contractual frictions ($z = 1$), the supervisor is able to perfectly match any changes in incentive in the interval $\left[0, \frac{1}{1+z}\right]$ with a commensurate change in side payments. All values of p in that interval result in the same number of visits. If this is the highest possible number of visits, then all $p \epsilon \left[0, \frac{1}{1+z}\right]$ are optimal.

Figure 1: Optimal Incentives ($\gamma > t$ and $z > 1$)



In sum, the model clarifies that, when efforts are strong strategic complements, it is optimal to offer a two-sided incentive scheme that rewards both players. Furthermore, in this case, we may observe that subjects' own efforts do not increase monotonically with the incentive that is offered to them. One final implication of the model, which we explore in Appendix C, is that the difference in output between the optimal two-sided incentive scheme and the one-sided scheme $p = 1$ increases in the complementarity γ. Thus, if in the experiment we find that a two-sided incentive scheme is optimal, we would also expect that the difference in output between this scheme and the worker

[19] In Figure A.1, we show how efforts and side payments change as p changes.

incentive scheme is larger for supervisor-worker pairs that have a high γ. We will explore these predictions empirically in Section 5.

Furthermore, the model sheds light on the important role played by side payments. In particular, two results will help us interpret the empirical data. First, the model shows that, when there are no contractual frictions, all incentive schemes that motivate positive side payments produce the same number of visits. In other words, if we observe positive side payments and equal impacts of the different treatments, it is likely that the supervisor and the worker can contract costlessly. Second, the model shows that there are limits to this balancing role of side payments. In the simple setting that we analyze in this section, side payments drop to zero when contractual frictions are large relative to the incentive that is received by the supervisor. Further, in the Appendix, we present an extension of the model that allows for heterogeneity in benefits and costs. This extended model shows that the supervisor will not offer any side payments when the benefit b_2 that they receive from household visits absent our intervention is lower than the benefit b_1 that is received by the worker.

4 Main Results

In this section, we assess the reduced form effect of our incentives treatments on output (visits). We start by estimating the following regression equation:

$$Y_{ij} = \alpha + \beta_1 T_{worker,j} + \beta_2 T_{supv,j} + \beta_3 T_{shared,j} + Z_j + \varepsilon_{ij} \tag{7}$$

where Y_{ij} represents the number of household visits provided by health worker i in PHU j. $T_{worker,j}$, $T_{supv,j}$, and $T_{shared,j}$ are indicators for whether incentives in PHU j were assigned to health workers only, supervisor only, or were shared between the two. In our model's notation, these correspond to $p = 1$, $p = 0$ and $p = 1/2$, respectively. Z_j are the stratification variables.[20] ε_{ij} is an error term clustered at the level of the treatment assignment, the PHU.

To measure the number of household visits provided by the health worker, we asked each sampled household to report the total number of natal- and disease-related visits received from the health worker in the six months preceding the endline survey, and aggregate these data at the worker level (mean of 7.3). We also study the coverage and range of services provided by the health worker, which we proxy with the share of households who were visited at least once (mean of 71%) and the number of different

[20] As mentioned, these include district fixed effects, whether the average distance between the residence of the supervisor and their health workers is above or below the median, and whether the number of health workers in the PHU is above or below the median.

visit types households received (mean of 1.7).

Our main results are reported in Table 2. Column (1) shows that introducing performance-based incentives significantly boosts the number of household visits provided by the health worker, regardless of whether the incentives are one- or two-sided. The mean number of visits per household in the control group is 5.334. This number increases by 2.090 (40.2%) in the worker incentives treatment, by 2.145 (39.2%) in the supervisor incentives treatment and by 3.356 (62.9%) in the shared incentives treatment. These coefficients are plotted in Figure 2. Interestingly, distributing the incentive to the health workers only is equally effective than distributing it only to the supervisor, while both are outperformed by the two-sided incentive scheme, which achieves 17% more visits. Relative to the control group, the boost in visits is 61% stronger in the two-sided incentive scheme than the two one-sided schemes.[21,22] In line with column (1), we find that health workers in the shared incentives treatment also achieve higher coverage (column 2) and provide a higher variety of services (column 3).

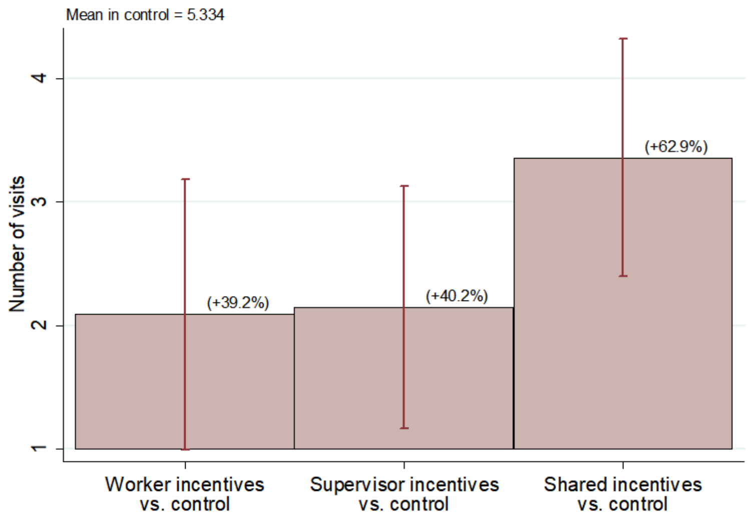
Visit length, trust, and targeting – The higher number of visits provided by workers in the shared incentives treatment may potentially come at the expense of visit length (which is not remunerated), so that the aggregate amount of time dedicated to the job remains unchanged. This would be problematic: workers are expected to follow a checklist when they visit a household and short visits are an indication that such checklist is not properly followed. Alternatively, the higher number of visits in the shared incentives treatment may come at the expense of mis-targeting: health workers may switch from visiting the most deserving households (i.e., poor households) to households who can be visited at a lower cost (i.e., households who are located close by or who are friends or family members). A reduction in visit length or an increase in mis-targeting could reduce the effectiveness of the program.

Table 2 (column 4) shows that, conditional on having received at least one visit, the average visit length reported by a household (23 minutes) did not decrease in the shared

[21] Table A.4 shows that health workers in the shared incentives treatment visit more pregnant women, accompany them more often to the health facility to give birth, and provide more post-natal visits (columns 1-3). They also provide more routine visits, treat more patients under the age of five, and make more referrals to the health facility (columns 4-6).

[22] These results estimate the treatment effects on the average number of visits provided by the health worker to a single sampled household in the six months preceding the endline survey. For completeness, in Table A.3, we also report the corresponding treatment effects on the *total* number of visits provided to sampled households per month (column 1) and on the *total* number of visits provided in the community per month (column 2). The latter outcome variable is measured as the number of visits per month in our sample divided by the share of households included in our sample. We estimate that health workers provide a total of 41 monthly household visits in the community in the control group. This number goes up to 59 in T_{worker} and T_{supv}, and to 67 in T_{shared}. We will use this last set of estimates to assess the cost-effectiveness of our treatments.

Figure 2: Effect of Incentives on the Number of Visits



Notes: The figure plots the difference in the number of visits provided by the health worker between each treatment group and the control group. The coefficients are estimated from a regression of the number of visits on the treatment dummies, controlling for stratification variables and clustering standard errors clustered at the PHU level. Bars are 95% confidence intervals. In brackets, we present the percentage increase in teach treatment group relative to the control group.

incentives treatment.[23] The average number of health topics the household discussed with the health worker also did not decrease; if anything, it increased (column 6). The increase in visit quantity without a reduction in visit quality positively affects trust: the share of households who report trusting the health worker in the shared incentives treatment is 7.1 percentage points (10%) higher in the shared incentives treatment than in the control, and 3.5 percentage points (5%) higher than in both one-sided incentives treatments (column 7). This is an important result because trust in the health worker is known to be one of the main determinants of the demand for health services (Alsan 2015; Lowes and Montero 2021; Martinez-Bravo and Stegmann Forthcoming).

To analyze targeting, we run a household-level regression of the number of visits received by the household on the treatments dummies interacted with whether the household is poor (wealth score below median), lives within 30 minutes of the health worker's home and is a family member or a friend of the health worker.[24] Table A.5 shows no evidence of differential targeting across treatments. That is, health workers who are paid a higher incentive are not more likely to target households who are socially

[23]When we assign an average visit length of zero to the 29% of households who were never visited by the worker, we obtain that the shared incentives increase visit length (see column 5) but this captures both the intensive and the extensive margin of effort.

[24]More precisely, we run the following household-level regression: $Y_{hij} = \alpha + \beta_1 T_{worker,j} + \beta_2 T_{supv,j} + \beta_3 T_{shared,j} + \beta_4 X_h + \beta_5 T_{worker,j} * X_h + \beta_6 T_{supv,j} * X_h + \beta_7 T_{shared,j} * X_h + Z_j + \varepsilon_{hij}$, where Y_{hij} represents the number of visits that the household h received from health worker i in PHU j and X_h is a household characteristic (e.g., poor, social/geographical distance to health worker). All the other variables are defined as in equation (7). ε_{hij} is an error term clustered at the PHU level.

or geographically close to them, nor are they less likely to target the poorest.

Overall, these results alleviate concerns related to quantity-quality trade-offs.

Access to natal-care services and disease incidence – We now test whether the increase in the number of natal- and disease-related services provided by the health worker in the shared incentives treatment translates into better health access and health outcomes.

We start by analyzing households' access to pre- and post-natal care, which we estimate with an equally weighted average of z-scores of six variables: the share of pregnant or expecting women who report having received at least four pre-natal visits from any provider, who have given birth in a health facility, who have received at least one post-natal visit within two days of birth, who have breastfed their infant for at least six months, who have infants who are up-to-date on the vaccination schedule. Table 3 (column 1) shows that the increase in the number of natal-related services provided by the health worker in the shared incentives treatment translates into better access to pre- and post-natal care. Columns 2-6 presents the results for each each single component of the index.[25]

Next, we analyze diseases incidence among children under the age of five, which we proxy with an equally weighted average of z-scores of three variables: the share of households who report that at least one child under five years of age had fever, diarrhea or cough in the past month.[26] Table 3 (column 7) shows that disease incidence index is lower in the shared incentives treatment. This is driven by households in the shared incentives treatment reporting fewer fever instances, while we see no effect for diarrhea and cough (columns 8-10). Note that we find no effect of our treatments on infant or under-five mortality rates, which is not surprising given the timeframe of the experiment.[27]

To corroborate these results, we leverage administrative records from the local health facility (PHU-level data). The results are presented in Table A.6. In line with the household survey data, we find that the number of recorded pregnant women services, institutional births and fully immunized children at the health facility is higher in the

[25]Relative to the worker incentives treatment, the shared incentives treatment increases the share of women who report having received at least four pre-natal visits from any provider by 4 percentage points (5.3%, significant at the 10% level), having given birth in a health facility by 1.6 percentage point (2%, not statistically significant), receiving at least one post-natal visit within two days of birth by 1.4 percentage point (3.2%, not statistically significant), having breastfed her infant for at least six months by 4.2 percentage points (9.4%, significant at the 10% level), and having infants who are up-to-date on vaccination by 3.3 percentage points (15.3%, significant at the 10% level).

[26]The three most serious and common diseases in Sierra Leone are malaria, pneumonia and diarrhea. Because households may not be aware of which disease a child suffered from, we asked them to report whether any child had common symptoms associated with each disease (fever, cough and diarrhea).

[27]Results available upon request.

shared incentives treatment than in the other groups, albeit the results are less precisely estimated. Finally, all three incentives treatments appear to increase the number of diseases treated at the health facility relative to the control group. Given the lower disease incidence rate reported by our sampled households, the positive coefficients are consistent with health workers referring sick children to the health facility more frequently in the treatment groups than in the control.

Reporting – As explained, the incentive scheme in every treatment group pays 2,000 SLL per visit *reported* by the health worker. Such incentive is allocated either to the worker (T_{worker}), to the supervisor (T_{supv}), or to both (T_{shared}). To assess the costs of our three treatment groups relative to the control group without incentives, it is thus crucial to assess the effect of our treatments on the number of visits *reported*, which may differ from the *actual* number of visits due to misreporting.

Table 4 (column 1) presents the results on the average number of SMS reports sent by the health worker each month. Such number is equal to 1.5 in the control group, 3.7 in T_{supv}, 6.3 in T_{shared}, and 8.7 in T_{worker}; and appears quite stable over time with only a slight downward slope in T_{worker} (see Figure A.2). Overall, this indicates that the most expensive incentive scheme for the organization is T_{worker}. We come back to this in the next section.

Table 4 (column 2) and the corresponding Figure 3 present the results on the reporting rate, i.e., the ratio between the number of SMS reports per month (column 1) and the *actual* number of visits per month (Table A.3, column 2).[28] This is an important variable to look at because it gives us a sense of which treatment is most cost-effective. Incentives schemes that reward the person in charge of the reporting (here, the worker) may turn out more expensive than the other schemes if it leads to more over-reporting or less under-reporting.

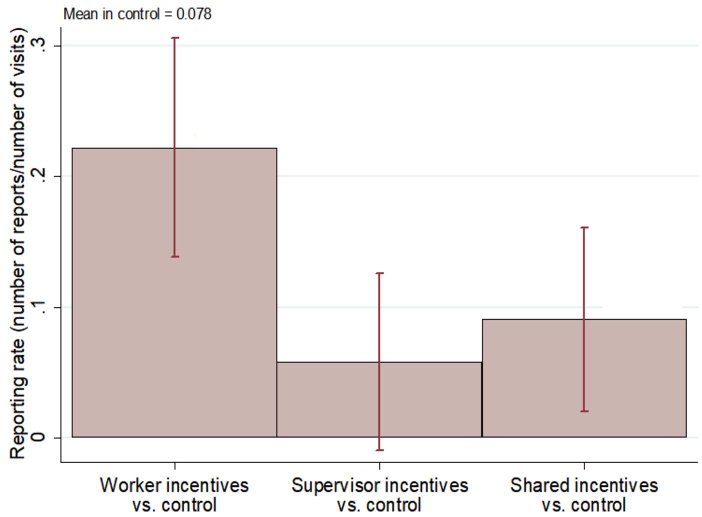
We find that health workers tend to under-report the number of visits provided, especially when they are not incentivized to do so: they report 30.3% of the actual visits in T_{worker}, 17.1% in T_{shared}, 13.8% in T_{supv}, and 8% in the control group. Moreover, the share of workers who under-report is 8 times smaller than the share of workers who over-report (Table 4, columns 3-4).[29] These results are consistent with reporting being costly, even in T_{shared} and $T_{workers}$ where the health workers receive an incentive for reporting. When asked whether the incentive they receive is "too little" or "high

[28]The latter is calculated as the number of actual visits among the random sample of households we interviewed scaled up for the number of households in the community. While the reporting rate we obtain from this calculation may be over- or under-estimated for a single health worker, average differences across treatments are meaningful and accurate.

[29]Again, the exact number of workers who over- vs. under-reporter may not be precisely calculated, but differences in these two numbers across treatments are meaninful.

enough," 88% of the health workers in the control group and in T_{supv} report that it is "too little" (column 5). This number remains high – at 75% and 63% – in T_{shared} and $T_{workers}$.[30]

Figure 3: Effect of Incentives on the Reporting Rate



Notes: The figure plots the difference in the reporting rate between each treatment group and the control group. The coefficients are estimated from a regression of the reporting rate on the treatment dummies, controlling for stratification variables and with standard errors clustered at the PHU level. Bars are 95% confidence intervals.

Cost per visit – The results on the reporting rate are tightly connected to those on the "average cost per visit," which is an important metric to estimate the cost of each treatment.

Among the two one-sided incentives treatments, the cost per visit is twice as high in the worker than in the supervisor incentives treatment: 356 vs. 177 SLL (Table 5, column 5). This is because they cost more due to the higher reporting (column 4); while achieve a similar number of completed visits. We also estimate that supervisor incentives have a similar cost per visit than the shared incentives treatment: 177 SLL in T_{shared} vs. 210 SLL in T_{supv} (column 5), and the difference is not statistically significant. This is because the total cost of supervisor incentives is lower (column 4) but they also achieve fewer visits.

Overall, this indicates that the shared incentives treatment not only outperform the one-sided incentives in terms of output but also in terms of cost-effectivevess: they lead to more visits at the same or lower cost. While worker incentives appear the least cost-effective from the point of view of the organization, note that they maximize health workers' pay relative to the other treatments (columns 1-2), and thus presumably also

[30]Refer to Section 2 for a discussion of what causes reporting to be costly in our context.

lead to higher worker welfare. In the very long run, this may translate in higher worker retention. However, in our medium-run experiment, we do not observe differences in retention across treatments.[31]

Assuming that the Community Health Program were to introduce a shared incentive nationwide, and that the effects at scale remain equivalent to those identified in our experiment, we calculate that the national wage bill of the program would increase by 13% (342 million SLL) and that this would trigger a 63% increase in the number of nationwide household visits (from 615 thousand visits per month without incentives to 1 million visits with shared incentives).[32]

5 Mechanisms

The previous section showed that health workers provide significantly more household visits under the two-sided incentives scheme than under the one-sided incentives schemes. In this section, we explore the mechanisms underlying this result. Guided by the theoretical framework developed in Section 3, we provide evidence consistent with the presence of *both* complementarities in the effort exerted by the supervisor and the health worker and contractual frictions ($z > 1$ and $\gamma > 0$). We then present evidence against two alternative mechanisms that are not considered in our model but are consistent with our reduced-form evidence: non-linear utility and cost functions, and inequality aversion.

5.1 Effort Complementarities

Three pieces of evidence point to the presence of effort complementarities in our setting.

Supervisor effort – First, we estimate the effects of our three incentive schemes on the levels of effort exerted by the supervisor. If effort complementarities were absent ($\gamma = 0$), the effort of the supervisor should monotonically increase with the level of the

[31]Table A.7 shows that the health worker's inactivity rate is lower in our three treatments than in the control but not statistically different across treatments (column 1). Inactivity rate takes value one if the health worker self-reports having dropped out at endline or if all interviewed households report having received no visit from the health worker in the past six months. Inactivity is very low among supervisors and does not vary across treatment groups (column 2).

[32]The Community Health Program employs 15,000 community health workers and 1,500 supervisors nationwide. Without incentives, the total wage bill is estimated at 2,625 million SLL [15,000 (workers) * 150,000 SLL (fixed wage) + 1,500 (supervisors) * 250,000 SLL (fixed wage)]. Introducing shared incentives is estimated to increase the wage bill by 341.7 million SLL [15,000 (workers) * 67 (visits per worker; see Footnote 22) * 17% (reporting rate; see Table 4) * 2,000 SLL (incentive)]. With shared incentives, the number of nationwide visits is estimated to increase by 63%, from 615 thousand visits per month [15,000 (workers) * 41 (visits per worker without incentives; see Footnote 22)] to 1 million visits [15,000 (workers) * 67 (visits per worker with shared incentives; see Footnote 22)].

supervisor's incentives, i.e., be higher in the supervisor incentives treatment relative to the other groups. We show next that this is not the case.

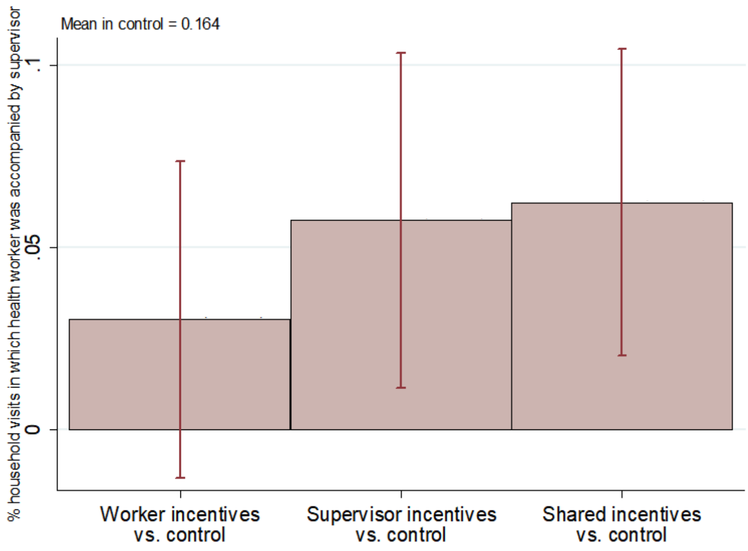
Recall from Section 2.1 that supervisors have three main tasks: (i) they provide in-the-field visits by accompanying health workers on household visits (henceforth, a "joint visit"), (ii) organize one-to-one meetings with each health worker and (iii) organize monthly one-day general trainings. Table 6 column (1) and the corresponding Figure 4 show that the share of household visits in which the health worker was accompanied by the supervisor (as reported by the households) is 5.7 percentage points (35%) and 6.2 percentage points (38%) higher in T_{supv} and T_{shared} respectively, relative to the control group. Importantly, the coefficients for T_{supv} and T_{shared} are nearly identical, and this is despite the fact that the supervisor is paid an incentive which is twice as high in the former than in the latter. This suggests that the overall returns to supervisor effort are similar in the supervisor and shared incentive schemes, which is consistent with the existence of effort complementarities ($\gamma > 0$) that indirectly compensate the supervisor in the shared incentive scheme for the lower monetary payment.

Columns (2) and (3) of Table 6 show that our treatments do not affect other dimensions of supervisor effort: the number of times the supervisor provided one-to-one meetings to the health worker, and the likelihood that the supervisor organized a training in the past month (reported by the health worker) are unchanged. The latter is not surprising as almost all supervisors (99%) organize such trainings.

Finally, column (4) shows that only 16% of the supervisors ever helped the health workers with the SMS reporting. This is also not surprising as all health workers received extensive training on how to report at the start of the experiment (see Section 2). Interestingly, the share of supervisors who helped health workers with reporting is comparable in the two one-sided treatments relative to the control group and is slightly lower in the shared incentives treatment. This indicates that the introduction of supervisor incentives did not divert supervisor's time away from productive tasks (e.g., training workers on health issues) into training them on reporting uniquely.

To shed more light on the role played by the supervisor, we report separate treatment effects on household visits for (i) households that receive at least one visit where the worker is accompanied by the supervisor and (ii) households that do not receive any joint visit. If the supervisor accompanies the worker mostly to monitor their effort, then we would expect to observe a larger treatment effect among households that received at least one joint visit. This is because the supervisor has been in direct contact with these households and hence would find it easier to recontact them at a later date to verify the effort of the worker. Knowing this, the worker should target their visits to these households. On the other hand, if joint visits are mostly a form of training, we should not expect a differential treatment effect since the supervisor would not have

Figure 4: Effect of Incentives on Supervisor Effort



Notes: The figure plots the difference in the fraction of household visits in which health worker was accompanied by supervisor between each treatment group and the control group. The coefficients are estimated from a regression of supervisor's effort on the treatment dummies, controlling for stratification variables with standard errors clustered at the PHU level. Bars are 95% confidence intervals.

a particular reason to target these households for future visits. In Table A.5 (column 5), we show that the boost in household visits is similar for households that receive at least one joint visit compared to households that do not receive any joint visit. This result suggests that the increase in joint visits we document is more likely to capture an increase in training than an increase in monitoring.

Effects by health worker's experience – As a second evidence in favor of effort complementarities, we present heterogeneous treatment effects by the extent to which the effort of the supervisor and the health worker complement each other. We proxy effort complementarity with the lack of health workers' experience. Health workers with little experience are less knowledgeable about health issues and less-known in the community, and they thus plausibly benefit more from the training and advice of the supervisor. We thus expect the shared incentives treatment to be more effective in boosting output and supervisor effort for these less experienced health workers.[33]

In line with this, Figure 5 and Table 7 (columns 1-3) show that the positive effect of the shared incentives treatment on visits and on supervisor effort attenuates as experience increases. For workers with experience below the median, the shared incentives treatment increases the number of household visits provided by the health worker by 4

[33]The median number of years of experience in our sample at baseline is 4 years, with a third of the workers having 1 year of experience and another third having more than 7 years of experience. Workers with above-median experience score 20% higher than workers with below-median experience on a knowledge test and are 16% more likely to be trusted as a valid health provider by the community.

(85%) and increases supervisor effort (measured with the share of households in which the health worker was accompanied by the supervisor) by 9.2 percentage points (70%) relative to the control group. For workers with experience above the median, these effects are significantly lower: they are about half the magnitude for visits and one third of the magnitude for supervisor effort.

Table 7 presents a number of robustness checks for these heterogeneous effects. First, the results are robust to using a continuous measure of experience rather than splitting experience below vs. above the median (columns 3-4). Second, the results are robust to controlling for potential correlates of health worker experience and their interaction with the three treatment dummies (columns 5-8).[34] This is reassuring as it indicates that our heterogeneous effects by worker experience are robust to holding other characteristics fixed. Finally, the results are robust to measuring supervisor effort with an index that bundles all three supervisor tasks (in-the-field visits, one-to-one meeting and general training). Overall, the results confirm that the shared incentives treatment is particularly effective in boosting output and supervisor effort when effort complementarity between the layers of the organization is high.

Mediation analysis – As a final evidence in favor of effort complementarities, we perform a mediation analysis to test whether the effectiveness of workers increases as supervisors exert more effort. Following Acharya, Blackwell, and Sen (2016), we estimate the Controlled Direct Effect (CDE) of the worker incentives treatment on visits *net of a mediator* — here, supervisor's effort. This quantity captures the treatment effect that would be observed if supervisor effort was fixed at an exogenous level, while worker's effort (which is not directly observable in our setting) was allowed to respond to the incentives.[35] We then present this "de-mediated" effect at different points of the supervisor's effort distribution. In the presence of effort complementarities ($\gamma > 0$), we would expect the increase in visits generated by the worker to grow in supervisor effort.

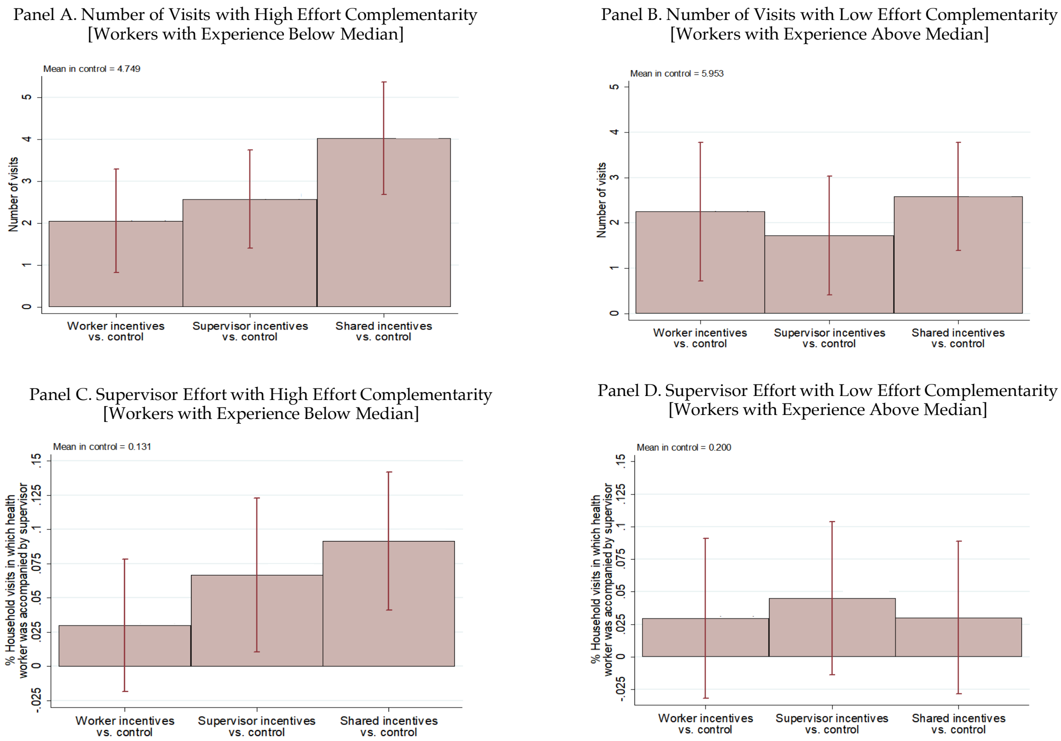
In line with this, Figure 6 shows that the effect of worker effort on output increases substantially with supervisor effort as measured with the fraction of household visits in which the health worker was accompanied by the supervisor. Indeed, the CDE of the worker incentives treatment on visits is close to zero when 0% of the household visits were accompanied by the supervisor and goes up to more than 2 at the opposite extreme when 100% of the household visits were accompanied.[36] This is consistent with

[34]The correlates we control for are all those presented in Table 1: age, gender, completion of primary/secondary education, wealth score, number of households the health worker is responsible for, visits per week, hours worked per week.

[35]We focus on the comparison between the worker incentives treatment and the control group to shut down any role played by the presence of supervisor's incentives in our estimation of effort complementarities.

[36]We produce Figure 6 by following the steps outlined in Acharya, Blackwell, and Sen (2016). First,

Figure 5: Effect of Incentives on Visits and Supervisor Effort by Worker Experience



Notes: The figure plots the difference between each treatment group and the control group in the number of visits (Panels A and B) and supervisor effort (proxied with the fraction of household visits in which the health worker was accompanied by the supervisor; Panels C and D); for workers with experience below the median (Panels A and C) and workers wih experience above the median (Panels B and D). The coefficients are estimated from a regression of the number of visits/supervisor effort on the treatment dummies interacted with experience being above the median, controlling for stratification variables with standard errors clustered at the PHU level. Bars are 95% confidence intervals.

a strategic complementary between worker effort and the in-the-field-training offered by the supervisor. We also find a strong complementarity between worker effort and the general training provided by the supervisor, while we see no complementarity with respect to the one-to-one meetings (see Figure A.3).

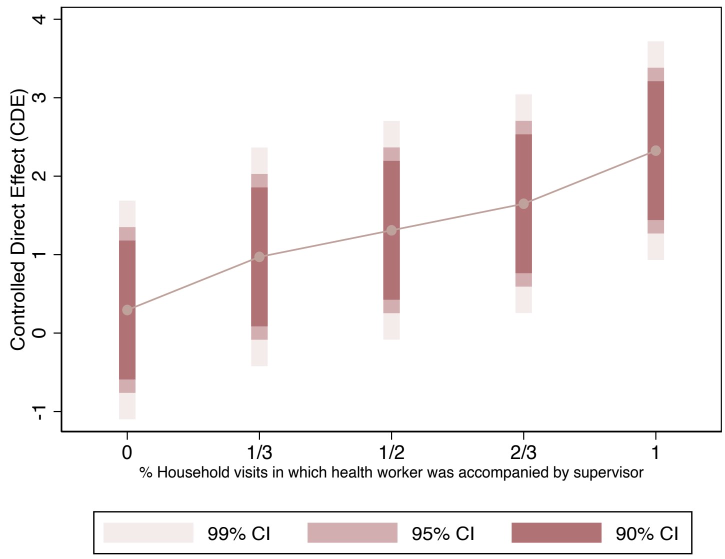
All in all, the empirical evidence is consistent with the presence of effort complementarities. This is in line with our model which predicts that sharing the incentive between the worker and the supervisor should achieve the highest output (i.e., highest number of visits) only if effort complementarities are present ($\gamma > 0$) (see Figure 1).

5.2 Contractual Frictions

Two pieces of evidence support the presence of contractual frictions in our setting.

we regress the number of visits provided by a health worker on the worker incentives treatment, the mediator (supervisor's effort), and their interaction. From this, we obtain a mediation function, defined as the difference between actual and predicted visits based on the coefficients of all covariates (except the treatment) estimated at different levels of the mediator. Finally, we run a regression of the de-mediated function on the treatment and report the coefficients for different levels of the mediator.

Figure 6: Mediation Analysis



Notes: This figure plots the controlled direct effect (CDE) of the worker incentives treatment on the number of visits provided by a health worker for different values of supervisor's effort.

Side-payments – First, the use of side payments is limited. To document this point, we exploit data on side payments collected from the supervisors and the health workers. At endline, all supervisors were asked whether they transferred part or all of their incentive (if any) to health workers since baseline. If they did, we then asked each health worker to assess the value (in-cash or in-kind) of this side payment.[37]

Figure 7 and the corresponding Table 8 (column 1) show that the share of supervisors who make side payments to health workers increases with the level of the supervisor incentive: 19.4% in T_{supv}, 11.3% in T_{shared}, and 1.6% in T_{worker} and 1.1% in the control group. The average amount that a supervisor transfers to a worker *over an entire month* is 702 SLL (resp., 431 SLL) in T_{supv} (resp., T_{shared}): see Table 8 (column 3).[38] These amounts are very small if one considers that the supervisor earns an incentive of 2,000 SLL (resp., 1,000 SLL) per visit reported in T_{supv} (resp., T_{shared}). Workers also occasionally make side payments to their supervisor when they are paid an incentive, but the amount of such transfers is negligible (average of 151 SLL transferred to the supervisor in T_{worker}). Overall, this evidence shows that side payments do happen in our context, but their frequency and magnitude is limited.

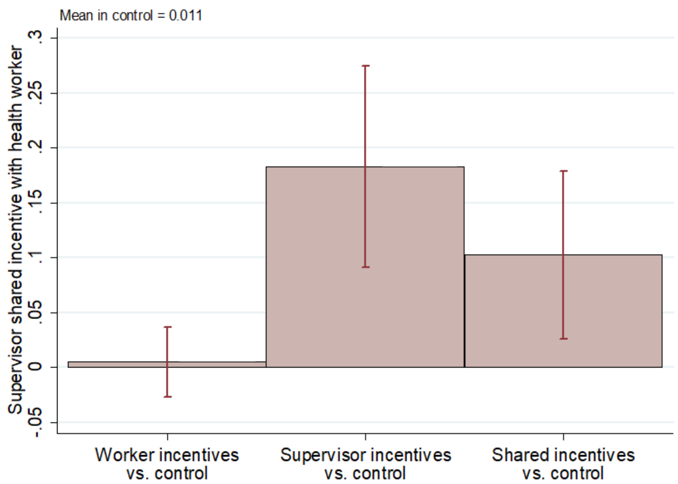
In Section 3, we discussed two different explanations for limited side payments: high contractual frictions or an asymmetry in the benefits that the supervisor and worker

[37]This was asked to health workers rather than supervisors to limit recall bias. To make sure workers and supervisors did not under-report transfers, everyone was made aware from the very start of the experiment that supervisors are entirely free to share incentives with their workers. See Section 2.2 for details.

[38]Supervisors who make no transfer are assigned a value of zero.

derive from household visits in the absence of the intervention. We also showed that, as long as side payments stay positive, we can disentangle these two potential explanations by looking at the impacts on visits of the different treatments. If side payments are low due to contractual frictions, we expect that changes in p (the share of the incentives allocated to the worker) can generate large differences in visits. In contrast, if side payments are low due to asymmetric benefits and there are no contractual frictions, we should observe the same number of visits for all incentive schemes that generate positive side payments. Our results in the previous section, which show that visits are far from being equalized across treatments (see Table 2), point to the likely presence of contractual frictions in our setting.

Figure 7: Effect of Incentives on Side Payments

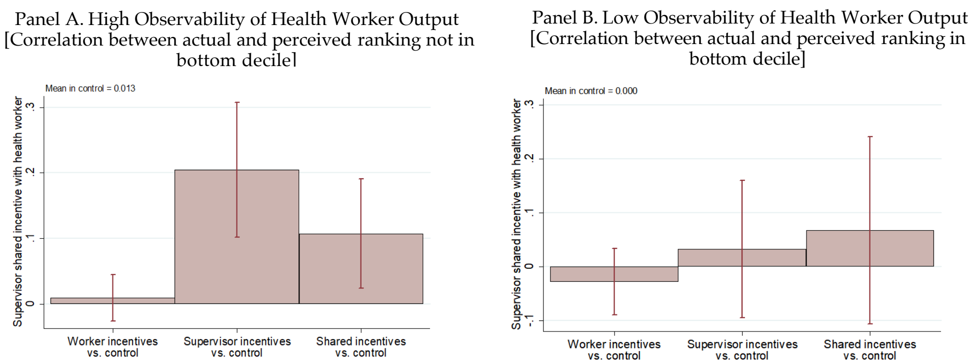


Notes: The figure plots the difference between each treatment group and the control group in the likelihood that a supervisor shared part or all of her incentive with the health worker. The coefficients are estimated from a regression of aupervisor's effort on the treatment dummies, controlling for stratification variables with standard errors clustered at the PHU level. Bars are 95% confidence intervals.

Effect by supervisor's observability of worker output – Second, we test whether supervisors make fewer side payments when contractual frictions are likely to be strong. We proxy contractual frictions with the extent to which the supervisor observes the output of the health workers. At endline, we asked each supervisor to rank the workers she supervises from the best to the worse in terms of their "overall work as a health worker." We correlate this *perceived* rank with the *actual* rank of health workers obtained on the basis of the number of households visits completed at endline. The correlation is positive and significant for over 60% of the supervisors, indicating that most supervisors have some information about worker output. However, a group of supervisors has very limited information: for 10% of the supervisors in our sample, the correlation is negative, indicating that they ranked high-performing workers as being among the worse. Table A.8 (column 1) shows that these poorly-informed supervisors tend to live further away from the health workers and are responsible for more subordinates, while they are equally experienced and work the same number of hours as the supervisors who are better informed. This suggests that our measure of supervisor information captures exogenous differences in the observability of worker output, rather than the supervisor's lack of experience or motivation. And, plausibly, supervisors who find it costlier to acquire information about worker output would also find it costlier to stipulate output-based contracts with their workers.

As expected, Figure 8 (Panels A and B) and Table 9 (column 1) show that side payments in both T_{supv} and T_{shared} are inexistent for the supervisors who observe worker output poorly. In contrast, side payment are positive (even though limited) for the remaining supervisors, who can better observe worker output. These heterogeneous effects are robust to controlling for the number of workers the supervisor is responsible for (which correlates with the observability of worker output; columns 4-6). Overall, these results are consistent with side payments being larger when worker output is more observable and hence contractual frictions are likely to be weaker. This corroborates the hypothesis that contractual frictions are large for several supervisors in our sample.

Figure 8: Effect of Incentives on Side Payments by Contractual Frictions



Notes: The figure plots the difference in the share of supervisors who transfer incentives to health workers between each treatment group and the control grou, for supervisors with high (resp,. low) observability of output in Panel A (resp., Panel B) measured with whether the correlation between actual and supervisor health worker performance ranking is not in the bottom decile (resp,. is in the bottom decile). The coefficients are estimated from a regression of these variables on the treatment dummies interacted with high observability of effort, controlling for stratification variables with standard errors clustered at the PHU level. Bars are 95% confidence intervals.

Interestingly, we do not observe any heterogeneity in side payments with respect to whether the supervisor and the worker know each other well or not. Table A.9 shows indeed that supervisors in T_{supv} and T_{shared} are equally likely to transfer money to workers who are friends/family members relative to those who are not. Similarly, we do not observe any heterogeneity with respect to the number of years the supervisor has known the health worker. This suggests that relational contracts — measured with social distance – have limited ability to overcome the intertemporal contracting problem in our setting and do not attenuate contractual frictions.[39] We also find no heterogeneity in side payments — but also in supervisor effort and output — with respect to the supervisor's span of control (i.e., the number of workers per supervisor). Finally, unlike Bandiera, Barankay, and Rasul (2007a), we find that supervisors who are paid a higher incentive are *not* more likely to target side payments or effort towards workers who they perceived as "high performers" at baseline.

Overall, the empirical evidence of this and the previous section is consistent with Figure 1, where the optimal incentives are not one- but two-sided due to the presence of *both* contractual frictions and effort complementarities ($z > 1$ and $\gamma > 0$).

5.3 Alternative Mechanisms

We examine two alternative mechanisms for why two-sided incentives outperform one-sided incentives: non-linear utility functions and inequality aversion. We discuss these in turn and provide evidence against both of them.

Non-linear utility or cost functions – In the absence of effort complementarities, the two-sided incentives could outperform the one-sided incentives if both the supervisor and the workers have utility functions that are sufficiently concave such that raising the incentive from 0 to 1,000 SLL substantially boosts their effort, but raising it from 1,000 to 2,000 SLL does not translate into any extra effort. This could explain why the supervisors provide the same amount of effort in T_{supv} and T_{shared} (Figure 4). Furthermore, if the production function of household visits is additive in terms of worker and supervisor effort, this could also explain why the shared incentives treatment leads to more visits relative to the other treatments.

To explain the same level of supervisor's effort in T_{supv} and T_{shared}, this mechanism requires the marginal utility to become flat at or before an incentive of 1,000 SLL. While having a very low marginal utility from additional income could be reasonable for wealthy workers or at high levels of income (e.g., corporate managers refusing to work additional hours even for a large pay raise), this is unlikely in the context of doubling a relatively large piece-rate incentive for workers in rural villages in one of the world's poorest countries.[40] To corroborate this, Figure A.4 (Panel A) displays

[39] The evidence on the role of relational contracts in mitigating contractual frictions is mixed. Some papers, like ours, find that relational contracts do not play any role (e.g., between farmers buying or selling an irrigation technology in Indian villages in Bubb, Kaur, and Mullainathan 2018), while others find that they play a role, e.g., between processing mills and farmers in Rwanda's coffee sector in Macchiavello and Morjaria (2021); between sellers and buyers in the Kenyan rose export sector in Macchiavello and Morjaria (2015); or between firms and customer in the Vietnamese informal credit market in McMillan and Woodruff (1999).

[40] Workers in our sample earn an average of 1,443 SLL per hour in their alternative occupation. This is low relative to the amount they can earn for providing one household visit (1,000 or 2,000 SLL for

non-parametric plots of the treatment effects on the number of visits and on supervisor effort, by worker and supervisor wealth score. If the utility function was non-linear and flattening at low levels of income (as this hypothesis implies), we should observe wealthier workers being less responsive to one-sided or two-sided incentives than poorer workers. The figure shows instead that the treatment effects are constant and smooth over the worker and supervisor wealth score distribution. If anything supervisor effort appears to slightly increase with supervisor wealth.

Alternatively, the two-sided incentives could outperform the one-sided incentives if both supervisors and workers have a cost of effort that exhibits a sharp convexity such that any incentive beyond 1,000 SLL does not compensate for the additional cost of effort. This would be the case if, for example, the geographical distribution of households in the village is such that some households are located so close from the health worker's home that the 1,000 SLL is large enough for the health worker to assist them all, and the rest of the households are so far away that even a 2,000 SLL incentive would not incentivize the worker to travel to them. Again, this is a very unlikely scenario as it requires a sharp convexity in the workers' and supervisor's cost of effort around the 1,000 SLL cutoff. Panel B of Figure A.4 presents non-parametric plots of the treatment effects on the number of visits and the supervisor effort over the distribution of household-worker distance (a proxy for the worker's cost of visiting a household) and worker-supervisor distance (a proxy for the supervisor's cost of training/monitoring a health worker). We see no sharp discontinuity in any of the treatment effects by cost of effort.

Inequality aversion – Another mechanism that could explain our core findings is that both the supervisor and the worker display aversion to pay inequality. For example, in the supervisor incentives treatment, the health workers may perceive as unfair the fact that the supervisor earns money for services they provide while they do not earn anything. Similarly, the supervisor may perceive unfair the fact of not being paid any incentive in the worker incentives treatment (while the worker is). If this was the case, then one-sided incentives may *reduce* the effort of the non-incentivized worker, while raising the effort of the incentivized one. This could, in turn, explain why one-sided incentives are outperformed by two-sided incentives.

We provide three pieces of evidence against this mechanism. First, recall from Section 2.2 that health workers were not told about the introduction of supervisor incentives and few seemed to have learned it from the supervisor.in T_{supv} (resp., T_{shared}), only 15% (resp., 20%) of workers reported that their supervisor receives an incentive. Of these, only 2% (resp., 10%) were aware of the exact amount earned by the supervisor

a 15 minutes visit).

while the rest under-estimated this amount. Second, we do not observe any differential treatment effects on visits depending on the worker's estimated level of inequality aversion (see Figure A.5).[41] Third, we observe that the supervisor's effort is higher (and not lower) in T_{worker} relative to the control group, which cannot be reconciled with supervisors being demotivated by workers receiving incentives. All in all, these three pieces of evidence make it unlikely that inequality aversion alone drives our results.

6 Structural Model

In this section, we use the exogenous variation generated by the interventions to structurally estimate the model presented in Section 3. First, we present our identification and estimation strategy. We then discuss the fit of the empirical and simulated moments. Finally, we present parameter estimates, and conclude with a set of counterfactual policy exercises.

6.1 Identification and Estimation

Our main objective is to estimate the parameters capturing the complementarity of effort γ and the size of the contractual friction z. Additionally, we want to estimate the remaining parameters of the model: the cost of effort c_1 and c_2, the baseline incentives b_1 and b_2, and α, the marginal product of worker effort in the absence of complementarities.

We identify these seven parameters using a core set of eight empirical moments. The mean of output (household visits) in the four experimental conditions gives us an initial set of four moments, and the mean of supervisor effort in the four conditions gives us an additional four moments.[42] The parameters are jointly identified by these empirical moments.

[41] Inequality aversion is measured by asking each health worker the following hypothetical questions: "There is a local farm that hires workers to help with the potato harvest. Sheka accepts a contract to work at the farm for 20,000 SLL per day. He arrives at work the next morning. The farm is very big and there is one supervisor for the 20 workers helping with the harvest. He learns that his supervisor gets paid [*amount*] SLL per day. Do you think Sheka will show up to work the next day?," and *amount* = {20,000; 30,000; 120,000}. Our measure of inequality aversion takes value 0 if the worker answers that Sheka would always show up to work, regardless of the amount; value 1 if worker reports that Sheka would not show up only if $amount = 120,000$ and would show up otherwise; value 2 if the worker reports that Sheka would not show up if $amount \geq 30,000$. Inequality aversion takes value 0 for 26% of the observations, value 1 for 66% of the observations and value 2 for 8%.

[42] The measure of visits we use for the structural analysis is total visits per month. We obtain this by multiplying the number of visits per month per surveyed household by the number of households served by the CHW (as reported in Table A.3, column 2). We then adjust this number down to reflect the fact that some visits are not reported. The measure of supervisor effort we use is the fraction of household visits in which the health worker was accompanied by the supervisor (as reported in Table 6, column 1).

The model makes additional predictions about side payments and worker effort. We have an empirical measure of side payments and, in a robustness exercise, we estimate a model that uses the eight original moments plus three moments capturing the average side payment in each treatment group. However, we do not have good data on worker effort, since for the worker it is hard to obtain a clean measure of effort that is empirically distinct from output (household visits). We thus do not use any moment describing worker effort.

To estimate the model we use a classical minimum distance estimator (Wooldridge 2010). We save the empirical moments in a vector $\boldsymbol{m}$. For a parameter vector $\boldsymbol{\theta}$, we solve the model and calculate the simulated moments $\boldsymbol{m}_S(\boldsymbol{\theta})$. We update $\boldsymbol{\theta}$ in order to solve:

$$\hat{\boldsymbol{\theta}} = \min_{\boldsymbol{\theta}} \left[\boldsymbol{m}_S(\boldsymbol{\theta}) - \boldsymbol{m}\right]' \cdot J(\boldsymbol{m})^{-1} \cdot \left[\boldsymbol{m}_S(\boldsymbol{\theta}) - \boldsymbol{m}\right]. \tag{8}$$

$J(\boldsymbol{m})$ is a diagonal matrix that contains the variance of each moment, ensuring that more precisely estimated moments get a greater weight in estimation.[43] We calculate $J(\boldsymbol{m})$ using a bootstrap with 1,000 replications. Table 10 presents our main structural results and Table 11 describes the empirical fit of the simulated moments.

The estimated model fits the empirical moments tightly. In particular, the model matches both the moments related to supervisor effort and those related to household visits. Crucially, the model is able to reproduce the key result that visits are maximized by the shared incentives treatment. Finally, the estimated model predicts that the supervisor does not offer any side payment to the worker. Indeed, side payments are a very small fraction of the incentive and thus cannot constitute a major motivating factor for the worker. The model is able to reproduce this qualitative fact without using any information on side payments. Importantly, when we estimate the model again using side payments, we obtain a set of parameter estimates that is remarkably similar to the original set of estimates (see Tables A.10 and A.11).

6.2 Parameter Estimates

Our structural estimates show that worker and supervisor effort are strongly complementary, and that contracting through side payments is very costly. We show all parameter estimates in Table 10.

The estimated complementarity parameter γ determines a substantial increase in the marginal product of worker effort. Compared to a setting where $\gamma = 0$, the number

[43] To reflect the fact that visits are more precisely measured than supervisor effort, we apply a penalty to the weight that supervisor effort gets in the loss function. This ensures the estimator prioritizes fitting the moments related to visits, compared to the moments related to supervisor effort.

of household visits generated by a unit of worker effort is 17% larger when the supervisor exerts the control level of effort, and 36% larger when the supervisor exerts the shared incentives level of effort. Supervisor effort thus plays a key role in enabling the worker to carry out household visits, and this results in a strong strategic complementarity between the efforts of the two agents.

Second, we find that the baseline incentive of the supervisor to exert effort in the control group (b_2) is lower than that of the worker (b_1). This is not surprising, since their role is probably harder to monitor and incentivize, especially if the organization under-appreciates the strategic complementarity of effort and thus the possibility to incentivize the supervisor through piece rates based on final output. The supervisor also has a high unit cost of effort (c_2). As a result, interventions that fail to incentivize the supervisor are bound to be highly ineffective: the contribution of the supervisor is key to ensure the worker can be productive, but, absent additional incentive, the supervisor will under-provide her key support to the worker.

Finally, our estimate of the parameter z highlights the important role of contractual frictions. Offering a side payment is very expensive to the supervisor: to transfer 1,000 SLL the supervisor has to pay a total cost of 3,000 SLL. This severely limits the supervisor's ability to redistribute the incentive — in line with the limited amount of side payments we documented in Table 8. However, it should be noted that our estimate of z is not well identified in the current model. Given the asymmetry in benefits documented in the previous point, changes in z do not immediately lead to changes in side payments, which remain set to zero. As a result, this estimate should be considered provisional.

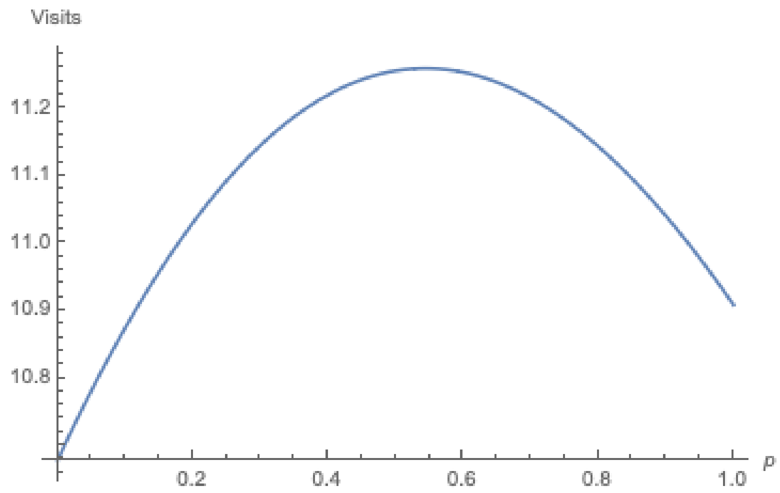
6.3 Counterfactual Policies

We conduct three counterfactual policy experiments that explore, in turn, how to optimally share the incentive between the two agents, how the optimal incentive changes as key structural parameters vary, and the impact of an alternative policy that directly incentivizes effort.

We find that offering an equal share of the incentive to the worker and the supervisor is almost optimal. In Figure 9, we show that, in order to maximize household visits, the worker should be offered 54% of the overall incentive, which is very close to the equal share that we offered in the shared incentives treatment. In other words, given the strong complementarity and large contractual frictions we have estimated, the optimal incentive scheme is one that rewards both agents with a similar payment.[44]

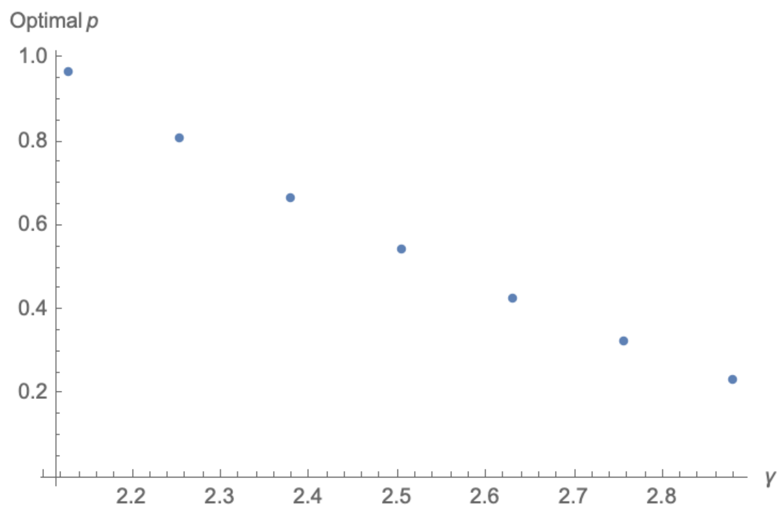
[44]This is a similar exercise than the one done in the simulations shown in Figure 1, but here we are using the estimated parameters from the model to simulate the optimal incentive split between the layers.

Figure 9: Optimal Incentive p^*



This result, however, depends strongly on the strength of the complementarity between worker and supervisor effort.[45] We illustrate this point with our second counterfactual experiment in Figure 10. Here, we plot the optimal share of the incentive offered to the worker (p) for different levels of complementarity. A key result that emerges from this analysis is that, as the complementarity parameter shrinks, the optimal incentive offered to the worker increases substantially. A second important result is that, if the complementarity was stronger than we estimated, it would be optimal to give a substantially larger share of the payment to the supervisor.

Figure 10: Optimal Incentive p^* by Complementarity γ

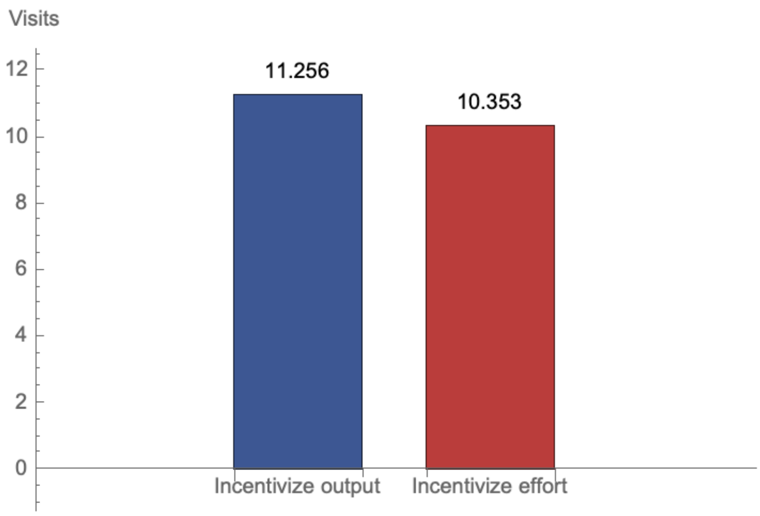


Our final key result highlights that tying incentives to joint output is more effective

[45]The optimal incentive split is expected to differ across organizations, depending on the level of complementarity across layers. It can also differ over time within a given organization, e.g., if the organization increases complementarities by improving the communication systems across layers, increasing team cohesion, etc.

than directly incentivizing effort (e.g., incentivizing supervisors on the number of times they advise/train the health workers and incentivizing health workers on the number of times they attempt to approach a household). In Figure 11, we compare the maximum number of visits that are generated through (i) a scheme that equally shares a payment of 2,000 SLL per visit between the worker and the supervisor, and (ii) a scheme of the same cost that optimally offers incentives directly tied to individual effort.[46] What emerges is that, at the current level of complementarity, incentivizing output through an equally-shared piece rate generates 8% more visits that optimally incentivizing effort, for the same cost. This is because, when efforts are highly complementary, output incentives implicitly help agents internalize the positive external effect that their effort has on the other player. This makes output incentives particularly effective.

Figure 11: An Alternative Policy that Targets Effort



7 Conclusion

This paper provides novel evidence on the optimal structure of performance incentives in large organizations, explicitly taking into account the presence of multiple hierarchical layers. Many incentive schemes economists studied in the past target frontline workers (e.g. teachers, health workers, tax collectors) rather than their superiors. Take, for example, Glewwe, Ilias, and Kremer (2010); Muralidharan and Sundararaman (2011); Duflo, Hanna, and Ryan (2012); Ashraf, Bandiera, and Jack (2014) among many others. In line with this, 52% of economists who participated in a prediction survey forecasted the one-sided worker incentives to be the most effective in our context (vs. 28% who

[46]In this comparison, we assume that effort can be observed and is perfectly predictive of output. Hence, we abstract from issues related to asymmetric information, which may decrease the effectiveness of both incentive schemes.

chose the shared incentives, 4% who chose the one-sided supervisor incentives, and 18% who forecasted either two or all three treatments to have the same effect).[47] However, we find in this paper that one-sided incentives are not always optimal. In the context of the national Community Health Program in Sierra Leone, we show that a piece-rate incentive that is equally shared between workers and supervisors leads to an increase of health visits that is 61% higher compared to the increase when the same piece-rate incentive is either only given to workers or only given to supervisors.

Our theoretical model highlights the two conditions that need to hold for shared incentives to be the optimal incentive structure of an organization in general. First, the production function of the final output needs to exhibit effort complementarities between the worker and the supervisor. Second, there need to be contractual frictions or benefit asymmetries that constrain the redistribution of incentives by the supervisor through side payments. Thus, compared to the context we study, we would expect the effectiveness of shared incentives to be weaker in organizations in which the production function of final output does not exhibit effort complementarities between workers and their supervisors. One example for this are organizations in which the role of the supervisor is limited to monitoring, distributing tasks or to making personnel decisions, but not to train and advise workers. Similarly, we would also expect the choice of which layer to incentivize to be less relevant in organizations in which side payments regularly occur across layers and where there is scope for Coasian bargaining.

A key insight that follows from this discussion is that real-world organizations should assess the extent to which effort complementarities are present when deciding about the structure of performance incentives for their workforce. In our context, local experts appear capable of making these assessments: 92% of the supervisors who participated in the study predicted the shared incentives to maximize household visits.[48] Whether organizations should rely on local experts in designing the incentives, or whether they should calibrate incentives using more sophisticated tools, is an open question and requires further research.

[47] Appendix D provides details about the prediction survey.

[48] 4% of the supervisors predicted the one-side worker incentives to maximize visits, while another 4% predicted the one-side supervisors incentives to maximize visits. This is again consistent with the presence of contractual frictions: supervisors know that the shared incentives are the best split but still do not transfer half of the incentive payments to their subordinates.

References

Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2016. "Explaining Causal Findings without Bias: Detecting and Assessing Direct Effects." *The American Political Science Review* 110 (3):512.

Alchian, Armen A. and Harold Demsetz. 1972. "Production, information costs, and economic organization." *American Economic Review* 62 (5):777–795.

Alsan, Marcella. 2015. "The Effect of the Tsetse Fly on African Development." *American Economic Review* 105 (1):382–410.

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

Babcock, Philip, Kelly Bedard, Gary Charness, John Hartman, and Heather Royer. 2015. "Letting Down the Team? Social Effects of Team Incentives." *Journal of the European Economic Association* 13 (5):841–870.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2007a. "Incentives for Managers and Inequality Among Workers: Evidence from a Firm Level Experiment." *The Quarterly Journal of Economics* 122 (2):729–773.

———. 2007b. "Incentives for managers and inequality among workers: Evidence from a firm-level experiment." *The Quarterly Journal of Economics* 122 (2):729–773.

Bandiera, Oriana, Michael Carlos Best, Adnan Qadir Khan, and Andrea Prat. 2021. "The Allocation of Authority in Organizations: A Field Experiment with Bureaucrats." *The Quarterly Journal of Economics* Forthcoming.

Banerjee, Abhijit, Esther Duflo, Clement Imbert, Santhosh Mathew, and Rohini Pande. 2020. "E-governance, Accountability, and Leakage in Public Programs: Experimental Evidence from a Financial Management Reform in India." *American Economic Journal: Applied Economics* 12 (4):39–72.

Basinga, Paulin, Paul J. Gertler, Agnes Binagwaho, Agnes L.B. Soucat, Jennifer Sturdy, and Christel M.J. Vermeersch. 2011. "Effect on Maternal and Child Health Services in Rwanda of Payment to Primary Health-Care Providers for Performance: An Impact Evaluation." *The Lancet* 377 (9775):1421–1428.

Behrman, Jere R., Susan W. Parker, Petra E. Todd, and Kenneth I. Wolpin. 2015. "Aligning Learning Incentives of Students and Teachers: Results from a Social Experiment in Mexican High Schools." *Journal of Political Economy* 123 (2):325–364.

Bertrand, Marianne. 2009. "CEOs." *Annual Review of Economics* 1 (1):121–150.

Breza, Emily, Supreet Kaur, and Yogita Shamdasani. 2018. "The Morale Effects of Pay Inequality." *The Quarterly Journal of Economics* 133 (2):611–663.

Bubb, Ryan, Supreet Kaur, and Sendhil Mullainathan. 2018. "The Limits of Neighborly Exchange." Working Paper.

Callen, Michael, Saad Gulzar, Ali Hasanain, Muhammad Yasir Khan, and Arman Rezaee. 2020. “Data and Policy Decisions: Experimental Evidence from Pakistan.” *Journal of Development Economics* 146.

Campbell, Jim, Gilles Dussault, James Buchan, Francisco Pozo-Martin, Maria Guerra Arias, Claudia Leone, Amani Siyam, and Giorgio Cometto. 2013. “A Universal Truth: No Health without a Workforce.” Tech. rep., Global Health Workforce Alliance and World Health Organization.

Casaburi, Lorenzo and Rocco Macchiavello. 2019. “Demand and Supply of Infrequent Payments as a Commitment Device: Evidence from Kenya.” *American Economic Review* 109 (2):523–55.

Christensen, Darin, Oeindrila Dube, Johannes Haushofer, Bilal Siddiqi, and Maarten Voors. 2021. “Building Resilient Health Systems: Experimental Evidence from Sierra Leone and The 2014 Ebola Outbreak.” *The Quarterly Journal of Economics* 136 (2):1145–1198.

Cilliers, Jacobus, Ibrahim Kasirye, Clare Leaver, Pieter Serneels, and Andrew Zeitlin. 2018. “Pay for Locally Monitored Performance? A Welfare Analysis for Teacher Attendance in Ugandan Primary Schools.” *Journal of Public Economics* 167:69–90.

Coase, Ronald H. 1937. “The Nature of the Firm.” *Economica* 4 (16):386–405.

Cullen, Zoë B. and Ricardo Perez-Truglia. 2019. “The Old Boys’ Club: Schmoozing and the Gender Gap.” Working Paper.

———. 2021. “How Much Does Your Boss Make? The Effects of Salary Comparisons.” *Journal of Political Economy* (Forthcoming).

Dal Bó, Ernesto, Frederico Finan, Nicholas Li, and Laura Schechter. 2021. “Information Technology and Government Decentralization: Experimental Evidence from Paraguay.” *Econometrica* (Forthcoming).

Deserranno, Erika. 2019. “Financial Incentives as Signals: Experimental Evidence from the Recruitment of Village Promoters in Uganda.” *American Economic Journal: Applied Economics* 11:277–317.

Deserranno, Erika, Philipp Kastrau, and Gianmarco León-Ciliotta. 2021. “Promotions and Productivity: The Role of Meritocracy and Pay Progression in the Public Sector.” Working Paper.

Deserranno, Erika, Aisha Nansamba, and Nancy Qian. 2020. “Aid Crowd-Out: The Effect of NGOs on Government-Provided Services.” Working Paper.

Dodge, Eric, Yusuf Neggers, Rohini Pande, and Charity Troyer Moore. 2018. “Updating the State: Information Acquisition Costs and Public Benefit Delivery.” Working Paper.

Duflo, Esther, Rema Hanna, and Stephen P. Ryan. 2012. “Incentives Work: Getting Teachers to Come to School.” *American Economic Review* 102 (4):1241–78.

Friebel, Guido, Matthias Heinz, Miriam Krueger, and Nikolay Zubanov. 2017. “Team Incentives and Performance: Evidence from a Retail Chain.” *American Economic Review* 107 (8):2168–2203.

Frydman, Carola and Dirk Jenter. 2010. “CEO Compensation.” *Annual Review of Financial Economics* 2 (1):75–102.

Gibbons, Robert. 1996. “Incentives and Careers in Organizations.” Working Paper.

———. 2005. “Four Formal(izable) Theories of the Firm?” *Journal of Economic Behavior & Organization* 58 (2):200–245.

Glewwe, Paul, Nauman Ilias, and Michael Kremer. 2010. “Teacher Incentives.” *American Economic Journal: Applied Economics* 2 (3):205–27.

Hines, James R. and Richard H. Thaler. 1995. “The Flypaper Effect.” *Journal of Economic Perspectives* 9 (4):217–226.

Holmström, Bengt. 2017. “Pay for Performance and Beyond.” *American Economic Review* 107 (7):1753–77.

Itoh, Hideshi. 1991. “Incentives to Help in Multi-Agent Situations.” *Econometrica* 59 (3):611–636.

Jackson, Matthew O. and Simon Wilkie. 2005. “Endogenous Games and Mechanisms: Side Payments Among Players.” *The Review of Economic Studies* 72 (2):543–566.

Karing, Anne. 2021. “Strengthening State Capacity in Health with Administrative Data in Sierra Leone.” Working Paper.

Lafontaine, Francine and Margaret Slade. 2007. “Vertical Integration and Firm Boundaries: The Evidence.” *Journal of Economic Literature* 45 (3):629–685.

Lameke, Aimable A., Albert J. Malukisa, Raúl Sánchez de la Sierra, and Kristof Titeca. 2020. “Corruption (with a Hierarchy).” Working Paper.

Lazear, Edward P. 2000. “Performance Pay and Productivity.” *American Economic Review* 90 (5):1346–1361.

Lowes, Sara Rachel and Eduardo Montero. 2021. “The Legacy of Colonial Medicine in Central Africa.” *American Economic Review* 111 (4):1284–1314.

Luo, Renfu, Grant Miller, Scott Rozelle, Sean Sylvia, and Marcos Vera-Hernández. 2019. “Can Bureaucrats Really Be Paid Like CEOs? Substitution Between Incentives and Resources Among School Administrators in China.” *Journal of the European Economic Association* 18 (1):165–201.

Macchiavello, Rocco and Ameet Morjaria. 2015. “The Value of Relationships: Evidence from a Supply Shock to Kenyan Rose Exports.” *American Economic Review* 105 (9):2911–45.

———. 2021. “Competition and Relational Contracts in the Rwanda Coffee Chain.” *The Quarterly Journal of Economics* 134 (2):895–951.

Martinez-Bravo, Monica and Andreas Stegmann. Forthcoming. “In Vaccines We Trust? The Effects of the CIA’s Vaccine Ruse on Immunization in Pakistan.” *Journal of the European Economic Association* .

McMillan, John and Christopher Woodruff. 1999. “Interfirm Relationships and Informal Credit in Vietnam.” *The Quarterly Journal of Economics* 114 (4):1285–1320.

Milgrom, Paul and John Roberts. 1988. “An Economic Approach to Influence Activities in Organizations.” *American Journal of Sociology* 94:S154–S179.

Muralidharan, Karthik, Paul Niehaus, Sandip Sukhtankar, and Jeffrey Weaver. 2021. “Improving Last-Mile Service Delivery Using Phone-Based Monitoring.” *American Economic Journal: Applied Economics* 13 (2):52–82.

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

Nyqvist, Martina Björkman, Andrea Guariso, Jakob Svensson, and David Yanagizawa-Drott. 2019. “Reducing Child Mortality in the Last Mile: Experimental Evidence on Community Health Promoters in Uganda.” *American Economic Journal: Applied Economics* 11 (3):155–92.

Ord, Toby. 2020. *The Precipice: Existential Risk and the Future of Humanity*. Hachette Books.

Rasul, Imran and Daniel Rogger. 2018. “Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service.” *The Economic Journal* 128 (608):413–446.

Ray, Debraj, Jean-Marie Baland, and Olivier Dagnelie. 2007. “Inequality and Inefficiency in Joint Projects.” *The Economic Journal* 117 (522):922–935.

Tirole, Jean. 1986. “Hierarchies and Bureaucracies: On the Role of Collusion in Organizations.” *Journal of Law, Economics, and Organization* 2 (2):181–214.

———. 1992. “Collusion and the Theory of Organizations.” In *Advances in Economic Theory*, vol. 2, edited by Jean-Jacques Laffont, chap. 3. Cambridge University Press.

Weaver, Jeffrey. 2021. “Jobs for Sale: Corruption and Misallocation in Hiring.” *American Economic Review* (Forthcoming).

Williamson, Oliver E. 1973. “Markets and Hierarchies: Some Elementary Considerations.” *American Economic Review* 63 (2):316–325.

Wilson, James Q. 1989. *Bureaucracy: What Government Agencies Do And Why They Do It.* Basic Books New York.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data.* MIT press.

World Health Organization. 2016. *Health in 2015: From MDGs, Millennium Development Goals to SDGs, Sustainable Development Goals*. Geneva: WHO Press, 2016.

———. 2017. "Global Health Observatory." URL https://www.who.int/data/gho/data/indicators.

Table 1: Summary Statistics and Balance Checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Sample of villages:	**All**		**Control**		**Worker Incentives Treatment**		**Supervisor Incentives Treatment**		**Shared Incentives Treatment**		F-test of joint significance
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	
A. Characteristics of the health workers											
Number of health workers	2,970		703		777		745		745		
Male = {0, 1}	0.708	0.455	0.727	0.446	0.721	0.449	0.710	0.454	0.675	0.469	*0.407*
Age (in years)	37.121	11.471	35.946	11.137	37.785	11.719	37.482	11.720	37.174	11.206	*0.062*
Completed primary education = {0, 1}	0.697	0.460	0.727	0.446	0.694	0.461	0.703	0.457	0.666	0.472	*0.301*
Completed secondary education or above = {0, 1}	0.077	0.267	0.070	0.255	0.076	0.265	0.078	0.268	0.085	0.278	*0.867*
Wealth score (0 to 8)	2.454	1.167	2.430	1.231	2.400	1.116	2.438	1.120	2.550	1.199	*0.273*
Number of households the health worker is responsible for	55.186	78.587	62.723	120.184	54.077	62.922	53.159	56.374	51.258	60.238	0.375
Number of visits provided per week	20.870	19.588	19.650	20.586	22.388	23.088	21.140	17.824	20.172	15.919	*0.271*
Number of hours worked as health worker per week	17.625	31.187	20.717	44.984	18.322	33.688	15.452	21.081	16.154	18.097	*0.102*
B. Characteristics of the supervisors											
Number of supervisors	372		93		93		93		93		
Male = {0, 1}	0.919	0.273	0.925	0.265	0.925	0.265	0.914	0.282	0.914	0.282	*0.988*
Age (in years)	37.839	8.856	37.914	9.329	37.462	7.869	36.849	8.690	39.129	9.433	*0.316*
Completed primary education = {0, 1}	0.739	0.440	0.763	0.427	0.731	0.446	0.785	0.413	0.677	0.470	*0.405*
Completed secondary education or above = {0, 1}	0.253	0.435	0.226	0.420	0.269	0.446	0.215	0.413	0.301	0.461	*0.533*
Wealth score (0 to 8)	3.013	1.227	3.097	1.269	2.914	1.222	2.914	1.239	3.129	1.182	*0.507*
Number of health workers supervisor is responsible for	7.984	2.861	7.559	2.799	8.355	2.831	8.011	2.902	8.011	2.899	*0.289*
Number of hours worked as supervisor per week	11.156	33.967	8.432	19.539	8.337	11.556	7.581	15.691	19.795	60.693	*0.378*

Notes: Each row states the sample mean and standard deviation of a variable in the entire sample of villages, and by treatment group. Col. 11 reports the p-value from the F-test of joint significance of the treatment dummies. This is calculated from a regression of each variable on the 3 treatment dummies, controlling for the stratification variables and clustering standard errors at the PHU level in worker-level regressions (Panel A).

Table 2: Visits

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Household visits provided by the health worker in the past 6 months						
Dep. Var.	Number of visits	% households visited	Number of visit types	Average visit length (conditional on at least 1 visit)	Average visit length (imputed with 0 when no visit received)	Number of health topics discussed	% households who trust the health worker as a health provider
Worker incentives	2.090***	0.072***	0.250***	0.603	2.024**	0.164	0.037
	(0.558)	(0.025)	(0.094)	(1.423)	(0.941)	(0.127)	(0.023)
Supervisor incentives	2.145***	0.082***	0.325***	-0.070	1.933**	0.173	0.031
	(0.501)	(0.025)	(0.100)	(1.409)	(0.926)	(0.130)	(0.024)
Shared incentives	3.356***	0.127***	0.565***	1.590	4.134***	0.528***	0.071***
	(0.492)	(0.023)	(0.092)	(1.290)	(0.927)	(0.134)	(0.024)
Unit	Worker	Worker	Worker	Worker	Worker	Worker	Worker
Observations	2,926	2,926	2,926	1,803	2,926	2,926	2,926
Mean dep. var.	7.296	0.709	1.745	23.404	14.388	2.248	0.745
Mean dep. var. in Control	5.334	0.637	1.448	22.736	12.324	2.015	0.707
p-value Worker = Supervisor	0.932	0.710	0.492	0.635	0.927	0.946	0.808
p-value Supervisor = Shared	0.040	0.060	0.024	0.196	0.024	0.017	0.102
p-value Worker = Shared	0.046	0.026	0.002	0.451	0.033	0.013	0.147

Notes: All regressions include stratification variables. Standard errors clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Health Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Pre- and post-natal care in the past 2 years						Disease incidence in the past month			
Dep. Var.	Index (cols. 2-6)	% women who received at least 4 ante-natal visits before birth	% women with institutional birth	% women who received post-natal visit within 2 days of birth	% women with at least 6 months of breastfeeding	% households with up-to-date infant vaccination	Index (cols. 8-10)	% households with child under-5 who had fever	% households with child under-5 who had cough	% households with child under-5 who had diarrhea
Worker incentives	0.029	0.017	0.022	0.007	-0.002	-0.008	0.010	-0.028	0.016	0.005
	(0.028)	(0.024)	(0.021)	(0.026)	(0.025)	(0.018)	(0.032)	(0.022)	(0.012)	(0.007)
Supervisor incentives	0.042	0.032	0.035*	-0.022	0.016	0.015	-0.028	-0.014	-0.005	-0.005
	(0.029)	(0.026)	(0.019)	(0.024)	(0.025)	(0.019)	(0.032)	(0.028)	(0.012)	(0.005)
Shared incentives	0.092***	0.058**	0.036*	0.017	0.040	0.025	-0.053**	-0.058***	-0.007	-0.001
	(0.028)	(0.025)	(0.019)	(0.027)	(0.025)	(0.019)	(0.026)	(0.022)	(0.011)	(0.005)
Unit	Worker	Worker	Worker	Worker	Worker	Worker	Worker	Worker	Worker	Worker
Observations	2,499	2,499	2,499	2,499	2,499	2,499	2,826	2,823	2,825	2,826
Mean dep. var.	-0.006	0.778	0.868	0.305	0.666	0.230	-0.009	0.183	0.072	0.016
Mean dep. var. in Control	-0.048	0.750	0.845	0.303	0.652	0.222	0.009	0.208	0.071	0.017
p-value Worker = Supervisor	0.656	0.509	0.491	0.258	0.439	0.223	0.273	0.580	0.088	0.129
p-value Supervisor = Shared	0.077	0.243	0.963	0.138	0.330	0.630	0.397	0.086	0.868	0.343
p-value Worker = Shared	0.024	0.052	0.469	0.703	0.077	0.091	0.025	0.105	0.056	0.360

Notes: All regressions include stratification variables. Standard errors clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1. The index in col. 1 [resp., col. 7] estimates an equally weighted average of the z-scores of variables in cols. 1-5 [resp., cols. 7-9].

Table 4: Reporting

	(1)	(2)	(3)	(4)	(5)
Dep. Var.	Number of reports per month	Reporting rate = number of reports/ number of visits	Over-reporting = {0, 1} [number of reports > number of visits]	Under-reporting = {0, 1} [number of reports < number of visits]	Workers believes the incentive they receive is "too little" = {0, 1}
Worker incentives	7.182***	0.222***	0.046**	-0.002	-0.244***
	(0.759)	(0.043)	(0.018)	(0.024)	-0.028
Supervisor incentives	2.181***	0.058*	-0.006	0.041*	0.005
	(0.586)	(0.034)	(0.015)	(0.021)	(0.023)
Shared incentives	4.757***	0.090**	0.009	0.046**	-0.122***
	(0.829)	(0.036)	(0.015)	(0.021)	(0.028)
Unit	Worker	Worker	Worker	Worker	Worker
Observations	2,970	2,624	2,926	2,926	2,709
Mean dep. var.	5.213	0.177	0.069	0.862	0.784
Mean dep. var. in Control	1.525	0.078	0.055	0.840	0.877
p-value Worker = Supervisor	<0.001	<0.001	0.004	0.059	<0.001
p-value Supervisor = Shared	0.008	0.398	0.333	0.802	<0.001
p-value Worker = Shared	0.026	0.004	0.050	0.035	<0.001

Notes: In col. 2, the reporting rate is measured as the number of reports (col. 1) divided by the number of household visits. In cols. 3-4, under (over) reporting is an indicator for whether the number of reports is lower (higher) than the number of household visits. All regressions include stratification variables. Standard errors clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Incentives Payouts

	(1)	(2)	(3)	(4)	(5)
	Monthly incentive payments (in 1,000 SLL)				
Dep. Var.	Payment per health worker	Payment for all health workers	Payment per supervisor	Total cost (cols. 2+3)	Cost per visit (col. 4/number of visits)
Worker incentives	16.054***	132.903***	-1.310	131.593***	0.355***
	(1.352)	(12.000)	(1.832)	(11.962)	(0.045)
Supervisor incentives	-0.047	-1.172	55.280***	54.108***	0.177***
	(0.185)	(2.631)	(8.095)	(8.377)	(0.038)
Shared incentives	5.584***	46.857***	47.097***	93.953***	0.210***
	(0.691)	(6.680)	(6.534)	(13.164)	(0.032)
Unit	PHU	PHU	PHU	PHU	PHU
Observations	372	372	372	372	371
Mean dep. var.	5.431	45.535	25.858	71.392	0.185
Mean dep. var. in Control	0.000	0.000	0.000	0.000	0.000
p-value Worker = Supervisor	<0.001	<0.001	<0.001	<0.001	0.002
p-value Supervisor = Shared	<0.001	<0.001	0.428	0.009	0.507
p-value Worker = Shared	<0.001	<0.001	<0.001	0.032	0.008

Notes: All regressions include stratification variables. Standard errors clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Supervisor Effort

	(1)	(2)	(3)	(4)
	In-the-field supervision	*One-to-one meetings*	*General training*	
Dep. Var.	% household visits in which health worker was accompanied by supervisor in the past 6 months	Number of times supervisor visited health worker in the past 6 months	Supervisor organized training in the past month = {0, 1}	Supervisor ever helped health worker with SMS reporting = {0, 1}
Worker incentives	0.030	-0.050	0.004	-0.018
	(0.022)	(0.137)	(0.005)	(0.032)
Supervisor incentives	0.057**	-0.041	0.006	-0.007
	(0.023)	(0.137)	(0.005)	(0.035)
Shared incentives	0.062***	-0.043	0.003	-0.059**
	(0.021)	(0.139)	(0.005)	(0.029)
Unit	Worker	Worker	Worker	Worker
Observations	2,919	2,833	2,864	2,813
Mean dep. var.	0.204	1.375	0.994	0.158
Mean dep. var. in Control	0.164	1.417	0.991	0.181
p-value Worker = Supervisor	0.293	0.950	0.443	0.717
p-value Supervisor = Shared	0.846	0.987	0.463	0.075
p-value Worker = Shared	0.181	0.963	0.916	0.120

Notes: Variable in col. 1 is reported by the household and aggregated at worker level. Variables in cols. 2-4 are reported by the health worker. All regressions include stratification variables. Standard errors clustered at the PHU level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 7: Visits and Supervisor Effort by Worker Experience

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Output	Supervisor effort	Output	Supervisor effort	Output	Supervisor effort	Output	Supervisor effort
Dep. Var.	Number of visits	% household visits in which health worker was accompanied by supervisor	Number of visits	% household visits in which health worker was accompanied by supervisor	Number of visits	% household visits in which health worker was accompanied by supervisor	Number of visits	% household visits in which health worker was accompanied by supervisor
Definition of covariate X:	Years of experience as a health worker at baseline > median (4 years) = {0, 1}		Number of years of experience of the health worker at baseline (demeaned)		Years of experience as a health worker at baseline > median (4 years) = {0, 1}		Number of years of experience of the health worker at baseline (demeaned)	
Worker incentives	2.054***	0.030	2.174***	0.029	2.130***	0.027	2.235***	0.026
	(0.628)	(0.025)	(0.565)	(0.022)	(0.650)	(0.026)	(0.549)	(0.022)
Supervisor incentives	2.576***	0.067**	2.209***	0.057**	2.654***	0.061**	2.314***	0.055**
	(0.598)	(0.029)	(0.500)	(0.023)	(0.628)	(0.030)	(0.511)	(0.023)
Shared incentives	4.022***	0.092***	3.397***	0.064***	4.133***	0.089***	3.455***	0.063***
	(0.684)	(0.026)	(0.493)	(0.021)	(0.667)	(0.027)	(0.479)	(0.022)
X	1.057**	0.017	0.114*	0.002	1.038*	0.014	0.122**	0.002
	(0.532)	(0.025)	(0.059)	(0.003)	(0.562)	(0.026)	(0.060)	(0.003)
Worker incentives * X	0.192	-0.000	-0.084	-0.005	0.173	0.001	-0.092	-0.005
	(0.863)	(0.034)	(0.098)	(0.004)	(0.898)	(0.035)	(0.101)	(0.004)
Supervisor incentives * X	-0.856	-0.022	-0.089	-0.005	-0.811	-0.013	-0.089	-0.004
	(0.775)	(0.035)	(0.097)	(0.004)	(0.801)	(0.038)	(0.099)	(0.004)
Shared incentives * X	-1.439*	-0.061*	-0.179**	-0.007**	-1.549*	-0.058	-0.195**	-0.007**
	(0.857)	(0.036)	(0.083)	(0.003)	(0.849)	(0.036)	(0.083)	(0.003)
Unit	Worker	Worker	Worker	Worker	Worker	Worker	Worker	Worker
Extra Controls	No	No	No	No	Yes	Yes	Yes	Yes
Observations	2,909	2,902	2,909	2,902	2,886	2,879	2,886	2,879
Mean Dep. Var.	7.296	0.204	7.296	0.204	7.296	0.204	7.296	0.204
Mean Dep. Var. in Control	5.334	0.164	5.334	0.164	5.334	0.164	5.334	0.164
Mean X	0.473	0.473	0.000	0.000	0.473	0.473	0.000	0.000
p-value Worker*X = Supervisor*X	0.824	0.994	0.393	0.194	0.847	0.971	0.363	0.201
p-value Supervisor*X = Shared*X	0.270	0.535	0.356	0.192	0.312	0.735	0.372	0.310
p-value Worker*X = Shared*X	0.094	0.086	0.033	0.034	0.069	0.107	0.020	0.044
p-value Worker + Worker*X = 0	0.004	0.343	<0.001	0.301	0.003	0.347	<0.001	0.346
p-value Supervisor + Supervisor*X = 0	0.010	0.136	<0.001	0.027	0.006	0.102	<0.001	0.024
p-value Worker + Worker*X = 0	<0.001	0.313	<0.001	0.009	<0.001	0.287	<0.001	0.011

Notes: All regressions include stratification variables. Cols. 5-8 also control for the entire list of health worker characteristics listed in the summary statistics table, along with their interaction with the three treatment dummies. Standard errors clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Side Payments

	(1)	(2)	(3)	(4)	(5)
			Monthly transfer amount (in 1,000 SLL)		
Dep. Var.	Supervisor shared incentive with health worker = {0, 1}	Health worker shared incentive with supervisor = {0, 1}	… from supervisor to health worker	… from health worker to supervisor	Net transfer (col. 3-4)
Worker incentives	0.005	0.073***	0.110	0.151***	-0.042
	(0.016)	(0.014)	(0.090)	(0.056)	(0.077)
Supervisor incentives	0.183***	-0.001	0.702***	0.104**	0.598***
	(0.047)	(0.008)	(0.190)	(0.043)	(0.190)
Shared incentives	0.102***	0.041***	0.432***	0.084*	0.348**
	(0.039)	(0.015)	(0.158)	(0.043)	(0.164)
Unit	Worker	Worker	Worker	Worker	Worker
Observations	2,915	2,909	2,488	2,488	2,488
Mean dep. var.	0.084	0.049	0.308	0.101	0.207
Mean dep. var. in Control	0.011	0.019	0.000	0.016	-0.016
p-value Worker = Supervisor	<0.001	<0.001	0.004	0.484	0.001
p-value Supervisor = Shared	0.171	0.005	0.273	0.725	0.318
p-value Worker = Shared	0.013	0.100	0.068	0.325	0.026

Notes: All regressions include stratification variables. Standard errors clustered at the PHU level. The sample size drops in col. 3-5 because a number of health workers answered "don't know" to the question. *** p<0.01, ** p<0.05, * p<0.1

Table 9: Side Payments by Effort Observability

	(1)	(2)	(3)	(4)	(5)	(6)
	Side-payments	Supervisor effort	Output	Side-payments	Supervisor effort	Output
Dep. Var.	Supervisor shared incentive with health worker = {0, 1}	% household visits in which health worker was accompanied by supervisor	Number of visits	Supervisor shared incentive with health worker = {0, 1}	% household visits in which health worker was accompanied by supervisor	Number of visits
	Low observability of health worker output			*Low observability of health worker output*		
Definition of covariate X:	Correlation between actual and perceived ranking in bottom decile = {0,1}			Correlation between actual and perceived ranking in bottom decile = {0,1}		
Worker incentives	0.009	0.034	2.015***	-0.001	0.033	2.012***
	(0.018)	(0.024)	(0.586)	(0.020)	(0.024)	(0.589)
Supervisor incentives	0.205***	0.065**	2.640***	0.196***	0.064**	2.638***
	(0.052)	(0.026)	(0.540)	(0.051)	(0.026)	(0.540)
Shared incentives	0.107**	0.060**	3.615***	0.100**	0.059**	3.613***
	(0.042)	(0.023)	(0.537)	(0.040)	(0.024)	(0.540)
X	0.008	-0.015	1.618**	0.016	-0.014	1.620**
	(0.023)	(0.037)	(0.661)	(0.027)	(0.038)	(0.661)
Worker incentives * X	-0.037	-0.050	1.737	-0.027	-0.048	1.740
	(0.037)	(0.063)	(1.616)	(0.042)	(0.063)	(1.619)
Supervisor incentives * X	-0.173**	-0.062	-3.846***	-0.150*	-0.058	-3.840***
	(0.084)	(0.061)	(1.318)	(0.085)	(0.061)	(1.330)
Shared incentives * X	-0.040	0.020	-1.936	-0.030	0.022	-1.933
	(0.097)	(0.059)	(1.287)	(0.094)	(0.060)	(1.294)
Unit	Worker	Worker	Worker	Worker	Worker	Worker
Extra controls	No	No	No	Yes	Yes	Yes
Observations	2,915	2,919	2,926	2,915	2,919	2,926
Mean Dep. Var.	0.084	0.204	7.296	0.084	0.204	7.296
Mean Dep. Var. in Control	0.011	0.164	5.334	0.011	0.164	5.334
p-value Worker*X = Supervisor*X	0.315	0.425	0.283	0.526	0.442	0.283
p-value Supervisor*X = Shared*X	0.040	0.309	0.004	0.078	0.340	0.004
p-value Worker*X = Shared*X	0.680	0.733	0.133	0.745	0.715	0.136
p-value Worker + Worker*X = 0	0.372	0.784	0.013	0.447	0.786	0.013
p-value Supervisor + Supervisor*X = 0	0.618	0.953	0.308	0.497	0.923	0.313
p-value Worker + Worker*X = 0	0.448	0.141	0.145	0.416	0.140	0.145

Notes: All regressions include stratification variables. The regressions in col 4-6 also control for the number of health workers the supervisor is responsible for. Standard errors clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1

Table 10: Parameter Estimates

	(1)
Complementarity γ	2.5
Contractual friction z	3.0
CHW unit cost of effort c_1	14.6
PS unit cost of effort c_2	1612.5
CHW baseline incentive b_1	57.8
PS baseline incentive b_2	19.0
α	1.6
Δ in marginal product of CHW effort (shared incentive)	36 %
Δ in marginal product of CHW effort (no incentive)	17 %
Total CHW cost of effort (no incentive)	55.0
Total PS cost of effort (no incentive)	179.1

Table 11: Moment Fit

Moments	Targeted Real	Simulated
PS effort in CHW incentive group	0.198	0.159
PS effort in PS incentive group	0.225	0.278
PS effort in shared incentive group	0.228	0.230
PS effort in control group	0.164	0.111
Output in CHW incentive group	10.551	10.909
Output in PS incentive group	10.413	10.678
Output in shared incentive group	11.827	11.256
Output in control group	7.256	7.174
Side payment in CHW incentive group (1,000 SSL)		0.00
Side payment in PS incentive group (1,000 SSL)		0.000
Side payment in shared incentive group (1,000 SSL)		0.000
Value loss function	3.7	

Online Appendix

A Appendix Tables and Figures

Table A.1: Balance Checks (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)
	P-values from Pairwise Treatment Comparisons					
	Worker = Supervisor Incentives	Worker = Shared Incentives	Supervisor = Shared Incentives	Worker Incentives = Control	Supervisor Incentives = Control	Shared Incentives = Control
A. Characteristics of the health workers						
Male = {0, 1}	0.912	0.218	0.170	0.678	0.749	0.102
Age (in years)	0.472	0.338	0.838	0.009	0.067	0.088
Completed primary education = {0, 1}	0.812	0.329	0.201	0.405	0.528	0.059
Completed secondary education or above = {0, 1}	0.944	0.708	0.738	0.666	0.590	0.397
Wealth score (0 to 8)	0.915	0.138	0.112	0.835	0.736	0.094
Number of households the health worker is responsible for	0.532	0.353	0.711	0.291	0.138	0.096
Number of visits provided per week	0.379	0.114	0.363	0.085	0.258	0.715
Number of hours worked as health worker per week	0.078	0.174	0.549	0.405	0.048	0.087
B. Characteristics of the supervisors						
Male = {0, 1}	0.823	0.869	0.957	0.923	0.750	0.796
Age (in years)	0.664	0.151	0.074	0.615	0.370	0.371
Completed primary education = {0, 1}	0.399	0.454	0.109	0.592	0.748	0.195
Completed secondary education or above = {0, 1}	0.395	0.671	0.199	0.473	0.883	0.249
Wealth score (0 to 8)	0.901	0.285	0.215	0.371	0.295	0.888
Number of health workers supervisor is responsible for	0.375	0.450	0.904	0.054	0.304	0.253
Number of hours worked as supervisor per week	0.873	0.092	0.079	0.982	0.875	0.103

Notes: Each row presents p-values from pairwise treatment comparisons. These are calculated from a regression, where the variable is regressed on an indicator for worker, supervisor and shared incentives treatment. All regressions include stratification variables. Standard errors clustered at the PHU level in worker-level regressions (Panel A).

Table A.2: Summary Statistics and Balance Checks (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
	All		Control		Worker incentives		Supervisor incentives		Shared incentives		F-test of joint significance	P-values from pairwise treatment comparisons					
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.		Worker = Supervisor	Worker = Shared	Supervisor = Shared	Worker = Control	Supervisor = Control	Shared = Control
A. Characteristics of the villages																	
Number of observations	2,970		703		777		745		745								
Number of households in the village	214.148	5528.828	60.958	87.573	62.797	91.462	492.928	10316.584	246.736	4140.176	0.522	0.295	0.266	0.617	0.787	0.295	0.280
Accessible road to health facility= {0, 1}	0.766	0.424	0.778	0.416	0.762	0.426	0.775	0.418	0.747	0.435	0.797	0.784	0.511	0.361	0.809	0.991	0.400
Distance to health facility (in km)	6.678	21.687	4.540	8.384	6.802	20.490	10.518	34.632	4.105	6.468	0.468	0.043	0.629	0.145	0.555	0.195	0.881
Phone network available nowhere	0.007	0.084	0.006	0.075	0.005	0.072	0.003	0.052	0.015	0.121	0.492	0.583	0.288	0.166	0.821	0.354	0.319
Phone network available everywhere	0.407	0.491	0.464	0.499	0.362	0.481	0.393	0.489	0.413	0.493	0.247	0.604	0.368	0.666	0.052	0.122	0.280
B. Characteristics of the households, aggregated at village level																	
Number of observations	2,970		703		777		745		745								
Age (in years)	27.794	4.576	28.125	4.741	27.687	4.410	27.559	4.572	27.836	4.586	0.266	0.851	0.477	0.388	0.099	0.080	0.347
Completed primary education = {0, 1}	0.248	0.269	0.259	0.268	0.225	0.261	0.261	0.272	0.247	0.273	0.203	0.072	0.257	0.440	0.065	0.923	0.469
Completed secondary education = {0, 1}	0.035	0.119	0.039	0.126	0.033	0.118	0.036	0.118	0.033	0.116	0.912	0.924	0.776	0.712	0.669	0.755	0.470
Number of children under 5	1.178	0.626	1.163	0.637	1.194	0.630	1.196	0.621	1.156	0.618	0.753	0.937	0.413	0.415	0.479	0.473	0.880
Presence of a pregnant women	0.222	0.256	0.222	0.259	0.223	0.252	0.224	0.255	0.221	0.258	0.995	0.896	0.897	0.790	0.957	0.860	0.950
Presence of a woman in reproductive age	0.976	0.091	0.970	0.101	0.981	0.085	0.978	0.087	0.976	0.091	0.193	0.651	0.382	0.678	0.034	0.095	0.199
Main occupation is farming	0.601	0.373	0.582	0.390	0.613	0.364	0.606	0.369	0.603	0.369	0.612	0.526	0.707	0.766	0.600	0.209	0.318
Wealth score (0 to 8)	1.103	0.872	1.199	1.021	1.044	0.822	1.117	0.843	1.062	0.790	0.111	0.785	0.581	0.324	0.122	0.141	0.015
<30 min walking from health worker = {0, 1}	0.867	0.274	0.864	0.271	0.860	0.283	0.885	0.254	0.858	0.287	0.333	0.092	0.827	0.152	0.654	0.182	0.836
C. Services provided by local health facilities																	
Number of observations	372		93		93		93		93								
Number of pregnant women services per month	47.710	45.799	43.762	42.984	51.218	57.169	48.431	37.101	47.342	44.104	0.769	0.669	0.539	0.811	0.311	0.467	0.637
Number of institutional births per month	13.444	8.266	12.576	6.087	13.674	7.814	13.384	6.845	14.130	11.333	0.509	0.749	0.740	0.565	0.229	0.432	0.206
Number of fully immunized children per month	11.407	10.749	10.878	10.101	11.922	12.579	10.668	7.060	12.140	12.391	0.684	0.358	0.983	0.319	0.520	0.817	0.488
Number of fever cases treated per month	116.098	72.400	100.657	39.981	126.974	97.822	123.034	82.274	113.369	52.445	0.019	0.758	0.237	0.343	0.019	0.020	0.083
Number of malaria cases treated per month	45.888	32.029	42.289	26.879	51.443	38.876	46.632	31.517	43.233	29.378	0.286	0.345	0.102	0.458	0.076	0.368	0.860
Number of diarrhea cases treated per month	20.446	17.026	19.622	13.250	21.813	19.473	20.583	21.576	19.760	12.022	0.809	0.674	0.383	0.746	0.379	0.727	0.973

Notes: Each row states the sample mean and standard deviation of a variable, as well as the p-value from the F-test of joint significance of the treatment dummies and the p-values from pairwise treatment comparisons. These are calculated from a regression, where the variable is regressed on an indicator for worker, supervisor and shared incentives treatment. All regressions include stratification variables. Standard errors clustered at the PHU level.

Table A.3: Visits (Aggregated)

	(1)	(2)	(3)
Dep. Var.	Total number of visits provided by the health worker to sampled households per month	Total number of visits provided by the health worker in the entire village per month	Total number of visits provided by all health workers in the PHU per month
Worker incentives	0.982***	17.725***	172.036***
	(0.245)	(4.429)	(39.934)
Supervisor incentives	0.998***	18.007***	155.210***
	(0.229)	(4.139)	(38.812)
Shared incentives	1.414***	25.505***	215.576***
	(0.214)	(3.862)	(35.290)
Unit	Worker	Worker	Supervisor
Observations	2,926	2,926	372
Mean dep. var.	3.154	56.906	447.602
Mean dep. var. in Control	2.275	41.040	302.727
p-value Worker = Supervisor	0.957	0.957	0.733
p-value Supervisor = Shared	0.116	0.116	0.184
p-value Worker = Shared	0.116	0.116	0.349

Notes: Column 2 divides the total number of household visits provided by Health Worker to sampled households per month (column 1) by the share of households in the village sampled. All regressions include stratification variables. Standard errors clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1

Table A.4: Visit Type

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	% households who received [visit type] from the health worker in the past 6 months							
	Natal-related				Disease-related			
Visit Type	Pregancy visit	Accompanied woman for birth to the health facility	Pre and post-natal visit	Index (cols. 1-3)	Routine visit	Treatment/ referrals of under-5 children	Follow-up visit of under-5 children	Index (cols. 5-7)
Worker incentives	0.037**	-0.005	0.027	0.069	0.068**	0.053**	0.042*	0.155***
	(0.016)	(0.007)	(0.017)	(0.049)	(0.033)	(0.025)	(0.022)	(0.060)
Supervisor incentives	0.027*	0.004	0.037*	0.092*	0.089***	0.071**	0.031	0.178***
	(0.016)	(0.008)	(0.019)	(0.051)	(0.030)	(0.028)	(0.024)	(0.061)
Shared incentives	0.064***	0.008	0.051***	0.168***	0.151***	0.111***	0.079***	0.324***
	(0.017)	(0.008)	(0.017)	(0.051)	(0.029)	(0.024)	(0.020)	(0.056)
Unit	Worker	Worker	Worker	Worker	Worker	Worker	Worker	Worker
Observations	2,926	2,926	2,926	2,926	2,926	2,926	2,926	2,926
Mean dep. var.	0.179	0.041	0.132	0.000	0.517	0.504	0.201	0.000
Mean dep. var. in Control	0.145	0.038	0.103	-0.087	0.437	0.443	0.162	-0.171
p-value Worker = Supervisor	0.543	0.193	0.604	0.650	0.540	0.527	0.669	0.733
p-value Supervisor = Shared	0.038	0.665	0.452	0.157	0.047	0.143	0.039	0.023
p-value Worker = Shared	0.132	0.068	0.151	0.056	0.015	0.020	0.092	0.008

Notes: All regressions include stratification variables. Standard errors clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1. The index in col. 4 [resp., col. 8] estimates an equally weighted average of the z-scores of variables in cols. 1-3 [resp., cols. 5-7].

Table A.5: Visits and Household Targeting

	(1)	(2)	(3)	(4)	(5)
Dep. Var.	Number of visits				
Definition of covariate X:	Household's wealth score	Household is localed <30 min walking from health worker = {0, 1}	Household's respondent is a family member of the health worker = {0, 1}	Household's respondent is a friend of the health worker = {0, 1}	Household received no joint visit (health worker was never accompanied by supervisor to a visit)
Worker incentives	2.118***	1.078	2.113***	2.080***	1.658*
	(0.555)	(1.036)	(0.568)	(0.587)	(0.912)
Supervisor incentives	2.129***	1.115	1.889***	1.932***	0.671
	(0.509)	(1.064)	(0.505)	(0.505)	(0.807)
Shared incentives	3.385***	4.302***	3.515***	3.461***	2.219**
	(0.492)	(1.298)	(0.540)	(0.523)	(0.858)
X	0.046	0.590	2.583***	0.261	-3.300***
	(0.095)	(0.668)	(0.595)	(0.581)	(0.460)
Worker incentives * X	-0.208	1.223	-0.353	0.244	0.415
	(0.158)	(1.114)	(0.828)	(1.010)	(0.821)
Supervisor incentives * X	0.042	1.072	0.474	1.629	1.606*
	(0.134)	(1.132)	(0.772)	(1.073)	(0.867)
Shared incentives * X	-0.000	-1.165	-0.836	-1.066	1.158
	(0.138)	(1.270)	(0.853)	(0.977)	(0.832)
Unit	Household	Household	Household	Household	Household
Observations	8,559	8,159	8,601	8,601	8,606
Mean Dep. Var.	7.314	7.314	7.314	7.314	7.314
Mean Dep. Var. in Control	5.360	5.360	5.360	5.360	5.360
Mean X	0.000	0.871	0.308	0.112	0.798
p-value Worker*X = Supervisor*X	0.151	0.906	0.275	0.261	0.235
p-value Supervisor*X = Shared*X	0.783	0.115	0.096	0.028	0.655
p-value Worker*X = Shared*X	0.232	0.093	0.566	0.246	0.452

Notes: All regressions include stratification variables. Standard errors clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1

Table A.6: Services Provided by the Local Health Facility

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Pre- and post-natal care at the health facility in the past month				Disease treatments at the health facility in the past month			
Dep. Var.	Index (cols. 2-4)	Number of pregnant women services	Number of institutional births	Number of fully immunized children	Index (cols. 6-8)	Number of fever cases treated	Number of malaria cases treated	Number of diarrhea cases treated
Worker incentives	0.108	5.001	0.895	2.302	0.175*	22.269**	9.857*	4.309**
	(0.094)	(5.511)	(1.020)	(1.553)	(0.099)	(9.982)	(5.839)	(2.102)
Supervisor incentives	0.097	4.799	1.731	0.852	0.187**	25.311***	10.104*	2.420
	(0.083)	(4.220)	(1.090)	(1.352)	(0.094)	(9.108)	(5.767)	(2.231)
Shared incentives	0.244*	13.918*	2.552*	3.042*	0.223*	24.337**	9.455	6.309*
	(0.126)	(7.283)	(1.389)	(1.625)	(0.132)	(11.949)	(6.621)	(3.269)
Unit	Supervisor	Supervisor	Supervisor	Supervisor	Supervisor	Supervisor	Supervisor	Supervisor
Observations	371	371	371	371	371	371	371	371
Mean dep. var.	0.000	41.89	13.78	12.41	0.000	132	57.46	18.94
Mean dep. var. in Control	-0.110	36.06	12.51	10.89	-0.155	113.3	49.59	15.58
p-value Worker = Supervisor	0.916	0.971	0.483	0.386	0.908	0.776	0.967	0.494
p-value Supervisor = Shared	0.269	0.209	0.584	0.203	0.800	0.937	0.923	0.291
p-value Worker = Shared	0.323	0.263	0.246	0.690	0.733	0.873	0.953	0.569

Notes: All regressions include stratification variables. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The index in col. 1 [resp., col. 5] estimates an equally weighted average of the z-scores of variables in cols. 2-4 [resp., cols. 6-8]. The number of observations is 371 instead of 372 because of one missing variable.

Table A.7: Inactivity

	(1)	(2)
Dep. Var.	Health worker is inactive = {0, 1} (self-reported dropping out or no household visits in the past 6 months)	Supervisor is inactive = {0, 1} (no health worker visited in the past 6 months and no general training in last month)
Worker incentives	-0.056**	0.000
	(0.023)	(0.001)
Supervisor incentives	-0.071***	0.000
	(0.021)	(0.001)
Shared incentives	-0.068***	0.011
	(0.021)	(0.011)
Unit	Worker	Supervisor
Observations	2,970	372
Mean dep. var.	0.133	0.003
Mean dep. var. in Control	0.185	0.000
p-value Worker = Supervisor	0.443	0.965
p-value Supervisor = Shared	0.860	0.315
p-value Worker = Shared	0.549	0.316

Notes: All regressions include stratification variables. All regressions include stratification variables. Standard errors clustered at the PHU level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.8: Correlations with Worker Effort Observability

	(1)	(2)
Pairwise correlations	Correlation between actual and perceived ranking is in bottom decile = {0,1}	Correlation between actual and perceived ranking
Supervisor is a male = {0, 1}	0.066	-0.013
Supervisor age (in years)	-0.061	0.044
Supervisor completed primary education = {0, 1}	-0.008	0.086
Supervisor completed secondary education or above = {0, 1}	0.016	-0.098*
Supervisor wealth score (0 to 8)	-0.067	0.032
Number of hours worked as supervisor per week	-0.053	0.045
Number of years of experience of the supervisor	0.008	0.211**
Average number of years of experience of the health workers	0.068	-0.050
Number of health workers supervisor is responsible for	-0.245***	0.086
% health workers who are friends or family of the supervisor	-0.002	-0.075
Average distance between worker and supervisor	0.092*	-0.020
% health workers who are more than 1 km away	0.133**	-0.038

Notes: This table presents pairwise correlations. *** p<0.01, ** p<0.05, * p<0.1

Table A.9: side payments, Supervisor Effort and Visit by Social Distance, Span of Control and Performance Ranking

	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	Side-payments	Supervisor effort	Output	Side-payments	Supervisor effort	Output	Side-payments	Supervisor effort	Output	Side-payments	Supervisor effort	Output	Side-payments	Supervisor effort	Output
Dep. Var.	Supervisor shared incentive with health worker = {0, 1}	% household visits in which health worker was accompanied by supervisor	Number of visits	Supervisor shared incentive with health worker = {0, 1}	% household visits in which health worker was accompanied by supervisor	Number of visits	Supervisor shared incentive with health worker = {0, 1}	% household visits in which health worker was accompanied by supervisor	Number of visits	Supervisor shared incentive with health worker = {0, 1}	% household visits in which health worker was accompanied by supervisor	Number of visits	Supervisor shared incentive with health worker = {0, 1}	% household visits in which health worker was accompanied by supervisor	Number of visits
Definition of the covariate X:	Supervisor and health worker were friends or family at baseline = {0,1}			Number of years supervisor and health worker had known each other at baseline (demeaned)			Span of control: Number of health workers per supervisor at baseline (demeaned)			Health worker performance ranking reported by the supervisor at baseline (demeaned)			Health worker is ranked in the bottom half by the supervisor at baseline = {0, 1}		
Worker incentives	0.007	0.064**	2.559***	0.003	0.034	2.104***	0.006	0.025	2.099***	0.005	0.030	2.156***	0.002	0.020	1.939***
	(0.019)	(0.032)	(0.788)	(0.017)	(0.022)	(0.563)	(0.016)	(0.021)	(0.552)	(0.016)	(0.022)	(0.567)	(0.016)	(0.026)	(0.626)
Supervisor incentives	0.193***	0.086***	2.013***	0.180***	0.060***	2.192***	0.197***	0.044**	2.195***	0.180***	0.041*	2.016***	0.173***	0.033	1.589**
	(0.054)	(0.030)	(0.646)	(0.047)	(0.023)	(0.501)	(0.051)	(0.022)	(0.509)	(0.048)	(0.023)	(0.507)	(0.048)	(0.027)	(0.722)
Shared incentives	0.095**	0.076**	3.592***	0.099**	0.064***	3.361***	0.099***	0.060***	3.363***	0.103***	0.062***	3.541***	0.110**	0.044*	3.301***
	(0.043)	(0.032)	(0.752)	(0.039)	(0.021)	(0.493)	(0.038)	(0.021)	(0.484)	(0.040)	(0.021)	(0.511)	(0.043)	(0.026)	(0.668)
X	-0.004	0.065***	0.244	-0.002*	0.003**	0.021	-0.002	-0.009***	-0.034	-0.002	-0.002	-0.093	-0.004*	-0.009	-0.515
	(0.010)	(0.024)	(0.519)	(0.001)	(0.002)	(0.029)	(0.002)	(0.003)	(0.054)	(0.002)	(0.003)	(0.083)	(0.002)	(0.019)	(0.481)
Worker incentives * X	-0.003	-0.054	-0.834	0.001	-0.001	-0.016	0.001	0.012**	-0.015	-0.001	0.002	0.110	0.004	0.018	0.427
	(0.013)	(0.035)	(0.823)	(0.001)	(0.002)	(0.049)	(0.004)	(0.006)	(0.109)	(0.002)	(0.005)	(0.123)	(0.003)	(0.027)	(0.667)
Supervisor incentives * X	-0.023	-0.045	0.309	0.000	-0.002	0.027	0.014	0.000	0.088	0.010	-0.003	0.186	0.009	0.016	0.766
	(0.032)	(0.036)	(0.752)	(0.002)	(0.002)	(0.042)	(0.010)	(0.004)	(0.102)	(0.008)	(0.005)	(0.128)	(0.009)	(0.029)	(0.804)
Shared incentives * X	0.013	-0.024	-0.428	0.004	-0.003	-0.010	0.008	0.007	-0.039	0.010	0.004	0.029	-0.013	0.035	0.451
	(0.041)	(0.035)	(0.879)	(0.003)	(0.002)	(0.049)	(0.009)	(0.004)	(0.118)	(0.009)	(0.004)	(0.113)	(0.009)	(0.030)	(0.732)
Unit	Worker	Worker	Worker	Worker	Worker	Worker	Worker	Worker	Worker	Worker	Worker	Worker	Worker	Worker	Worker
Observations	2,904	2,908	2,915	2,909	2,913	2,920	2,915	2,919	2,926	2,685	2,689	2,696	2,685	2,689	2,696
Mean Dep. Var.	0.084	0.204	7.296	0.084	0.204	7.296	0.084	0.204	7.296	0.084	0.204	7.296	0.084	0.204	7.296
Mean Dep. Var. in Control	0.011	0.164	5.334	0.011	0.164	5.334	0.011	0.164	5.334	0.011	0.164	5.334	0.011	0.164	5.334
Mean X	0.566	0.566	0.566	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.526	0.526	0.526
p-value Worker*X=Supervisor*X	0.807	0.121	0.312	0.386	0.539	0.742	0.763	0.044	0.892	0.784	0.625	0.371	0.214	0.505	0.523
p-value Supervisor*X=Shared*X	0.473	0.210	0.681	0.830	0.318	0.518	0.151	0.977	0.390	0.204	0.625	0.145	0.281	0.576	0.341
p-value Worker*X=Shared*X	0.757	0.490	0.627	0.135	0.178	0.844	0.334	0.142	0.738	0.273	0.376	0.798	0.161	0.246	0.538
p-value Worker+Worker*X=0	0.828	0.701	0.005	0.806	0.135	<0.001	0.662	0.106	<0.001	0.792	0.162	<0.001	0.726	0.142	0.001
p-value Supervisor+Supervisor*X=0	<0.001	0.158	<0.001	<0.001	0.013	<0.001	<0.001	0.040	<0.001	<0.001	0.117	<0.001	<0.001	0.080	<0.001
p-value Worker+Worker*X=0	0.016	0.028	<0.001	0.009	0.004	<0.001	0.010	0.002	<0.001	0.013	0.003	<0.001	0.011	0.003	<0.001

Notes: All regressions include stratification variables. Standard errors clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1

Table A.10: Parameter Estimates

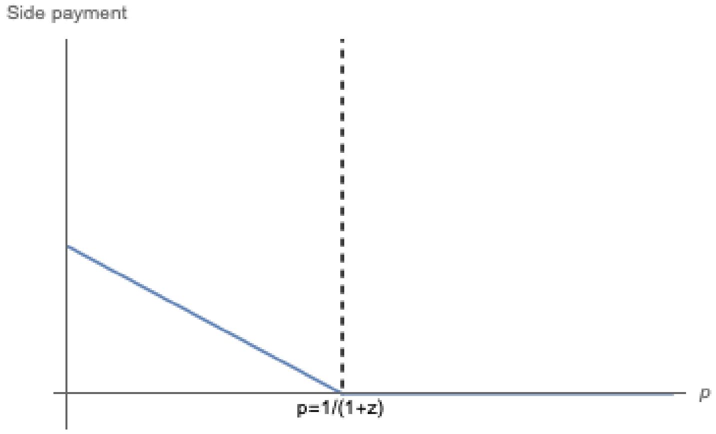
	(1)
Complementarity γ	3.1
Contractual friction z	3.0
CHW unit cost of effort c_1	23.0
PS unit cost of effort c_2	1610.9
CHW baseline incentive b_1	57.8
PS baseline incentive b_2	19.0
α	2.0
Δ in marginal product of CHW effort (shared incentive)	36 %
Δ in marginal product of CHW effort (no incentive)	17 %
Total CHW cost of effort (no incentive)	69.0
Total PS cost of effort (no incentive)	178.9

Table A.11: Moment Fit

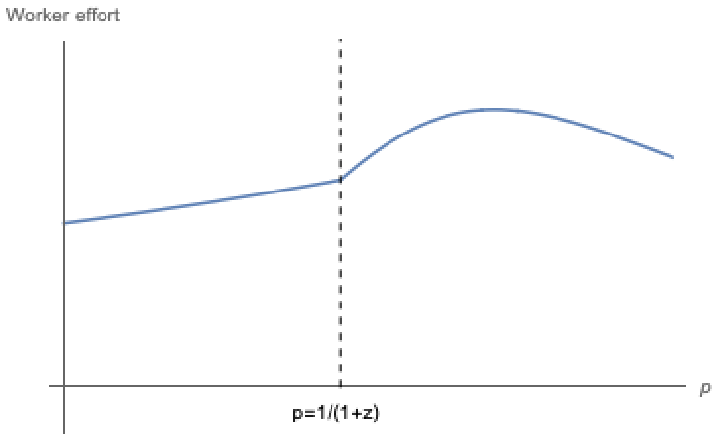
Moments	Targeted Real	Simulated
PS effort in CHW incentive group	0.198	0.159
PS effort in PS incentive group	0.225	0.278
PS effort in shared incentive group	0.228	0.230
PS effort in control group	0.164	0.111
Output in CHW incentive group	10.551	10.908
Output in PS incentive group	10.413	10.678
Output in shared incentive group	11.827	11.255
Output in control group	7.256	7.174
Side payment in CHW incentive group (1,000 SSL)	0.088	0.00
Side payment in PS incentive group (1,000 SSL)	0.715	0.000
Side payment in shared incentive group (1,000 SSL)	0.428	0.000
Value loss function	64.9	

Figure A.1: Side Payment and Efforts as a Function of the Share of the Incentive Offered to the Worker ($z > 1$, $\gamma > 0$)

(a) Side payment



(b) Worker effort



(c) Supervisor effort

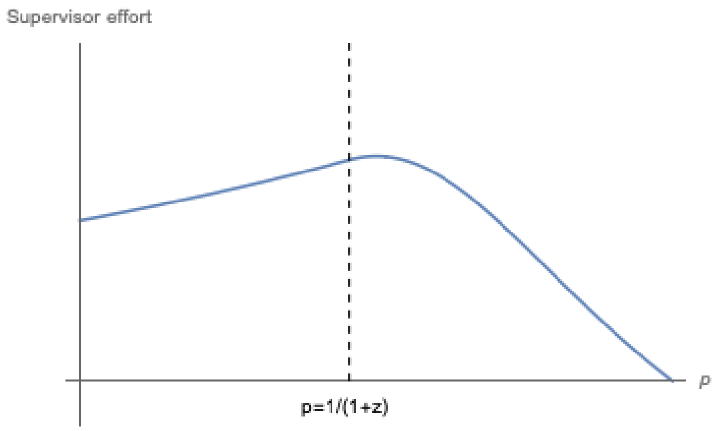
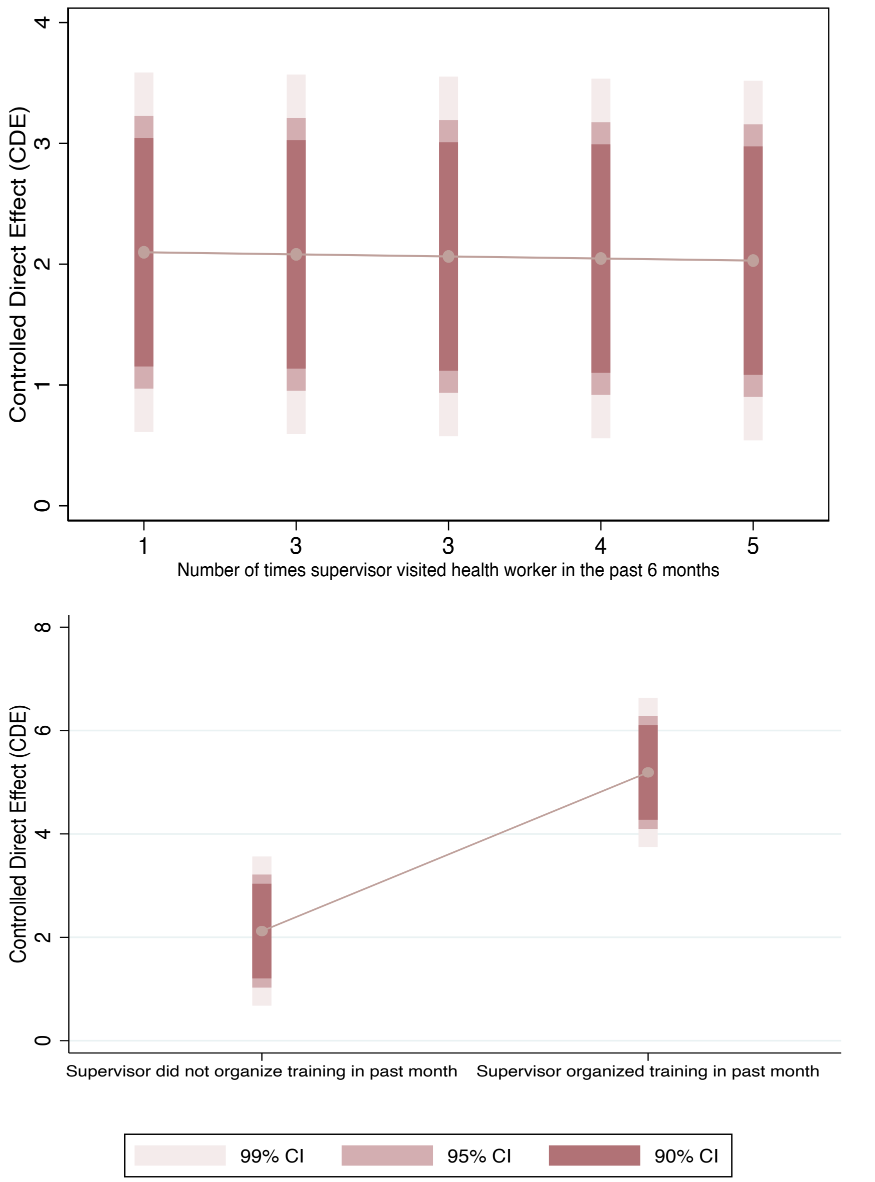


Figure A.2: Evolution of SMS Reporting

Notes: The figure plots the difference in the number of SMS reports between each treatment group and the control group. The coefficients are estimated from a regression of the number of SMS reports in each single month on the treatment dummies, controlling for the stratification variables and with standard errors clustered at the PHU level. Bars are 95% confidence intervals.

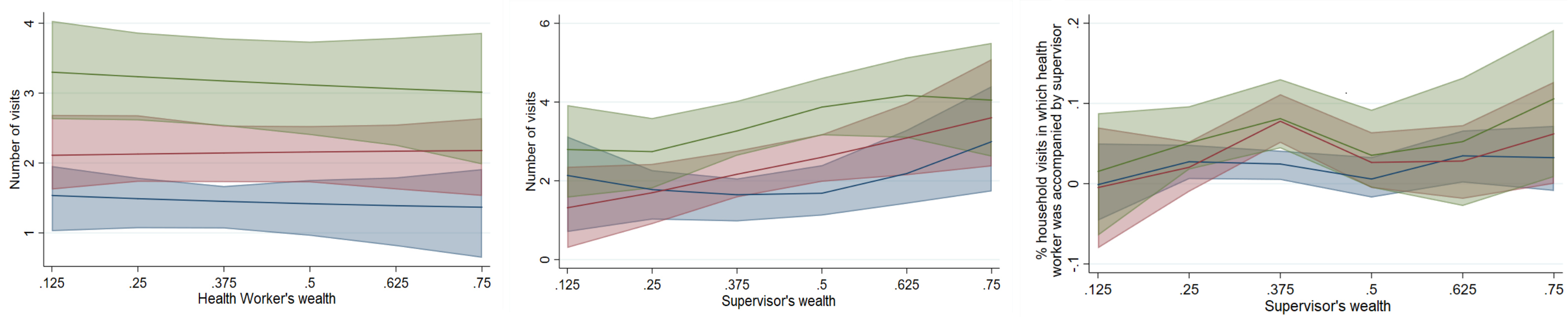
Figure A.3: Mediation Analysis (Continued)



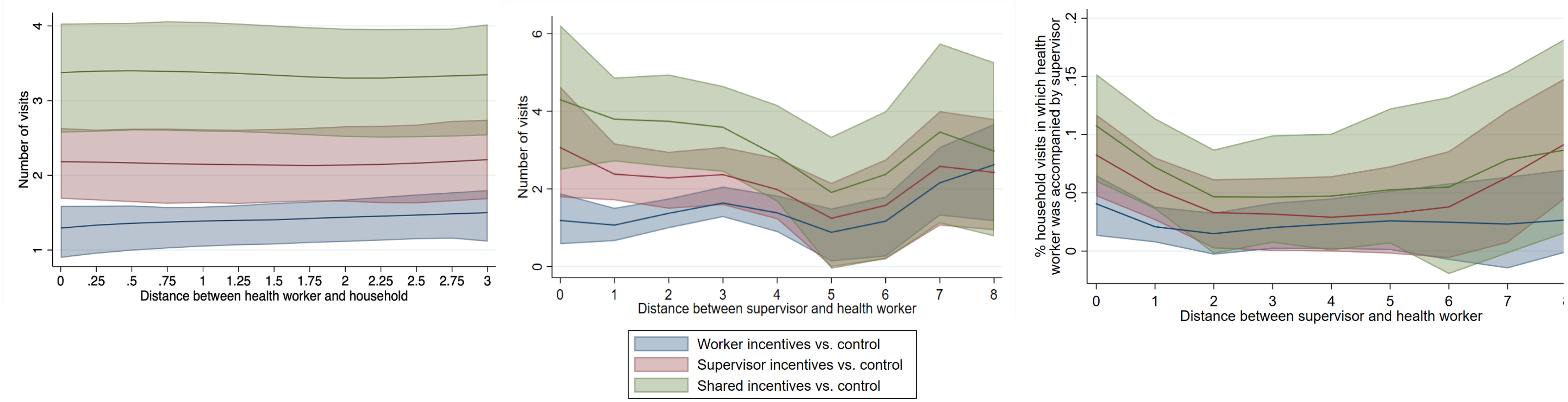
Notes: This figure plots the controlled direct effect (CDE) of the worker incentives treatment on the number of visits provided by a health worker for different values of supervisor's effort, measured differently in the two figures.

Figure A.4: Visits and Supervisor Effort by Wealth and Distance

A. Treatment Effects on Visits and Supervisor Effort by Wealth

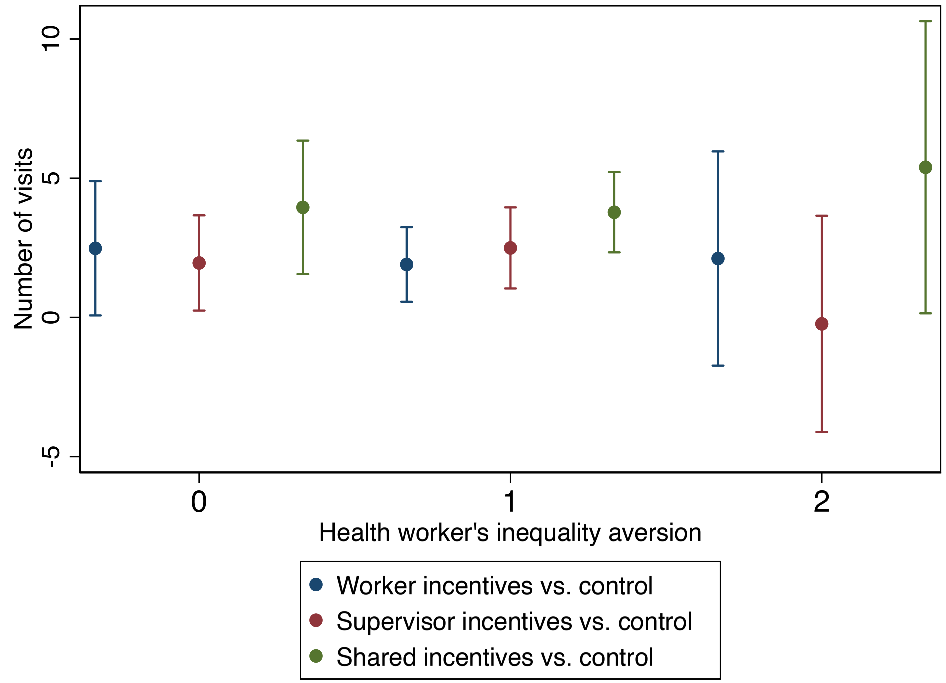


B. Treatment Effects on Visits and Supervisor Effort by Distance



Notes: The figure plots non-parametric estimates of the treatment effects on the number of visits (left and middle figures) and on the supervisor effort index (right figures) by wealth score and by distance. Standard errors are bootstrapped for each value of the x-axis, with 100 repetitions and the re-sampling is with replacement.

Figure A.5: Visits by Inequality Aversion



Notes: The figure plots the difference in the number of visits provided by the health worker between each treatment group and the control group, for workers with low, medium and high inequality aversion. Inequality aversion is measured by asking each health worker the following hypothetical questions: "There is a local farm that hires workers to help with the potato harvest. Sheka accepts a contract to work at the farm for 20,000 SLL per day. He arrives at work the next morning. The farm is very big and there is one supervisor for the 20 workers helping with the harvest. He learns that his supervisor gets paid [amount] SLL per day. Do you think Sheka will show up to work the next day?," and amount = {20,000; 30,000; 120,000}. Our measure of inequality aversion takes value 0 if the worker answers that Sheka would always show up to work, regardless of the amount; value 1 if worker reports that Sheka would not show up only if amount=120,000 and would show up otherwise; value 2 if the worker reports that Sheka would not show up if amount > 30,000. The coefficients are estimated from a regression of the number of visits on the treatment dummies interacted with inequality aversion, controlling for stratification variables and clustering standard errors clustered at the PHU level. Bars are 95% confidence intervals.

B Promotion Incentives Appendix

When a supervisor's position becomes available, one of the health workers in that PHU is promoted to take over the position. A random sample of 2,081 health workers out of the 2,970 health workers in this study were part of a separate evaluation that involved a change in the promotion system. More specifically, six month after the start of the piece-rate incentives, the promotion system became meritocratic in a random half of the 372 PHUs while the rest of the PHUs kept the status-quo system (in which the promotion decision is at the discretion of the PHU in-charge). Variation in the promotion system was cross-randomized with PHU-level variation in whether information about supervisors' fixed wage was revealed or not to the health workers. See Deserranno, Kastrau, and León-Ciliotta (2021) for more details.

Table A.12 shows that our main treatment effects on visits are orthogonal to the random variation in the promotion system and orthogonal to providing information about the supervisor's fixed wage. This is not surprising as the short-run incentives analyzed in this paper are paid by an external organization and have no role in the government promotion decision, nor do they influence the supervisor fixed wage. Table A.13 moreover shows that the effects of our incentives treatment persist if we restrict the analysis to the sub-sample of health workers that did not take part in this separate study.

Table A.12: Incentives and Household Visits, by Promotions

Dep. Var.	(1)	(2)	(3)	(4)	(5)	(6)
	Household visits provided by the health worker in the past 6 months					% households who trust the health worker as a health provider
	Number of visits	% households visited	Number of visit types	Average visit length	Number of health topics discussed	
Worker incentives	1.635	0.094*	0.305	1.221	0.006	0.005
	(1.125)	(0.048)	(0.189)	(2.128)	(0.203)	(0.045)
Supervisor incentives	1.664*	0.063	0.414*	2.116	0.386	0.064
	(0.992)	(0.051)	(0.237)	(2.157)	(0.312)	(0.045)
Shared incentives	3.335***	0.139***	0.611***	4.432**	0.521**	0.125***
	(1.186)	(0.047)	(0.190)	(2.041)	(0.238)	(0.044)
Meritocratic promotions	0.651	0.072*	0.264	2.369	0.224	0.070*
	(0.766)	(0.042)	(0.163)	(1.730)	(0.190)	(0.039)
Pay progression	-0.895	0.004	0.011	-1.980	0.026	0.020
	(0.844)	(0.048)	(0.182)	(1.905)	(0.265)	(0.043)
Meritocratic promotions + Info about supv. pay	0.272	-0.031	0.065	-0.914	0.080	-0.017
	(0.848)	(0.044)	(0.163)	(1.555)	(0.203)	(0.048)
Worker incentives * Meritocratic promotions	-0.784	-0.140**	-0.485*	-3.099	-0.216	-0.020
	(1.700)	(0.068)	(0.263)	(2.765)	(0.309)	(0.061)
Supervisor incentives * Meritocratic promotions	2.352	0.037	0.128	0.271	-0.194	-0.084
	(1.429)	(0.066)	(0.307)	(2.761)	(0.393)	(0.062)
Shared incentives * Meritocratic promotions	0.064	-0.068	-0.172	-2.104	-0.114	-0.158**
	(1.533)	(0.064)	(0.270)	(2.672)	(0.389)	(0.065)
Worker incentives * Info about supv. pay	0.491	-0.010	-0.033	3.265	0.322	0.045
	(1.427)	(0.073)	(0.263)	(2.829)	(0.356)	(0.065)
Supervisor incentives * Info about supv. Pay	-0.046	-0.018	-0.261	-1.068	-0.315	-0.068
	(1.248)	(0.071)	(0.293)	(2.744)	(0.412)	(0.067)
Shared incentives * Info about supv. pay	0.217	-0.045	-0.121	-0.481	-0.200	-0.082
	(1.376)	(0.067)	(0.259)	(2.795)	(0.356)	(0.062)
Worker incentives * Merit. + Info about supv. Pay	2.157	0.059	0.292	2.954	0.521	0.102
	(1.569)	(0.065)	(0.251)	(2.657)	(0.316)	(0.065)
Supervisor incentives * Merit. + Info about supv. Pay	-0.416	0.057	-0.233	-0.011	-0.354	0.017
	(1.303)	(0.070)	(0.279)	(2.559)	(0.372)	(0.067)
Shared incentives * Merit. + Info about supv. Pay	-0.290	0.058	0.080	1.039	0.289	0.016
	(1.510)	(0.064)	(0.253)	(2.475)	(0.337)	(0.064)
Unit	Worker	Worker	Worker	Worker	Worker	Worker
Observations	2,926	2,926	2,926	2,926	2,926	2,926
Mean dep. var.	7.296	0.709	1.745	14.39	2.248	0.745
Mean dep. var. in Control	5.334	0.637	1.448	12.32	2.015	0.707

Notes: All regressions include stratification variables. Standard errors clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1

Table A.13: Incentives and Household Visits, Restricted Sample of Health Workers

	(1)	(2)	(3)	(4)	(5)	(6)
	Household visits provided by the health worker in the past 6 months					% households who trust the health worker as a health provider
Dep. Var.	Number of visits	% households visited	Number of visit types	Average visit length	Number of health topics discussed	
Worker incentives	2.357***	0.119***	0.345***	3.915***	0.330**	0.051
	(0.815)	(0.040)	(0.132)	(1.359)	(0.165)	(0.033)
Supervisor incentives	2.350***	0.138***	0.452***	1.988	0.362**	0.048
	(0.686)	(0.038)	(0.129)	(1.233)	(0.171)	(0.038)
Shared incentives	3.122***	0.154***	0.636***	4.267***	0.603***	0.060*
	(0.655)	(0.037)	(0.134)	(1.251)	(0.165)	(0.035)
Unit	Worker	Worker	Worker	Worker	Worker	Worker
Observations	960	960	960	960	960	960
Mean dep. var.	6.753	0.682	1.660	13.702	2.168	0.730
Mean dep. var. in Control	4.777	0.571	1.268	11.060	1.811	0.680
p-value Worker = Supervisor	0.993	0.600	0.425	0.147	0.868	0.928
p-value Supervisor = Shared	0.258	0.661	0.178	0.062	0.209	0.746
p-value Worker = Shared	0.344	0.330	0.036	0.794	0.146	0.780

Notes: Sample restricted to health workers who did experience any change in the promotion system. All regressions include stratification variables. Standard errors clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1

C Model Appendix

C.1 The model: set up

We want to model the interaction between a supervisor (player 2) and a worker (player 1), in a situation in which the supervisor can pay an incentive (s) to the worker based on his effort, there are potential complementarities between the worker and the supervisor and the incentive payment of the supervisor to the worker might be subject to frictions or transaction costs (z). We will also assume that a principal can pay an incentive, based on the common output, to the worker and/or to the supervisor.

The objective of the principal will be to maximize the output, y. The number of output units produced will be given by: $e_1 + \gamma e_1 e_2$, that depends on the effort of players 1 and 2 (e_1, e_2 correspondingly) and on the level of complementarity between them (γ). Effort is costly to both the worker and the supervisor, and we assume the cost of effort is quadratic: ce_i^2.

For each unit of output, the principal will pay p to the worker and $1 - p$ to the supervisor, where $p \in [0, 1]$.

The payoff of the worker will look as follows:

$$\pi_1 = (e_1 + \gamma e_1 e_2)(s + p) - ce_1^2$$

And the payoff of the supervisor:

$$\pi_2 = (e_1 + \gamma e_1 e_2)(1 - p - sz) - ce_2^2$$

where we model the contracting frictions for the supervisor with $z > 1$, whenever $z = 1$ this friction is shut down. We will see later that a convenient assumption would be to set an upper bound of 2 for z.

We will assume that, in period 0, the principal chooses p that maximizes the output y. After this, the supervisor moves first and chooses both the incentives he wants to pay to the worker (s) and his effort (e_2) at a time. Then, the worker observes s and e_2 and decides how much effort to exert. We assume that s is paid proportionally to y. Also, we assume that transfers can only go from the supervisor to the worker $s \geq 0$.

We will make the following assumption:

Assumption 1: $8c^2 < \gamma^2$; $c, \gamma \in \mathbb{R}^+$

Claim: If assumption 1 ($8c^2 < \gamma^2$) holds; then, it is also true that:

a) $2c^2 - \gamma^2 p(1 - p) > 0$ when $p \in \left[\frac{1}{1+z}, 1\right]$

b) $8zc^2 - \gamma^2(1 + p(z - 1))^2 > 0$ when $p \in \left[0, \frac{1}{1+z}\right)$

Proof:

I will divide the proof in two parts. First, I will show that assumption 1 implies a). Then, I will show that it also implies b).

Part 1: Consider the following maximization problem:

$$\max_{p\in[0,1]} p(1-p)$$

The solution is $p=\frac{1}{2}$, such that, at its maximum, the objective function attains the value of $\frac{1}{4}$. By the definition of maximum, we have that:

$$\frac{\gamma^2}{4} \geq \gamma^2 p(1-p) \ \ \forall p \in [0,1]$$

By our assumption 1, we have that: $2c^2 > \frac{\gamma^2}{4}$. Thus, by the above and the transitivity of the inequality this also implies that $2c^2 > \gamma^2 p(1-p)$ (what we wanted to show).

Part 2: To start with, we W.T.S. that the following inequality holds for $p \in \left[0, \frac{1}{1+z}\right)$ when assumption 1 holds:

$$8zc^2 > 8c^2(1+p(z-1))^2 \iff z-1 > 2p(z-1)+p^2(z-1)^2 \iff p^2(z-1)+2p-1<0$$

To see whether $p^2(z-1)+2p-1<0$ holds, we examine the shape of this quadratic function. First, its roots with respect to p are the following:

$$\bar{p} = \frac{\sqrt{z}-1}{z-1} > 0; \underline{p} = \frac{-\sqrt{z}-1}{z-1} < 0 \ \ \forall z \geq 1$$

Now, to understand where the parabola takes negative values we plug in a point $p \in (\underline{p}, \bar{p})$, for instance, $p=0$. At this point, the function takes a negative value -1. This means that $p^2(z-1)+2p-1<0 \ \ \forall p \in (\underline{p}, \bar{p})$.

Next, we W.T.S. that $p^2(z-1)+2p-1<0 \ \ \forall p \in (\underline{p}, \bar{p}) \implies p^2(z-1)+2p-1 < 0 \ \ \forall p \in \left[0, \frac{1}{1+z}\right)$. This is equivalent to showing that $\underline{p}<0$ and $\bar{p} \geq \frac{1}{1+z}$:

Lower bound: as we already shown $\underline{p} < 0$.

Upper bound: we W.T.S. that $\bar{p} \geq \frac{1}{1+z}$, this is equivalent to showing:

$$(\sqrt{z}-1)(z+1) \geq z-1 \iff z(\sqrt{z}-2)+\sqrt{z} \geq 0$$

Consider now the minimum value of $z(\sqrt{z}-2)+\sqrt{z}$ with respect to z, as it is attained at the minimum possible value of z ($z=1$), it would be equal to $1(1-2)+1=0$. This means that $\forall z \geq 1 \ z(\sqrt{z}-2)+\sqrt{z} \geq 0$.

Finally, by assumption 1 and transitivity of the inequalities:

$$8c^2(1+p(z-1))^2 > \gamma^2(1+p(z-1))^2 \implies 8c^2z > \gamma^2(1+p(z-1))^2$$

What we wanted to show.

C.2 The model: main analysis

Let us solve the model by backward induction:

Period 2:

The maximization problem of the worker in the second period would be:

$$\max_{e_1}(e_1+\gamma e_1 e_2)(s+p)-ce_1^2$$

Thus, his optimal level of effort is given by:

$$e_1 = \frac{(s+p)(1+\gamma e_2)}{2c}$$

Period 1:

Anticipating the optimal action of player 1, the maximization problem of player 2 becomes:

$$\max_{e_2,s} \frac{(s+p)(1-p-sz)(1+\gamma e_2)^2}{2c} - ce_2^2$$

Thus, the optimal effort and incentive payment would be:

$$e_2 = \frac{\gamma(s+p)(1-p-sz)}{2c^2-\gamma^2(s+p)(1-p-sz)}$$

$$s = \begin{cases} \frac{1-p(1+z)}{2z}, & p \leq \frac{1}{1+z} \\ 0, & p > \frac{1}{1+z} \end{cases}$$

Let us first focus in the case where $p \leq \frac{1}{1+z}$. In this situation:

$$e_2 = \frac{\gamma(1+p(z-1))^2}{8zc^2-\gamma^2(1+p(z-1))^2}$$

And plugging e_2 into e_1:

$$e_1 = \frac{2c(1+p(z-1))}{8zc^2-\gamma^2(1+p(z-1))^2}$$

In this case, the output y as a function of p would be given by:

$$y = \frac{16zc^3(1+p(z-1))}{(8zc^2 - \gamma^2(1+p(z-1))^2)^2}$$

In the case in which $p > \frac{1}{1+z}$, we will assume that $s = 0$ (since we do not allow for $s < 0$). Thus, in this case:

$$e_2 = \frac{\gamma p(1-p)}{2c^2 - \gamma^2 p(1-p)}$$

$$e_1 = \frac{pc}{2c^2 - \gamma^2 p(1-p)}$$

And so the output would be:

$$y = \frac{2pc^3}{(2c^2 - \gamma^2 p(1-p))^2}$$

Period 0:

Recall that the objective of the principal is to maximize the units of output. Thus, his maximization problem would be divided in two parts: first, he maximizes the expression for y assuming that $p \leq \frac{1}{1+z}$; then, he does the maximization corresponding to $p > \frac{1}{1+z}$ and, finally, compares:

In the first maximization problem, the principal would want to solve:

$$\max_{p \leq \frac{1}{1+z}} \frac{16zc^3(1+p(z-1))}{(8zc^2 - \gamma^2(1+p(z-1))^2)^2}$$

Note that the derivative of the objective function with respect to p in this case is positive given the assumption that $8zc^2 - \gamma^2(1+p(z-1))^2 > 0$ for any p, such that the effort levels are positive for any p. To see this:

$$\frac{dy}{dp} = \frac{16zc^3(z-1)(8zc^2 + \gamma^2(1+p(z-1))^2)}{(8zc^2 - \gamma^2(1+p(z-1))^2)^3}$$

The numerator of this expression is always positive since $p, c \geq 0$ and $z \geq 1$ by assumption. And the denominator will be positive by the assumption above.

Thus, the solution to this maximization problem in this case would be to choose the maximum p possible: $p = \frac{1}{1+z}$.

The next problem would be:

$$\max_{p > \frac{1}{1+z}} \frac{2pc^3}{(2c^2 - \gamma^2 p(1-p))^2}$$

And so the optimal p in this case would be given by the solution to:

$$\frac{dy}{dp} = \frac{2c^3(2c^2 + \gamma^2 p(1-3p))}{(2c^2 - \gamma^2 p(1-p))^3} = 0$$

Assuming that $2c^2 - \gamma^2 p(1-p) > 0$, the solution for the optimal p will be given by the solution to:

$$3\gamma^2 p^2 - \gamma^2 p - 2c^2 = 0$$

The unique positive middle solution for the optimal p is then:

$$p^* = \frac{1}{6} + \frac{\sqrt{\gamma^2 + 24c^2}}{6\gamma}$$

Interestingly, in terms of comparative statics, when γ is larger, the optimal p^* will be smaller. This is easy to see from the derivative of p^* with respect to γ, which is always negative:

$$\frac{dp^*}{d\gamma} = \frac{-4c^2}{\sqrt{\gamma^2 + 24c^2}} < 0$$

Note that in order to say that p^* is the global maximum in the right-hand side problem (when $p > \frac{1}{1+z}$), we need to ensure that $\frac{d^2y}{d^2p} < 0$ (the second derivative is negative), that the objective function ($\frac{2pc^2}{(2c^2 - \gamma^2 p(1-p))^2}$) is continuous on $p \in [\frac{1}{1+z}, 1]$ and that $p = \frac{1}{6} + \frac{\sqrt{\gamma^2+24c^2}}{6\gamma} \leq 1$. This translates into:

- A negative second derivative at $p = \frac{1}{6} + \frac{\sqrt{\gamma^2+24c^2}}{6\gamma}$:

  $$\frac{d^2y}{d^2p} = \frac{2c^3\gamma^2((2c^2 - \gamma^2 p(1-p))(1-6p) - 3(2c^2 + \gamma^2 p(1-3p))(2p-1)))}{(2c^2 - \gamma^2 p(1-p))^4} < 0$$

  $$\iff (2c^2 - \gamma^2 p(1-p))(1-6p) - 3(2c^2 + \gamma^2 p(1-3p))(2p-1) < 0$$

  Note that $p = \frac{1}{6} + \frac{\sqrt{\gamma^2+24c^2}}{6\gamma} > \frac{1}{2}$. Now take the minimum of $(2c^2 - \gamma^2 p(1-p))(1-6p) - 3(2c^2 + \gamma^2 p(1-3p))(2p-1)$ with respect to $p \in [\frac{1}{2}, 1]$.

  As the first derivative of $(2c^2 - \gamma^2 p(1-p))(1-6p) - 3(2c^2 + \gamma^2 p(1-3p))(2p-1)$ is negative, its minimum is achieved at $p = \frac{1}{2}$. At this point:$(2c^2 - \gamma^2 p(1-p))(1-6p) - 3(2c^2 + \gamma^2 p(1-3p))(2p-1) = -\frac{1}{2}(8c^2 - \gamma^2) < 0$ since $8c^2 - \gamma^2 > 0$ by assumption 1.

- The objective function is continuous:

  $$2c^2 - \gamma^2 p(1-p) \neq 0 \iff p \neq \frac{1}{2} \pm \frac{\sqrt{\gamma^2 - 8c^2}}{\gamma^2}$$

A sufficient condition for this is to assume $\gamma^2 < 8c^2$ (exactly assumption 1).

- The condition $p = \frac{1}{6} + \frac{\sqrt{\gamma^2+24c^2}}{6\gamma} \leq 1$ is equivalent to $c^2 \leq \gamma^2$. This is, the complementarity between workers has to be high enough for a two-sided incentive to be optimal compared to paying all incentives to the worker.

To sum up, the possible candidates for the optimal p^* when $c^2 \leq \gamma^2$ are:

$$p^* = \frac{1}{1+z}$$

$$p^* = \frac{1}{6} + \frac{\sqrt{\gamma^2 + 24c^2}}{6\gamma}$$

Finally, the principal would need to compare the value of the output (y), under $p^*(p \leq \frac{1}{1+z})$ and $p^*(p > \frac{1}{1+z})$ and choose the highest. For simplicity we will call them $p_a^* = p^*(p \leq \frac{1}{1+z}) = \frac{1}{1+z}$ and $p_b^* = p^*(p > \frac{1}{1+z}) = \frac{1}{6} + \frac{\sqrt{\gamma^2+24c^2}}{6\gamma}$:

$$y(p_a^*) = \frac{8c^2 z(1+z)^3}{(2(1+z)^2c^2 - \gamma^2 z)^2}$$

$$y(p_b^*) = \frac{27c^3(\gamma + \sqrt{\gamma^2 + 24c^2})}{(24c^2 - \gamma(\gamma + \sqrt{\gamma^2 + 24c^2}))^2}$$

C.3 Comparative statics on the advantage of each optimal incentive candidate

Let $\mathcal{A}_{p,q}$ be the advantage of choosing the incentive that gives p to the worker and $1-p$ to the supervisor compared to choosing the incentive that pays q to the worker and $1-q$ to the supervisor. Using this tool we can compare different incentive schemes and analyze how certain parameters affect the advantage of one versus the other.

Comparing $p = p_a^*$ and $p = 1$:

$$\mathcal{A}_{p_a^*,1} = y(p_a^*) - y(1) = \frac{8c^2 z(1+z)^3}{(2(1+z)^2c^2 - \gamma^2 z)^2} - \frac{1}{2c}$$

We have that:

$$\frac{d\mathcal{A}_{p_a^*,1}}{d\gamma} = \frac{32\gamma c^2 z^2(1+z)^3}{(2(1+z)^2c^2 - \gamma^2 z)^3} > 0$$

since $2(1+z)^2c^2 - \gamma^2 z > 0$ by our previous assumption: $2c^2 - \gamma^2 p(1-p) > 0$.

In a similar fashion, comparing $p = p_b^*$ and $p = 1$:

$$\mathcal{A}_{p_b^*,1} = y(p_b^*) - y(1) = \frac{27c^3(\gamma + \sqrt{\gamma^2 + 24c^2})}{(24c^2 - \gamma(\gamma + \sqrt{\gamma^2 + 24c^2}))^2} - \frac{1}{2c}$$

$$\frac{d\mathcal{A}_{p_b^*,1}}{d\gamma} = \frac{27c^3(\gamma + \sqrt{\gamma^2 + 24c^2})(24c^2 - \gamma(\gamma + \sqrt{\gamma^2 + 24c^2}) + 2\gamma + 2\gamma^2\sqrt{\gamma^2 + 24c^2})}{(8c^2 - \gamma(\gamma + \sqrt{\gamma^2 + 24c^2}))^3\sqrt{\gamma^2 + 24c^2}} > 0$$

again using $8c^2 - \gamma(\gamma + \sqrt{\gamma^2 + 24c^2}) > 0$ by our previous assumption: $8zc^2 - \gamma^2(1 + p(z-1))^2 > 0$.

This means that the advantage of choosing the optimal $p^* \in (0,1)$ compared to $p^* = 1$ is increasing in γ: the larger γ is, the more harming it is (in terms of final output), to pay all the incentive to the worker.

Let us now try the analogous comparison between $p = p_a^*$, $p = p_b^*$ and $p = 0$.
For $p = p_a^*$ versus $p = 0$:

$$\mathcal{A}_{p_a^*,0} = y(p_a^*) - y(0) = \frac{8c^2 z(1+z)^3}{(2(1+z)^2c^2 - \gamma^2 z)^2} - \frac{16zc^3}{(8zc^2 - \gamma^2)^2}$$

We have that:

$$\frac{d\mathcal{A}_{p_a^*,0}}{d\gamma} = \frac{32\gamma c^2 z^2(1+z)^3}{(2(1+z)^2c^2 - \gamma^2 z)^3} - \frac{\gamma 64c^3 z}{(8zc^2 - \gamma^2)^3}$$

And comparing $p = p_b^*$ with $p = 0$:

$$\mathcal{A}_{p_b^*,0} = y(p_b^*) - y(0) = \frac{27c^3(\gamma + \sqrt{\gamma^2 + 24c^2})}{(24c^2 - \gamma(\gamma + \sqrt{\gamma^2 + 24c^2}))^2} - \frac{16zc^3}{(8zc^2 - \gamma^2)^2}$$

$$\frac{d\mathcal{A}_{p_b^*,0}}{d\gamma} = \frac{2c^3(\gamma + \sqrt{\gamma^2 + 24c^2})(56c^2 + \gamma^2 + 2\gamma + 3\gamma\sqrt{\gamma^2 + 24c^2})}{(8c^2 - \gamma(\gamma + \sqrt{\gamma^2 + 24c^2}))^3\sqrt{\gamma^2 + 24c^2}} - \frac{\gamma 64c^3 z}{(8zc^2 - \gamma^2)^3}$$

As one can see from the derivatives, the effect of γ on the advantage of $p = p^*$ with respect to $p = 0$ is unclear and will depend on the specific value of γ, but also on the cost of effort of the players z and the contracting cost of the supervisor c. Intuitively, when z is small it is more likely that γ has a positive effect on the advantage of $p = p^*$ with respect to $p = 0$; while a large z makes $p = 0$ more attractive and the increase in the advantage of $p = p^*$ with respect to $p = 0$ less sharp on γ.

C.4 Special Cases

$\gamma = 0$, $z = 1$:

In this case, the supervisor has no incentive to exert effort, since his effort is not leading to any rise in productivity $\gamma = 0$. Therefore, his optimal level of effort is $e_2 = 0$. And, as in the general case, he chooses to pay a positive side payment ($s > 0$) as long as his incentive is above $\frac{1}{1+z}$. In this case, this means that $p \leq \frac{1}{2}$.

On the other hand, the worker exerts:

$$e_1 = \frac{s+p}{2c}$$

Let us then analyze the maximization problem of the principal:

- If $p \leq \frac{1}{2}$ and so $s = \frac{1-2p}{2}$ the principal maximizes the output, that is equal to the effort of the worker when he is offered $s = \frac{1-2p}{2}$:

  $$\max . \frac{1}{4c}$$

  This is independent of p; it is, any $p \leq \frac{1}{2}$ would lead to the same output level y.

- If $p > \frac{1}{2}$ and $s = 0$, the principal's problem becomes:

  $$\max . \frac{p}{2c}$$

  which solution is $p^* = 1$ since the objective function is increasing in p. Note that, in this case, as $c > 0$, we have that $\gamma < c$ (unlike before).

Finally, the principal compares the two possible optimal p^*:

$$y(p^* \leq \frac{1}{2}) = \frac{1}{4c}$$

$$y(p^* = 1) = \frac{1}{2c}$$

And, as $y(p^* = 1) > y(p^* \leq \frac{1}{2})$, he chooses $p^* = 1$. This is intuitive given that the supervisor is not "useful" for production and, moreover, has the same understanding of the worker's effort as the principal (both of them observe e_1).

$\gamma = 0$, $z > 1$:

Again here, the supervisor chooses to exert no effort $e_2 = 0$ and a side payment of $s = \frac{1-p(1+z)}{2z}$, while the worker exerts $e_1 = \frac{s+p}{2c}$.

The two-step maximization problem of the principal is now:

- When $s > 0$ and $p \leq \frac{1}{1+z}$:
$$\max.\frac{1-p(1-z)}{4zc}$$
solved by $p^* = \frac{1}{1+z}$ as the objective function increases in p.

- When $s = 0$ and $p > \frac{1}{1+z}$:
$$\max.\frac{p}{2c}$$
just like in the previous case, maximized at $p^* = 1$.

Now, the principal would compare the output levels under the 2 candidate:

$$y\left(p^* = \frac{1}{1+z}\right) = \frac{1}{2c(1+z)}$$

$$y(p^* = 1) = \frac{1}{2c}$$

Also here $p^* = 1$ turns out to be the optimal incentive from the point of view of the principal, since $y(p^* = 1) > y(p^* = \frac{1}{1+z})$. Indeed, the result above would be nested in this example.

$\underline{\gamma > 0,\ z = 1}$:

Using the results above and plugging in for $z = 1$ one can obtain:

- When $p \leq \frac{1}{2}$ and so $s > 0$:
$$e_2 = \frac{\gamma}{8c^2 - \gamma^2}$$
$$e_1 = \frac{2c}{8c^2 - \gamma^2}$$
$$y = \frac{16c^3}{(8c^2 - \gamma^2)^2}$$

- While under $p > \frac{1}{2}$ and $s = 0$:
$$e_2 = \frac{\gamma p(1-p)}{2c^2 - \gamma^2 p(1-p)}$$
$$e_1 = \frac{pc}{2c^2 - \gamma^2 p(1-p)}$$
$$y = \frac{2pc^3}{(2c^2 - \gamma^2 p(1-p))^2}$$

The solution to the two-step principal's problem would be given by one of the following p^*:

- When $p \leq \frac{1}{2}$, any $p^* \in [0, \frac{1}{2}]$ would work.
- When $p > \frac{1}{2}$, $p^* = \frac{1}{6} + \frac{\sqrt{\gamma^2+24c^2}}{6\gamma}$, as long as $\gamma > c$

Finally, the actual optimum will be determined by comparing:

$$y\left(p^* = \frac{1}{6} + \frac{\sqrt{\gamma^2 + 24c^2}}{6\gamma}\right) = \frac{27c^3(\gamma + \sqrt{\gamma^2 + 24c^2})}{(24c^2 - \gamma(\gamma + \sqrt{\gamma^2 + 24c^2}))^2}$$

$$y\left(p^* \leq \frac{1}{2}\right) = \frac{16c^3}{(8c^2 - \gamma^2)^2}$$

The p^* generating the largest level of output y will be chosen and this will depend on the specific values of γ and c.

D Prediction Survey Appendix

In collaboration with the Social Science Prediction Platform,[49] we invited social scientists to forecast how our treatments affect household visits compared to the control group and the extent to which effort complementarities and contractual frictions play a role in our context as well as in a number of different public and private organizations. The participants made their forecasts before the results of this study were made public. Out of the 29 participants, 90% are economists; 41% of whom are faculty members and 45% are graduate students. The first set of questions asked participants to forecast the average number of household visits health workers conduct in T_{worker}, T_{supv}, and T_{shared} after giving them a 700-word description of the study and informing them about the average number of household visits and its standard deviation for control group workers. The average forecasts for the number of household visits by survey participants are 7.73 in T_{worker} (compared to 7.42 we find in the data), 6.28 in T_{supv} (7.48), and 7.41 in T_{shared} (8.7). 52% of participants forecasted T_{worker} to be the most effective treatment in our paper, 4% chose T_{supv}, 28% chose T_{shared}, and 18% forecasted either two or all three treatments to have the same effect.

The next set of questions asked participants to estimate the magnitude of effort complementarities and contractual frictions after providing them with a brief description of our model.[50] The exact question regarding effort complementarities was: *"By how much do you think the effort of the supervisor raises the marginal return to the effort of the CHW in the shared incentive treatment? Please express this in percentage terms."*

As we show in Section 6.2, the estimate we get from the data is 36%. The average estimate from the survey participants is lower at 25%. The participants were then asked to answer the same question for supervisors and frontline workers of four other organizations, namely a retail firm in the private sector, a garment factory, a public school, and a central ministerial administration. The average estimates about the extent to which supervisor effort raises the marginal return to effort for frontline workers in these organizations by the participants are quite similar to the one they gave for our context: 27% for both the private sector retail firm and the garment factory and 20% for both the public school and the central ministerial administration.

The exact questions regarding contractual frictions was: *"By how much do you think that contractual frictions raise the cost to the supervisor of providing side incentives to*

[49]See https://socialscienceprediction.org. This prediction platform enables the systematic collection and assessment of expert forecasts of the effects of untested social programs.

[50]After answering a set of questions, it was not possible to change previous answers. Therefore, the information we gave participants about the model did not affect their forecasts about the average treatment effects.

the worker (i.e. what is the value of the contractual friction parameter)? Please express this in percentage terms."

As we show in Section 6.2, the estimate we get from the data is 300%. The average estimate of 42% from the survey participants reveals that they vastly underestimate the magnitude of contractual frictions in our setting. Interestingly, however, they do believe that contractual frictions in three out of the four organizations are higher than our setting. The average estimates are 56% for the garment factory, 67% for the public school, and 51% for the central ministerial administration. The only organization in which participants believe contractual frictions are lower than in our setting is the private sector retail firm at 38%.