The Allocation of Incentives in Multi-Layered Organizations*

Erika Deserranno, ⓡ Stefano Caria, ⓡ Philipp Kastrau, ⓡ Gianmarco León-Ciliotta[†]

September 18, 2022

Abstract

Does the allocation of incentives across the hierarchy of an organization matter for its performance? In a field experiment with a large public-health organization in Sierra Leone, we find that health-care provision is highly affected by how incentives are allocated between frontline workers and their supervisors. Sharing incentives equally between these two layers raises output by 61% compared to the unilateral allocations that are typical in public organizations. These results are surprising under a Coasian view of organizations, but can be reconciled under an alternative theory that emphasizes the coexistence of effort complementarities and contractual frictions. We leverage the experiment to estimate a structural model that quantifies these two forces, and their implications for the optimal design of incentive policies in multi-layered organizations.

JEL Codes: O15, O55, I15, J31, M52. **Keywords:** incentives, multi-layered organizations, effort complementarities, side payments, output

*The paper has benefited from the comments of participants at the NBER Development SI, NBER Organizations, BREAD, BREAD Africa, CEPR Development, CEPR Symposium, RIDGE Political Economy, RIDGE Public Economics, Barcelona GSE Organizational Economics and SIOE, and from many seminar participants. For suggestions that have substantially improved this article, we are grateful to Oriana Bandiera, Michael Best, Michael Callen, Katherine Casey, Ernesto Dal Bó, Claudio Ferraz, Frederico Finan, Rachel Glennester, Jonas Hjort, Anne Karing, Eliana La Ferrara, Rocco Macchiavello, Paul Niehaus, Anant Nyshadham, Alessandro Tarozzi, and Yusuf Neggers. Luz Azlor, Andre Cazor Katz, Margaux Jutant, Mustapha Adedolapo Kokumo, and Raquel Lorenzo Vidal provided outstanding research assistance. This project would have not been possible without the support from the staff at the Sierra Leonean Ministry of Health and Sanitation and the CHW Hub. Financial support 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 (CEX2019-000915-S) and the Ramón Areces Foundation. 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-0003345. There are no conflicts of interest.

[†]Authors are listed in a random order (AEA random author order archive code 69q7bqNU79Fz). Deserranno: Northwestern University, E-mail: erika.deserranno@kellogg.northwestern.edu. Caria: Warwick University, E-mail: stefano.caria@warwick.ac.uk. Kastrau: IDinsight, E-mail: p.kastrau@gmail.com. León-Ciliotta: Universitat Pompeu Fabra & Barcelona School of Economics & IPEG, E-mail: gianmarco.leon@upf.edu.

1 Introduction

Financial incentives are ubiquitous in modern hierarchical organizations. The impact of the *level* of these incentives — both at the top and bottom layer of the organization — has been explored in an extensive empirical literature (Bertrand 2009; Bandiera, Barankay, and Rasul 2011; Duflo, Hanna, and Ryan 2012; Finan, Olken, and Pande 2017). In contrast, the effects of the *allocation* of incentives across the different layers of an organization have been surprisingly neglected. Under a Coasian view of organizations, any allocation of incentives should result in the same level of output, as transaction costs within the organization are low and thus agents can redistribute the incentive through side payments (Coase 1937, 1960; Gibbons 1999). Under an alternative view, however, frictions and misaligned interests within the organization prevent side contracts between workers, making the specific allocation of incentives consequential (Jackson and Wilkie 2005; Garicano and Rayo 2016; Atkin et al. 2017). This is especially true in the presence of effort complementarities — a key feature of many organizations — since low effort in one layer of the organization cannot be easily compensated by higher effort in the other layer (Alchian and Demsetz 1972; Kremer 1993; Brynjolfsson and Milgrom 2013; Gibbons and Roberts 2012). Disentangling which view holds more truth is important because it sheds light on the nature of hierarchical organizations and because many organizations, especially in the public sector, choose unilateral allocations (Wade 1985, 1992; Meyer-Sahling, Schuster, and Sass Mikkelsen 2018).[1] Whether such unilateral allocations are optimal or not remains theoretically and empirically unclear.

In this paper, we show that the allocation of financial incentives within a large public-health organization substantially affects the provision of health care services in poor communities across Sierra Leone. In particular, we document experimentally that equally sharing an output-based incentive between a health worker and a supervisor generates an increase in output — in our context, health visits — that is 61% larger than the gain in output achieved when the incentive is offered entirely to the worker or entirely to the supervisor. Through a structural model and detailed survey data, we show that these results are driven by the combination of strong complementarity in worker and supervisor effort, *and* a limited redistribution of the incentive across layers. Our results thus support the second view of hierarchical organizations: complementarities are large, but Coasian contracting is imperfect and insufficient to maximize output. Using the structural model, we explore the quantitative implications of this finding for optimal policy design.

Our results help us understand how to expand access to health care in low-income countries, a major objective of global public policy (Dupas and Miguel 2017; Roser 2021). The World Health Organization estimates that half of the world's population lacks coverage for

[1]For example, a recent survey reveals that about 88% of public-health organizations in developing countries provide performance-based incentives uniquely to frontline health workers and not their superiors (Perry 2020). We discuss this further in the final part of the introduction.

essential health services and that health expenses are high enough to push more than 100 million people into extreme poverty (World Health Organization 2021). In the context of a low-income, post-conflict country with one of the highest infant mortality rates in the world (Casey and Glennerster 2016), we show that health care access and health outcomes can be substantially improved by changing the allocation of incentives, without altering their level.

The field experiment we design creates random variation in the recipient of a new piece-rate scheme across 372 health units of Sierra Leone. Each unit comprises an average of 8 health workers, who directly carry out the health visits, and one supervisor, who provides training, support, and advice. The incentive pays 2,000 Sierra Leone Leones (SLL, about \$0.25) per health visit completed and reported by the health worker, and is paid either (i) only to the health worker who carried out the visit, (ii) only to the supervisor of this worker, or (iii) is shared equally between the worker and the supervisor. Following Asiedu et al. (2021), we detail key aspects of research ethics (e.g., the AEA pre-registration and IRB) in the Appendix.

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 allocate the incentive between the workers and the supervisor. In the model, supervisors and workers 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 number of services delivered at the end of the game. In the second period, the worker chooses how much effort to exert to provide services. A key intuition is that the optimal share of the incentive to be offered to each agent depends on (i) the strategic complementarity of worker and supervisor effort, and (ii) the extent to which side payments offset the initial allocation of the incentive. When complementarities are limited, output is maximized by offering the entire piece rate to the worker. When complementarities are large and supervisors can incentivize workers with side payments, offering the entire piece rate to the supervisor can be optimal. Finally, when complementarities are large but contractual frictions constrain supervisors' ability to use side payments, sharing the piece rate between the two layers of the organization maximizes output.[2]

Our empirical analysis is divided into three parts. In the first part, we present the causal effects of our treatments on the number of visits carried out by the health workers. Importantly, our analysis does not rely on the number of visits reported by the health worker,

[2]In our setting, effort complementarities are likely to stem from the fact that supervisors raise the health workers' ability to conduct household visits by training and advising them, providing the necessary skills to perform their tasks, and helping them build trust in the community. Further, contractual frictions can derive from 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 appropriateness of side payments or institutional rules that limit managerial autonomy (Banerjee et al. 2020; Bandiera et al. 2021), or flypaper effects whereby payments are expected to stay in the layer of the organization to which they are originally allocated (Hines and Thaler 1995).

as this may differ from the actual number of visits due to reporting costs and moral hazard. Instead, we collect independent measures of completed health visits (quantity and quality) by interviewing a random sample of households in each village. Our central empirical finding is that the shared incentives treatment maximizes the number of completed health visits. Workers in the control group without any performance-based incentive (status quo) carried out 5.3 visits per household in the six months prior to our endline survey. This number significantly increases to 7.4 visits (a 40% increase over the control condition) when the incentive is offered either only to the worker or only 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 the targeting of poor and deserving households. Moreover, the share of households who report trusting the health worker is the highest in the shared incentives treatment. This result is important because trust in health service providers 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 2022; León-Ciliotta, Zejcirovic, and Fernandez 2022). We also find that the health worker's knowledge about how to adequately provide health services to the community is the highest in the shared incentives treatment.

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) in the shared incentives treatment than in the one-sided incentives treatments or the control. Households in this treatment also report fewer instances of fever among children below the age of five, and have better knowledge about how to prevent diseases. Administrative records from the local health facilities corroborate these results by showing that the number of recorded pregnant women services, institutional births, and fully immunized infants at the health facility are higher in the shared incentives treatment.

Importantly, shared incentives outperform one-sided incentives also in terms of cost-effectiveness. The incentive is only paid when a visit is *reported* by the health worker. Thanks to a system of extensive back-checks, we find that over-reporting is minimal. Instead, health visits are often under-reported, plausibly due to high reporting costs, which we discuss in Section 2. Crucially, under-reporting decreases with the share of the incentive offered to the worker. This makes shared incentives particularly cost-effective: we find that each 2,000 SLL spent on the program generates 16.1 extra visits in the shared incentives treatment, 9.6 extra visits in the supervisor incentives treatment, and 6.5 extra visits in the worker incentives treatment.

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 model, we show that both effort complementarity and limited side payments play an important role.

Three key results point to the presence of large effort complementarities. First, shared incentives generate the same increase in supervisor effort as supervisor incentives. 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 supervisor effort and hence *indirectly* incentivizes the supervisor to raise 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. Third, we carry out a formal mediation analysis which shows that the boost in visits due to worker effort increases with the level of supervisor effort.

Next, we turn to the role of side payments. We leverage detailed survey data on interpersonal transfers to show that, on average, net transfers from the supervisor to the worker are positive, but very small: less than 10% of the overall incentive payment of the average supervisor. Why are transfers limited? One possibility is that the poor observability of worker effort makes contracting hard. In line with this, we show that supervisors who plausibly cannot observe worker output accurately make lower transfers. Additionally, as we argue below, in many cases the worker may have a higher stake in the production of output than the supervisor. Transfers from workers to supervisors, however, are almost never observed in the data, suggesting that frictions may also prevent bottom-up transfers. Overall, the evidence is consistent with the existence of contractual frictions that prevent Coasian bargaining in the organization.

We rule out two main alternative explanations of our results. First, one-sided incentives treatments could be ineffective due to a negative morale effect arising from pay inequality (Breza, Kaur, and Shamdasani 2018; Cullen and Perez-Truglia 2022). Our experimental design minimizes this concern, as workers are not informed of the presence of supervisor incentives (if any) and only few seem to learn about it from the supervisors. Moreover, we find no evidence suggesting that workers in the supervisor incentives treatment are less satisfied with their payment or their job compared to the control group. The absence of treatment effects on job satisfaction is also hard to reconcile with alternative behavioral mechanisms, e.g. based on positive reciprocity. Second, we consider the possibility of strong diminishing returns to individual-level incentives in the utility function, or strong non-linearities in the cost or production functions. 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 (product) of effort increases (decreases) steeply after the level of effort generated by a 1,000 SLL incentive. However, 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 non-linearities. Similarly, we do not find evidence of sharp non-linearities in the relationship between supervisor effort and visits completed.

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 estimated model is able to match these moments with precision. Crucially, the model is also able to reproduce the key result that visits are maximized by the shared incentives treatment. In contrast, a version of the model based on a production function where efforts are *not* strategic complements has a much worse fit and wrongly predicts that worker incentives generate the largest increase in visits.

The estimated model parameters confirm that our results are driven by strong effort complementarity. In particular, we estimate that the marginal return to worker effort is up to 116% higher due to the complementarity with supervisor effort. Second, our calibrated contractual friction parameter implies that side transfers are 45% more expensive due to difficulties in contracting. Third, we find that, in the absence of the intervention, supervisors have weaker incentives to provide effort than workers. This underscores the importance of incentive schemes that ensure supervisors are adequately incentivized.

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 59% of the value of the incentive to the worker, and 41% to the supervisor. Second, we study how the optimal policy changes for different levels of effort complementarity. We find that the optimal allocation of the incentive is 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 makes interventions that tie incentives to joint output 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 show that the allocation of incentives in an organization with multiple tiers is highly consequential due to a combination of effort complementarities and a limited redistribution of incentives. The existing empirical literature has largely been unable to shed light on this point, since most studies to date have explored the effects of raising incentives in one layer of the organization (the bottom or the top), while holding incentives in the other layer fixed.[3] Our results document that agents engage in very limited fine-tuning of the allocation of incentives through transfers, partly due to the presence of contractual frictions.[4] Thus, there is little scope for Coasian bargaining within the organization, and there are large returns from picking the optimal allocation of incentives. These results further our understanding of the structural features of hierarchical organizations, and identify an overlooked policy lever to maximize public-sector effectiveness in developing countries (Callen et al. 2016; Bandiera et al. 2019; Finan, Olken, and Pande 2017).

Second, we provide evidence on the *productive* role of middle managers in hierarchical organizations. In doing this, we contribute to the literature demonstrating the relevance of management practices (e.g., Bloom and Van Reenen 2007; Bloom et al. 2013; McKenzie and Woodruff 2017; Bruhn, Karlan, and Schoar 2018; Macchiavello et al. 2020; Cai and Wang 2020). Importantly, our paper is also related to but differs from the long-standing literature that focuses on the *monitoring* role of managers. This literature — which spans seminal theoretical contributions (e.g., Tirole 1986, 1992) and recent empirical papers (Cilliers et al. 2018; Bandiera et al. 2021; Dal Bó et al. 2021; Dodge et al. 2021; Rasul and Rogger 2018; Kala 2019) — studies how to optimally delegate authority and how to avoid harmful collusion between workers and supervisors. However, this literature remains silent on how supervisor effort can directly increase the returns to worker effort. In our experiment, we explicitly minimize the scope for collusion through frequent back-checks of worker reports. This allows us to shed light on how the top layer of the hierarchy enables the frontline layer to be more productive, and on the implications of this complementarity for the design of incentives. Recent findings on the spillover effects of training interventions across the organizational hierarchy support our emphasis on the productive role of public-sector managers (Espinosa and Stanton 2022).

Third, we advance the literature on effort complementarities in organizations. Seminal

[3]These include papers that study incentives for the bottom layer — e.g., frontline workers or sales associates — while holding incentives for the top layer fixed (e.g., Glewwe, Ilias, and Kremer 2010; Muralidharan and Sundararaman 2011; Lazear 2000; Duflo, Hanna, and Ryan 2012; Ashraf, Bandiera, and Jack 2014), and papers that study incentives for the top layer — e.g., high-level public sector officials, private sector CEOs/managers — while holding incentives for the bottom layer fixed (Bandiera, Barankay, and Rasul 2007; Bertrand 2009; Frydman and Jenter 2010; Rasul and Rogger 2018; Luo et al. 2019). In the education sector, Behrman et al. (2015) evaluates 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). Geng (2018) complements Behrman et al. (2015) by showing evidence in favor of the presence of effort complementarities between students and teachers.

[4]A large literature has focused on contractual frictions *across organizations or firms* (Coase 1937; Gibbons 2005; Lafontaine and Slade 2007; Lee, Whinston, and Yurukoglu 2021; and, most related to our developing country setting, Macchiavello 2021). In this paper, we document the presence of contractual frictions *within an organization* (Adhvaryu et al. 2020).

theoretical work by Alchian and Demsetz (1972); Itoh (1991); Milgrom and Roberts (1995); Ray, Baland, and Dagnelie (2007); Brynjolfsson and Milgrom (2013) has reflected on the implications of complementarities for incentive design. Empirically, several studies have demonstrated that in "horizontal" teams — composed of workers from the *same* layer of the organization — group incentives that reward joint (rather than individual) output are effective even if at the potential cost of increasing free-riding (Muralidharan and Sundararaman 2011; Babcock et al. 2015; Friebel et al. 2017). However, unlike our paper, this literature does not shed light on the optimal allocation of such incentives. This is partly because, in "horizontal" teams, offering asymmetric incentive schemes to workers performing comparable tasks is often not a policy option due to fairness concerns (Card et al. 2012; Breza, Kaur, and Shamdasani 2018) or other rigid contractual arrangements. Instead, in "vertical" teams, asymmetric incentives are more acceptable since workers in the different layers of the organization have different responsibilities and levels of experience.

Our results have important policy implications for the design of community health programs worldwide. These programs have been shown to be crucial in reducing the burden of common diseases and child mortality in developing countries (Nyqvist et al. 2019; Deserranno, Nansamba, and Qian 2020). Finding ways to optimize the performance of community health workers is hence a first-order policy priority. Of the 57 community health worker programs reviewed in Perry (2020), 48% incentivize community health workers based on their performance while only 1.7% incentivize the supervisors.[5] In line with this, many of the social scientists invited to forecast our results on the Social Science Prediction Platform expect worker incentives to maximize output, despite being informed of the key role played by supervisors in our context.[6] Contrary to common policy design and expert predictions, our paper shows that this unilateral allocation may not be optimal.

2 Context and Research Design

2.1 The Community Health Program

Sierra Leone is a low-income, post-conflict country, with the third-highest maternal mortality rate and the fourth-highest child mortality rate in the world (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

[5] Several papers have studied the impacts of financial incentives for frontline community health workers. Singh and Masters (2017); Singh and Mitra (2017) report positive effects, while Shapira et al. (2018); Wagner, Asiimwe, and Levine (2020); Khan (2021) document zero or negative impacts. The results of our worker incentives treatments are more aligned with the first set of studies. There is also a large related literature on pay-for-performance in the health sector of developing countries beyond community health worker programs, e.g., Basinga et al. (2011); Celhay et al. (2019); Mohanan et al. (2021).

[6] See the Conclusion for more details on the results and the platform.

limited access to health facilities throughout the country (World Health Organization 2016). 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, nurses, and midwives. Each PHU typically has a catchment area of seven to ten villages with one community health worker per village and one supervisor per PHU, for a total of approximately 15,000 health workers and 1,500 supervisors nationwide.

The health workers and the supervisors are part-time workers who work around 20 hours per week and typically maintain another secondary occupation (e.g., farming, shopkeeper). They are paid a fixed monthly allowance of 150,000 SLL and 250,000 SLL by the MoHS, respectively, corresponding to a standard local monthly salary of $18 and $29 per month (January 2019 exchange rate). Health workers are hired locally, typically have no experience in the health sector prior to joining the program, and are trained/monitored by the supervisor after joining the program. Supervisors usually have experience working as a health worker.

Role of the health workers (bottom layer) The role of the health workers is to provide a package of basic health-care services in their community. They do so by making home visits to expecting mothers or mothers who recently gave birth, during which they provide: (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 a health facility when necessary. To ensure high-quality visits, workers are asked to follow a checklist each time they provide a service. We describe the checklists in Appendix B.1.

Role of the supervisors (top layer) The role of the supervisors is to train and advise health workers in their PHU (typically, seven to ten health workers per supervisor). They do so in three ways. (i) They organize monthly one-day "general trainings" at the local health facility which cover vital health topics, such as diagnosing, treating, and recognizing danger signs for referral to health facilities. (ii) They organize "one-to-one trainings" with health workers monthly in their respective villages. (iii) They provide "in-the-field supervision" by accompanying health workers on household visits. During these household visits, supervisors are neither tasked to provide services themselves to the households, nor are they in charge of scheduling or setting up the visits. Instead, their role consists in providing health workers with concrete feedback on how to improve service delivery and continuous on-site training. Supervisor's presence during these household visits also helps build trust toward 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. Thus, overall, a substantial share of the support offered to the worker is personalized, which limits the potential for economies of scale in supervisor effort. Personnel decisions (hiring, firing, promotions, etc.) are taken by the head of the PHU and not by the supervisors. We provide more details on the supervisors' tasks in Appendix B.2.

Complementarities across layers In the study setting, supervisors are mostly engaged in supporting frontline workers. This is a common arrangement in many organizations, and sets our paper apart from recent literature that focuses on the monitoring role played by middle managers (Callen et al. 2020; Muralidharan et al. 2021; Bandiera et al. 2021; Dal Bó et al. 2021; Dodge et al. 2021). In our context, supervisors generate demand for the workers' services by training the workers and building trust towards them in the community. This can create a strategic complementarity between worker and supervisor efforts. When a supervisor increases her effort, the worker is able to generate more visits for the same amount of time spent in the community. Similarly, the effort of the supervisor has a larger return when the worker is motivated and makes the most of the stronger demand for their services created by the supervisor.

2.2 Intervention and Research Design

We study the introduction of a new incentive scheme that pays a piece rate of 2,000 SLL ($0.25) for each reported household visit. We have four experimental conditions. In the *worker incentives treatment* (T_{worker}), the incentive of 2,000 SLL is paid entirely to the health worker who provides the visit.[7] 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.[8]

[7] The size of the piece rate is substantial: a health worker can earn up to 14% of her monthly fixed allowance if she provides one visit every other day.

[8] An alternative design to study effort complementarities would offer (i) an incentive of 2,000 SLL to the worker *and* the supervisor in the shared incentives treatment, (ii) an incentive of 2,000 SLL to the worker in the worker incentives treatment, and (iii) an incentive of 2,000 SLL to the supervisor in the supervisor incentives treatment. In this design, complementarities can be detected by testing whether the impact of the two-sided incentive is greater than the sum of the impacts of the one-sided incentives. However, this would entail varying the amount of the incentive per visit disbursed by the organization, and hence would not shed light on the key question we aim to answer, i.e., how a given incentive should be allocated across the layers of an organization with limited liquidity. Appendix B.3 discusses why we opted for equally splitting the incentives across layers in the shared incentives treatment.

Our experiment takes place in 372 PHUs, with the intervention running from May 2018 to August 2019. The 372 PHUs are located throughout Sierra Leone and were randomly assigned to one of the four experimental groups in equal proportions. Because staff interactions are common within a PHU but minimal across PHUs, the randomization was performed at the PHU level to limit spillovers across treatments. The randomization was stratified by district, average distance between the residence of the supervisor and the health workers in the PHU, and number of health workers in the PHU. Importantly, a 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 we study in Deserranno, Kastrau, and León-Ciliotta (2021). In Appendix B.4, we describe the change in the promotion system and show that the results of this paper are orthogonal to this variation.[9] Appendix B.4 details the experiment's location and the randomization.

Note that the incentive scheme we study rewards health workers and supervisors for output (household visits), rather than direct measures of effort (e.g., number of households the health worker attempts to visit, number of trainings the supervisor provides to the health workers). Output incentives have the advantage of rewarding workers based on a measure that is more verifiable than effort, and are widespread both in the private and public sectors.[10] As we will show later, output incentives also have the advantage of incentivizing both the worker and the supervisor to internalize some of the positive spillovers of their effort on the productivity of other subjects.

Description of the intervention The incentive scheme has two important features.

First, the incentives were disbursed by a reputable external organization independent from the government. Subjects were paid monthly through mobile money and without any delay. This enabled us to establish the credibility of the new incentive scheme in the eyes of all experimental participants.

Second, incentives were paid based on worker self-reports. This is a common arrangement for incentive schemes with decentralized delivery agents, as directly monitoring output is typically expensive and impractical (e.g., Soeters and Griffiths 2003; Shapira et al. 2017). To report a visit, the worker must send an SMS from their main phone number to a toll free number. To trigger a payment, the SMS needs to indicate the date of the service and

[9]Specifically, we 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.

[10]In the financial sector, for example, a large fraction of the pay of financial analysts is variable and proportional to the amount of capital 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 salespeople 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.

the contact number of the patient, and needs to be sent from the worker's registered phone number. The latter implies that supervisors or households are unable to report services on behalf of the workers. All health workers in our study (including those in the control group) were asked to report their visits, but only those in the treatments were incentivized based on the SMSs. We present more information about the reporting system in Appendix B.5.

Our set-up discourages over-reporting through extensive back-checks and strong penalties. A random 25% of reports are verified by contacting the household mentioned in the report, and a worker caught reporting a visit that did not occur would not be eligible for any further incentive payment and would be reported to the MoHS. Back-checks and penalties were discussed extensively during the training on the reporting system that workers received prior to the start of the intervention. To keep things as comparable as possible across experimental groups, all workers received the same training and the same number of back-checks, including those in the control group.

We will later show that the threat of being caught "cheating" was credible and nearly eliminated over-reporting. Our design, however, does not prevent under-reporting. Even though the SMS reporting tool is free to use, reporting is inherently costly. First, reporting takes time and requires gathering information on the patients' name and phone number, which patients may not always be willing to share. Second, mobile phone coverage is unreliable and unpredictable in rural areas of Sierra Leone, thus limiting health workers' ability to send the SMS on the spot. This can lead to under-reporting if the worker subsequently forgets to send the SMS or sends an incomplete SMS with missing information. In Section 4.2, we will show that under-reporting is frequent in our setting. Similarly low reporting rates have been documented in other low-income countries. Karing (2021), for example, shows that local health facilities in Sierra Leone under-report vaccination entries, despite the presence of financial incentives, and that this is likely due to hassle costs.

Transparency of the incentive scheme To mirror most workplace environments where supervisors have information about the pay structure of the subordinates, but subordinates are not informed about their superior's compensation (Cullen and Perez-Truglia 2019, 2022), we informed all supervisors in the study about the worker incentives but did not inform the workers about the supervisor incentives. As we will later show, this limits negative morale concerns resulting from pay inequality. Workers could only learn about the presence of supervisor incentives from the supervisors themselves, and few supervisors seem to have shared this information with their workers (see Section 5.3).

Importantly, we did not provide supervisors with information about the number of SMSs sent by each worker or about information about worker earnings from the incentive scheme. We did so because, as we will later show, workers' reporting behavior varies considerably across experimental groups. As a result, disclosing information about the number of visits reported by each worker to supervisors would have introduced differential observability of worker effort across treatments, and hence would have confounded the interpretation of our results. The fact that supervisors are unaware of worker earnings also further minimizes the possibility that the supervisor and the worker collude to report visits that have not actually been carried out.

Side payments We clarified to all supervisors that they could share all or part of their incentives with workers at their discretion. These transfers could potentially be used to incentivize worker effort. However, whether supervisors will provide such payments will depend largely on the existence of contractual frictions. In our context, these frictions can derive from three factors that are common in organizations. First, supervisors have limited ability to precisely observe the worker's level of effort and reporting behavior since production is decentralized (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). Third, there may be social norms or psychological factors that limit redistributions within the boundaries of the same organization (Hines and Thaler 1995). The second and third features are also likely to inhibit transfers from workers to supervisors.

2.3 Data and Balance Checks

2.3.1 Data Sources

We leverage three main sources of data.

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

Household surveys. A random sample of three eligible households per village (∼7% of the households) were surveyed at endline in June-September 2019. The respondent of the survey was the female household head, who is typically the most knowledgeable about health topics. Each respondent was asked questions on the number of visits received by the health worker and the quality of these visits, trust in the health worker, disease incidence among young children, access to pre- and post-natal care. We will later use these data as our primary measures of health worker performance.

Administrative data. We have access to two sources of administrative data. First, we observe the number of valid SMS reports each health worker sent throughout the experiment, 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 (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).

2.3.2 Summary Statistics and Balance Checks

Table 1 reports summary statistics and balance checks for the characteristics of the supervisors (Panel A), health workers (Panel B), households (Panel C), and villages (Panel D).[11] Panel E reports statistics on the number of health services provided by the local health facility (one per PHU) in the month before the start of the experiment.

Panel B shows that 71% of the health workers in our sample are male, 70% have completed primary education, and 8% have completed secondary school. On average, health workers are 37 years old, are responsible for 55 households each, and live 3.4 km away from the supervisor. Panel A shows that the supervisors are more likely to be men than the health workers (92%) and are more likely to have completed secondary school (25%). They are responsible for an average of 8 health workers each. Panel C shows that 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. Households live on average 1.4 km away from the health worker.

Panel D shows that 77% of the villages in our experiment have an accessible road to the health facility. Phone network is available in 84% of the villages but is mostly unreliable. We will later show that the lack of reliable network substantially increases the cost of SMS reporting. Finally, Panel E shows that health facilities record 47 pregnant women visits per month, 13 institutional births, 11 infants immunized, and 66 cases of malaria and diarrhea among children under five.

To perform the balance checks, we regress each baseline characteristic on a dummy variable for each of the three treatments, controlling for the stratification variables and clustering

[11]Given the absence of a baseline household survey, we asked households in our endline survey a set of retrospective questions that are unlikely to vary over time (i.e., age, education, location) and report those in Panel C.

standard errors at the PHU level in worker/village level regressions. Column (11) of Table 1 reports the p-value from a joint F-test of the equality of all treatment groups. The baseline characteristics are balanced across treatments except for the age of the health worker (p-value of 0.062). In Table A.1, we report the p-value for each pairwise treatment comparison. Out of 156 pairwise comparisons, 16 are statistically significant with a p-value below 0.1.

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 effort complementarities and contractual frictions 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).[12] 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 and advising 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 she offers, 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 supervisor's 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 $z \geq 1$. 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 efforts. We capture this with the following output function:

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

where α is 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

[12] This departs from our empirical setting, in which supervisors are responsible for multiple workers. As explained at the end of the section, this simplification does not affect the main results of the model.

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 players 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).[13] 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*, $p = 0.5$; 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.[14] 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.[15] 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:

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

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

We solve the model using backward induction. 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 costs or benefits are asymmetric. We will relax these assumptions when we take the model to the data in Section 6.

In this simplified setting, the optimal side payment is given by:

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

This formula shows that the optimal side payment decreases with the contractual frictions (z) and the incentive offered to the worker (p). 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 payment. We derive optimal efforts for these two cases — positive side payments ($p \leq \frac{1}{1+z}$)

[13] In the empirical setting, agents also receive a fixed wage. Given the linear utility specification, the introduction of this additional term does not affect our conclusions.

[14] In the empirical setting, agents are paid uniquely for the visits they performed *and* subsequently reported. We abstract from modeling worker reporting behavior because it complicates the model without affecting its main results. See the discussion at the end of the section.

[15] In practice, transfers from supervisors to workers could be fixed (not proportional to visits) or based on the number of visits reported by the worker. Again, such extensions do not affect the central intuition of the results.

and zero side payments ($p > \frac{1}{1+z}$) — and present the complete mathematical analysis of the model in Appendix D. As expected, the efforts of both players increase in the strength of the complementarity. Further, due to the complementarity, agents' efforts do not necessarily increase monotonically in the share of the incentive offered to them.

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 aims 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." Finally, we restrict attention to values of γ and c such that $z\gamma^2 < 8c^2$. This condition limits the relative size of the complementarity, guaranteeing positive optimal efforts (as we show in Appendix D.2). 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 equal or larger than t, there is always a two-sided scheme that 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.*

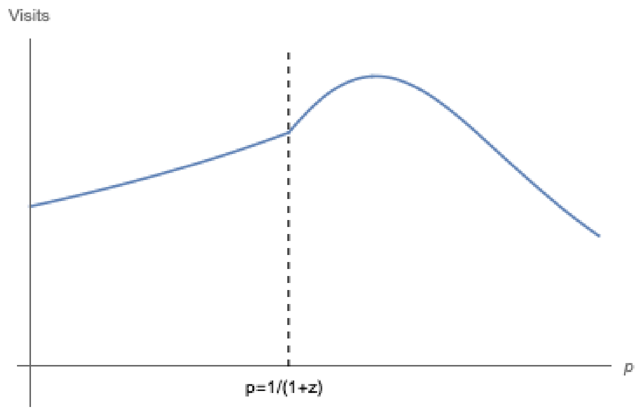
This result is established in two steps, which are discussed in detail in Appendix D and summarized here. When complementarities are low ($\gamma < t$), supervisor effort has only a limited effect on the worker's ability to carry out household visits. In this case, it is straightforward to show that household visits are maximized by offering the entire incentive to the worker.

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 shown in Figure 1.[16]

If complementarities are large ($\gamma \geq t$) and 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 (as shown, for instance, in the example analyzed in Figure A.1b), then all $p \, \epsilon \left[0, \frac{1}{1+z}\right]$ are optimal.

[16]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 $(\frac{1}{1+z}, 1]$). In Figure A.2, we show how optimal efforts and side payments change as p changes.

Figure 1: Optimal Incentives ($\gamma \geq t$ 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 D, 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 incentive scheme is larger for supervisor-worker pairs that have a high γ. We will explore these predictions empirically in Section 5.

The model also sheds light on the important role played by side payments. In particular, two predictions will help us interpret our experimental results. 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 contrast, when there are contractual frictions, changes in the allocation of incentives always affect output. In other words, if we observe positive side payments and differential treatment effects on output, this indicates that the supervisor and the worker cannot contract costlessly. Second, the model shows that there is an additional factor that can limit side payments. In Appendix D.7, 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 payment when the benefit b_2 that she receives from household visits absent our intervention is low compared to the benefit b_1 that is received by the worker.[17] In these cases, it would be optimal for the worker to pay the supervisor to exert effort — an action which we do not allow in the model and do not observe in the data, presumably because the frictions preventing transfers from the bottom

[17]Both b_1 and b_2 may stem from the agents' worry that low output will result in their dismissal from the organization. As supervisors tend to be more experienced and established in the organization, they are likely less concerned by the possibility of being fired, compared to workers. Alternatively, supervisors may have better outside options and would hence suffer a smaller utility loss if they lose their position.

to the top of the hierarchy are even larger than the frictions that impede transfer from the top to the bottom. In sum, the lack of side payments is theoretically consistent either with high contractual frictions preventing the supervisor from offering side payments, or with an asymmetry in how much workers and supervisors value output. However, in the latter case, these limited side payments are sufficient to equalize output.

Finally, we note that we depart from our empirical setting in two main ways to keep the model tractable. First, we abstract from the fact that each supervisor has multiple workers. This prevents us from exploring the optimal targeting of supervisor effort across heterogeneous workers but does not affect the model's main predictions. Second, in the model, the incentive is paid based on the number of actual visits completed rather than the number of visits reported. In the structural estimation Section 6, we present a version of the model in which incentives are based on the number of visits reported. To model under-reporting, we posit that the reporting process suffers from random shocks (e.g., bad network), which prevent some visits from being reported. We allow the reporting rate to differ by treatment since presumably, workers can take costly actions to override the shock, but their willingness to take these actions is a function of the incentive they get paid for each report. Indeed, empirically, the reporting rate increases in p. This raises the relative attractiveness of the worker incentive scheme compared to the other schemes. However, as long as the elasticity of reporting with respect to p does not exceed a threshold, all model results remain qualitatively unchanged.

4 Main Results

4.1 Output

We estimate the effect of our treatments on output with the following regression equation:

$$Y_{ij} = \alpha + \beta_1 T_{worker,j} + \beta_2 T_{supv,j} + \beta_3 T_{shared,j} + Z_j + \varepsilon_{ij}, \tag{5}$$

where Y_{ij} measures the quantity or quality 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 discussed in Section 2.2. We estimate standard errors clustered by PHU (level of the randomization), and report p-values corrected for multiple hypothesis testing using three alternative procedures.[18] The outcome variables and the analysis were

[18] The three procedures are: Bonferroni, Romano and Wolf (2016), and Benjamini, Krieger, and Yekutieli (2006). The Bonferroni procedure controls for the familywise error rate. This procedure is conservative, as it assumes that test statistics are independent. Therefore, we also present corrected p-values following the

pre-registered in the AEA RCT Registry.[19]

Quantity of visits We start by assessing the treatment effects on the incentivized measure of output, i.e., the *quantity* of visits provided by the health worker. To measure the latter, we do not rely on the number of visits reported by the worker because this often differs from the actual number of visits due to under-reporting, as discussed in Section 4.2. Instead, we asked each sampled household the total number of natal- and disease-related visits they received from the health worker in the six months preceding the endline survey.[20] For each health worker, we then calculate the mean number of visits received by a household (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 visit types received by a household (mean of 1.7).

Our main results are reported in Table 2 column (1) and the corresponding Figure 2. They show that introducing performance-based incentives significantly boosts the number of household visits the health worker provides, 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 (39%) in the worker incentives treatment, by 2.145 (40%) in the supervisor incentives treatment, and by 3.356 (63%) in the shared incentives treatment. These results are all statistically significant at the 1% level. Interestingly, offering the whole incentive to the health workers is equally effective than offering the whole incentive to the supervisor. Both interventions, however, are outperformed by the two-sided incentive scheme, which achieves 17% more visits overall. Relative to the control group, the boost in visits generated by the two-sided incentive scheme is 61% larger than the boost in either of the one-sided schemes.[21]

When we break down household visits by their type, we find that, compared to the one-

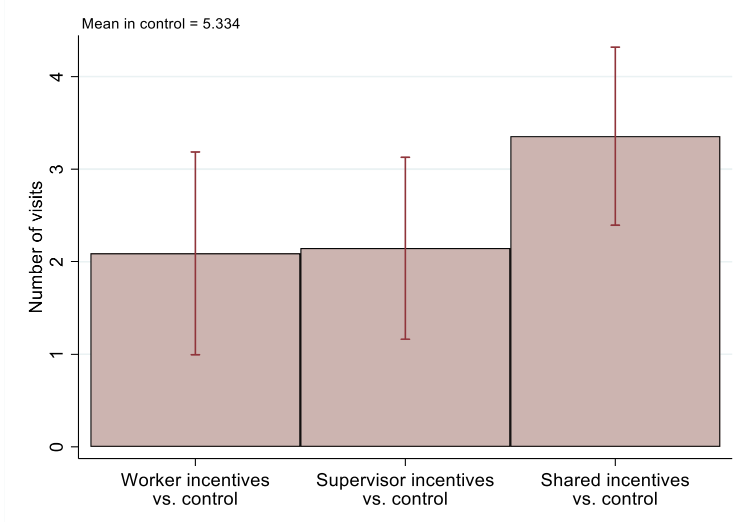
procedure in Romano and Wolf (2016), which accounts for dependence across test statistics. Furthermore, we include sharpened q-values following the approach used in Benjamini, Krieger, and Yekutieli (2006). This procedure controls for the false discovery rate, and it typically preserves even more power at the cost of some type I errors.

[19]The only deviation from the pre-analysis plan is that we do not study treatment effects on the "number of hours that the workers report dedicating to their job". This is because the variable exhibits limited variability in the data, presumably due to a self-reporting bias (as also reported by enumerators in the field). See Appendix C for more details.

[20]To minimize recall bias, households were asked about visits received "since the start of the year," which roughly corresponds to the past six months.

[21]The results in Figure 2 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.2, 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}.

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 with standard errors clustered at the PHU level. Bars are 95% confidence intervals.

sided treatments, shared incentives generate significant gains over both natal-related and disease-related visits (Table A.3). Health workers in the shared incentives treatment also achieve higher coverage and provide a higher variety of services (Table 2, columns 2 and 3). The results are robust to correcting p-values for multiple hypothesis testing (Table A.4, Panel A).

Quality of the visits and targeting The larger number of visits provided by workers in the shared incentives treatment may come at the expense of visit length (which is not incentivized), so that the aggregate amount of time dedicated to the job remains unchanged. This would be problematic: as discussed earlier, workers are expected to follow a checklist when they visit a household. Short visits may indicate that such checklist is not properly followed, and thus, the service provided may be of lower quality.

We do not find a quantity-quality trade-off. Table 2 (column 5) shows that, conditional on having received at least one visit, the average visit length reported by a household (23 minutes) does not decrease in the shared incentives treatment relative to the control group, while the number of health topics discussed per visit increases by 15% (column 7). The latter is consistent with the health workers receiving more training from the supervisor, as further discussed in Section 5.1. If we set the average visit length and the number of health topics discussed to zero for households who were never visited, we obtain that the shared incentives increase visit length by 34% (column 4) and the number of health topics discussed by 26% (column 6). This captures both the intensive and the extensive margin of effort. Importantly, the shared incentives also maximize trust: the share of households who report trusting the health worker in the shared incentives treatment is 7.1 percentage points (10%) higher than in the control, and 3.5 percentage points (5%) higher than in both one-sided incentives treatments (column 8). The results are robust to multiple hypothesis testing corrections (Table A.4, Panel A).

Next, we examine the possibility that the higher number of visits in the shared incentives treatment comes at the expense of worse household targeting, i.e., health workers switching from visiting poor and deserving households to visiting households who are geographically and socially close to them (who are presumably less costly to visit).[22] Table A.5 shows that households who are geographically and socially close to the health workers are visited more often than households who are not, but that this equally true in all treatments. This further alleviates concerns related to quantity-quality tradeoffs and to misreporting driven by worker-household collusion.

A last possibility is that the higher number of visits in the shared incentives treatment comes at the expense of health workers diverting their time away from providing long and complex pre- and post-natal checks into short and easy routine visits. As shown earlier, the shared incentives treatment does not affect the mix of services provided relative to the control (Table A.3), and does not reduce the visit length (Table 2, column 5).

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 access to health services and better health outcomes.

We start by analyzing households' access to pre- and post-natal care. We measure access with an equally-weighted average of the z-scores of key indicators of natal care quality according to the World Health Organization framework (four pre-natal visits, institutional birth, post-natal care within two days of delivery, up-to-date vaccination, breastfeeding).[23] Table 3 (column 1) shows that the shared incentives treatment leads to better access to pre- and post-natal care. More precisely, the pre- and post-natal care index is 0.092 standard deviations higher in the shared treatment relative to the control (significant at the 1% level).

[22] 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., distance to the health worker, social relationship with the health worker). All the other variables are defined as in equation (5). ε_{hij} is an error term clustered at the PHU level.

[23] Questions on pre-natal and post-natal care were asked to households composed of a woman who gave birth in the year preceding the endline survey. Table A.6 (column 1) shows that our treatments do not affect this fertility measure. Pre-natal care is measured by asking women who gave birth in the past year whether they received at least four pre-natal visits from any provider, and post-natal care is measured by asking them whether they gave birth in a health facility (vs. at home), whether they received at least one post-natal visit within two days of delivery, whether they breastfed their infant for at least six months, and whether their infants are up to date on the vaccination schedule.

Columns (2) to (6) present the results for each single component of the index.

Next, we analyze disease 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.[24] Table 3 (column 7) shows that the disease incidence index is 0.053 standard deviations lower in the shared incentives treatment than in the control group (significant at the 5% level). 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). These households also have better knowledge about how to prevent malaria (i.e., sleep under a treated bednet) and diarrhea (i.e., wash hands with soaps, drink clean water): see Table A.6 (column 2). We find no significant effects on under-five mortality rates (Table A.6, column 3), presumably due to the relatively short timeframe of the experiment. The results are robust to multiple hypothesis testing corrections (Table A.4, Panel B).

We corroborate these health results with administrative records from the local health facility (PHU-level data), which do not suffer from any recall or response bias. The results are presented in Table A.7, columns (1)-(7). In line with the household survey data, we find that the number of recorded pregnant women services, institutional births, and fully immunized infants at the health facility is higher in the shared incentives treatment than in the other groups, albeit the results are less precisely estimated. All three incentive treatments appear to increase the number of malaria and diarrhea cases treated at the health facility relative to the control group. Given the lower disease incidence rate reported by our sampled households, these positive coefficients are consistent with health workers referring sick children to the health facility more frequently in the treatment groups than in the control group.

4.2 Reporting and Cost-Effectiveness of the Intervention

This section assesses the relative cost-effectiveness of the three incentive schemes. All schemes pay 2,000 SLL ($0.25) per visit reported by the health worker. Cost-effectiveness is thus a function of both the actual number of visits carried out and the number of visits reported.

We start by evaluating, in Table 4, whether our treatments impact the number of visits reported. Column (1) shows that reported visits are highest in the worker incentives treatment, even though we have shown that actual visits are maximized by shared incentives. More precisely, we find that, in the six months preceding the endline survey, workers send an average of 8.7 SMS reports per month in T_{worker}, 6.3 in T_{shared}, and 3.7 in T_{supv}. The reporting differences across treatments are relatively stable over time (see Figure A.3). These results imply that the most expensive incentive scheme for the organization is T_{worker}. More

[24]The three most common diseases among children in Sierra Leone are malaria, pneumonia, and diarrhea. Because households may not know 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).

precisely, Table 4 (column 8) shows that the new incentive scheme costs the organization an average of 131,593 SLL in T_{worker}, 93,953 SLL in T_{shared}, and 54,108 SLL in T_{supv}.

In Table 4 (column 2), we present 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.[25] We also present results on dummy variables capturing whether a worker under-reports or over-reports actual visits. This analysis shows that health workers generally 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}. Moreover, the share of workers who under-report is 12 times larger than the share of workers who over-report (columns 3-4).

Low over-reporting is also observed in the back-checks data.[26] Only 3.6% of the health workers in the worker incentives treatment ever reported a visit that the recipient household did not confirm having received during the back-check, vs. 2.7% in the shared incentives treatment, and 2% in the supervisor incentives treatment.

Overall, the results confirm that over-reporting is minimal but, as discussed earlier, under-reporting is more common. This is consistent with the fact that reporting is inherently a costly action, so that workers under-report even in T_{shared} and T_{worker} when they receive a monetary incentive for reporting. A key driver of the reporting costs appears to be the lack of reliable mobile phone network, which prevents health workers from sending the SMSs. We find indeed that the reporting rate is close to zero in villages where network connectivity is virtually absent, regardless of the level of the incentives. In villages with some network connectivity (even if unreliable), the reporting rate is higher and varies with the level of the incentives.[27]

Given that under-reporting is widespread and differential across treatment groups, in the structural estimation (Section 6), we extend our basic model to explicitly take under-reporting into account.

Policy choice What policy should the organization adopt based on these results? Suppose the organization wants to maximize household visits, conditional on the payment per actual

[25] The actual number of visits per month 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, as in Table A.2 column (2). While the reporting rate we obtain from this calculation may be over- or under-estimated for a single health worker due to a sampling error, average differences across treatments are meaningful and accurate. Note that households have no strategic incentive to misreport the number of visits received by the health worker and that the survey was not announced beforehand so that the health worker could not have influenced households to give favorable answers during the survey.

[26] Recall that the SMS reports were back-checked by a team of phone monitors who contacted a random 25% of patients, and asked them to confirm the date and type of service received.

[27] See Table A.8. In villages with network connectivity, we estimate an elasticity of reporting to incentives of 0.75, and hence estimate that an incentive of 3,800 SLL would lead health workers to report 100% of their visits. The results are robust to controlling for correlates of network connectivity (e.g., distance to urban area) interacted with the treatment dummies (Table A.8, column 2).

visit not exceeding 2,000 SLL. In this case, the shared incentive intervention is unambiguously optimal for the organization. On the one hand, over-reporting is minimal in all treatments, so the cost per actual visit never exceeds 2,000 SLL. On the other hand, shared incentives maximize actual visits, and so satisfy the organization's objective.[28]

The shared incentive intervention is also optimal if the organization's objective is to maximize cost-effectiveness, i.e., to maximize the number of actual visits generated per dollar spent. In Table 4 (column 9), we show that in the worker incentives treatment, the organization obtains an additional 6.5 visits per worker for each 2,000 SLL spent on incentives. This figure goes up to 9.6 visits for each 2,000 SLL spent in the supervisor incentives treatment, and to 16.1 visits for each 2,000 SLL spent in the shared incentives treatment (a significant difference of 9.6 visits compared to worker incentives).[29]

Shared incentives, however, impose a larger total cost compared to supervisor incentives. If this cost breaks the organization's budget constraint, the organization could either opt for supervisor incentives, which offer a similar increase in visits as worker incentives, for a lower cost; or it may decrease the amount of the incentive paid in the shared incentive scheme.

5 Mechanisms

The previous section showed that health workers provide significantly more household visits under shared incentives than one-sided incentives schemes, with no concomitant reduction in visit quality. 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 limited side payments. We then present evidence against two alternative mechanisms.

5.1 Effort Complementarities

Three pieces of evidence point to the presence of strong effort complementarities in our setting. We discuss each in turn.

[28] Note that over-reporting is minimal in our setting thanks, in part, to the presence of a relatively sophisticated and independent monitoring system. In the absence of such a monitoring system, over-reporting may occur more frequently, especially in the worker incentives treatment. This would presumably make the worker incentives even less attractive relative to the other treatments.

[29] A key caveat is that these results are partly driven by the differential rate of under-reporting. Due to under-reporting, the organization saves on incentive payouts that should instead accrue to workers and supervisors. If under-reporting was reduced, e.g., by lowering the reporting costs, differences across treatments in the number of additional visits produced for each 2,000 SLL spent would be smaller. Further, the organization may want to design a scheme to reimburse agents for the incentives that have not been claimed, for equity reasons or to comply with labor laws, which would also reduce differences in cost-effectiveness. Finally, differences in cost-effectiveness would be smaller in settings where output is observable and hence incentives can be tied to actual output.

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 weak ($\gamma < t$), the effort of the supervisor should monotonically increase with the level of the 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 training and advising by accompanying health workers on household visits (henceforth, an "accompanied visit"), (ii) organize one-to-one meetings with each health worker, and (iii) organize monthly one-day general trainings. We measure (i) with the fraction of households who report having received an accompanied visit (mean of 20%).[30] We measure (ii) and (iii) by asking health workers the number of times the supervisor provided them one-to-one meetings in the six months preceding the endline survey (mean of 1.4) and whether the supervisor organized a general training in the last month (mean of 99.4%). As a proxy measure of training quality, we administered a quiz on health knowledge to all health workers at baseline and endline, and estimate the difference in health knowledge over time.[31]

Table 5 column (1) shows that the share of households who report having received an accompanied visit 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, while there is no difference between T_{worker} and the control group. Notably, 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 that indirectly compensate the supervisor in the shared incentive scheme for the lower monetary payment. Columns (2) and (3) of Table 5 show that our treatments neither affect the number of times the supervisor provided one-to-one meetings to the health worker nor do they affect the likelihood that the supervisor organized a general training. The latter is not surprising as supervisors are required to organize such trainings on a monthly basis, and nearly all of them do so.

Two pieces of evidence provide direct support to the fact that supervisors play an "en-

[30] Among households who received at least one accompanied visit, 97% received one visit only, and 3% received two visits. Further, no household reports having received a visit from the supervisor without the presence of the health worker. Overall, this implies that roughly one-fifth of the households have seen the supervisor once in the past six months, and the vast majority of the remaining households have never seen the supervisor.

[31] The quiz counts the number of correct answers given to the following five questions: (i) Assuming there are otherwise no danger signs, after how many days of fever should the health worker refer a child under 5 to the PHU? (ii) Assuming there are otherwise no danger signs, after how many days of loose or watery stools should the health worker refer a child under 5 to the PHU? (iii) For a child between 2 and 11 months of age, how many breaths per minute is the threshold from which it counts as fast breathing? (iv) For how many years after giving birth should a woman wait before falling pregnant again? (v) Is a composting toilet an improved or unimproved sanitation facility?

abling" role in our context, and not only a "monitoring" role. First, Table 5 column (4) shows that health workers improve their health knowledge the most in the shared incentives treatment. This is consistent with health workers receiving higher-quality training from the supervisor in the shared incentives treatment, and health workers being able to discuss more health topics during the household visits (as shown in Table 2).

Second, Table A.5 (column 5) shows that the boost in the number of visits (accompanied or unaccompanied) in the shared incentives treatment is similar for households that ever received an accompanied visit and those that did not, with the obvious caveat that this variable is endogenous. This implies that the boost in visits in the shared incentives treatment is not entirely driven by more accompanied visits (which are few) but rather by more unaccompanied visits. Moreover, if the role of the supervisor was limited to monitoring, we would expect health workers to target their visits towards households that were in direct contact with the supervisor in the past, since presumably, the supervisor would find it easier to contact these households again and to monitor whether the worker has visited them. That shared incentives boost visits for households who were never in direct contact with the supervisor suggests instead that health workers in this treatment have received better training and are able to raise demand for their services even when unaccompanied.

Finally, note that only 9% of the health workers report that their supervisors ever helped them with SMS reporting (Table 5, column 5). This is not surprising as all health workers received extensive training on reporting at the start of the experiment (see Section 2 and Appendix B.5). Moreover, the share of supervisors who help health workers with reporting is comparable across treatment groups. 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) towards helping with reporting.

Heterogeneity by health worker's experience Next, we present heterogeneous treatment effects by an empirical proxy of effort complementarity: limited health worker's experience. Health workers with little experience are less well-trained about health issues and less known as a health worker in the community, and they thus plausibly benefit more from the training and advice of the supervisor. We therefore expect the shared incentives treatment to be more effective in boosting output and supervisor effort for these health workers, compared to their more experienced counterparts.

Table A.10 estimates a fully interacted model and presents the treatment effects for workers with experience below the median (4 years) in Panel A and for workers with experience above the median in Panel B. For inexperienced workers, the shared incentives treatment increases the number of household visits provided by the health worker by 4 (85%) relative to the control group (column 1), and increases supervisor effort (share of visits accompanied by the supervisor) by 9.2 percentage points (70%; column 3). For experienced workers, these effects are significantly lower: they are about half the magnitude for visits and one-third of the magnitude for supervisor effort. Columns (2) and (4) of Table A.10 show that the results are robust to controlling for correlates of health worker experience listed in Table A.9 (column 1), and their interaction with the three treatment indicators. 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.[32]

Mediation analysis As a final piece of evidence in favor of effort complementarities, we perform a mediation analysis to test whether the boost in visits attributable to worker effort increases when 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. We then present this "de-mediated" effect for different levels of supervisor's effort. In the presence of effort complementarities, we would expect the increase in visits generated by the worker to grow in supervisor effort (when supervisor effort increases, the worker exerts more effort and the return to worker effort increases).[33]

In line with this, Panel A of Figure A.4 shows that the effect of worker effort on output increases substantially with supervisor effort (share of visits 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. This is consistent with a strategic complementary between worker effort and the supervisor's in-the-field training. We also find evidence of a strong complementarity between worker effort and the general training provided by the supervisor, while we see no complementarity for the one-to-one meetings (Figure A.4, Panels B and C).

[32]In contrast with Bandiera, Barankay, and Rasul (2007), Table A.11 shows that supervisors are not more likely to target their effort towards health workers who they perceived as highly ranked in terms of performance at baseline, and are also equally likely to target their friends/family members. We also find no heterogeneity in supervisor effort and household visits with respect to the supervisor's span of control (the number of workers per supervisor). This might be explained by limited variation in span of control in our setting.

[33]We focus on the comparison between the worker incentives treatment and the control group since a mediation analysis performed on the other treatments would be confounded by the fact that in those treatments, the supervisor is directly incentivized to exert effort. We follow the steps outlined in Acharya, Blackwell, and Sen (2016) to perform the analysis. First, we regress the number of visits a health worker provides on the worker incentives treatment, the mediator (supervisor's effort), and their interaction. Second, we obtain a de-mediated outcome, defined as the difference between actual visits and the number of visits predicted by the regression model for a given level of the mediator. Third, we run a regression of the de-mediated outcome on the treatment. This regression identifies the CDE of the intervention for a given level of the mediator. We repeat this three-step procedure multiple times, changing the level at which we fix the mediator. In Figure A.4 we report the CDE estimates corresponding to these different levels of the mediator.

5.2 Limited Side Payments

In this section, we document that side payments are limited in our context and present suggestive evidence that this is because of the presence of contractual frictions. To measure side payments, we collected detailed data on monetary and in-kind transfers between the supervisors and the health workers. At endline, we asked all supervisors whether they transferred a portion of their incentive to health workers since baseline. If they did, we then asked each health worker to assess this side payment's value (in-cash or in-kind).[34]

Side payments are generally small and infrequent. In Table 6 column (1), we show that the share of supervisors who make positive side payments increases with the level of the supervisor incentive (1.1% in the control group, 1.6% in T_{worker}, 11.3% in T_{shared}, and 19.4% in T_{supv}), but the large majority of supervisors do not make any transfer across all treatment groups. In column (3), we document that 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}). These amounts are minimal 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}), and that the average supervisor earns 55,280 SLL per month in T_{supv} (resp., 47,097 SLL in T_{shared}) from the incentive payment, as shown in Table 4 (column 7). 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 in T_{worker}; see Table 6 column 4). Overall, this evidence shows that side payments happen in our context, but their frequency and magnitude are minimal.

Why are side payments limited? In the theoretical framework, we discussed two possible explanations: (i) the supervisor may find it optimal to offer a sizable side payment to the worker, but contractual frictions limit her ability to provide these payments; (ii) the optimal side payment may be small or even zero if the value the supervisor attaches to household visits is small compared to that of the health worker.[35] We also pointed out that, as long as we observe positive net payments from the supervisor to the worker, we can disentangle these two potential explanations by looking at the impacts on visits of the different treatments. Under explanation (i), varying the share of the incentives allocated to the worker should generate large differences in visits. Under explanation (ii), we should instead observe the same number of visits for all incentive schemes that generate positive side payments. As shown in Figure 2, visits are far from equalized across treatments despite side payments being positive. This supports explanation (i) and points to the presence of contractual frictions. In what follows, we present two additional pieces of evidence pointing to the presence of these frictions.

[34]This was asked to health workers rather than supervisors to limit recall bias. Supervisors and workers have no incentive to misreport transfers because transfers were neither banned nor encouraged in our setting. See Section 2.2 for details.

[35]If the asymmetry is large enough, the worker may actually find it desirable to offer a payment to their supervisor, but bottom-to-top contractual frictions may prevent them from doing so.

Heterogeneity by supervisor's observability of worker output First, we study the sensitivity of side payments to proxies for "top-to-bottom" contractual frictions. These frictions are more likely present when worker effort or output is not observable to the supervisor since this makes contracts hard to enforce. To measure the observability of output, we asked each supervisor at endline to rank the workers she supervises from the best to the worse in terms of their "overall work as a health worker." We then correlate this *perceived* rank of worker performance/output with the *actual* rank obtained on the basis of the number of household visits completed at endline. The correlation is positive for most supervisors, except for 10% for whom the correlation is negative and who thus have poor observability.[36]

Table A.12 (column 1, Panel A) shows that side payments in both T_{supv} and T_{shared} are inexistent for the supervisors who observe worker output poorly. In contrast, side payments are positive (even though limited) for the remaining supervisors, who can better observe worker output (column 1, Panel B). These heterogeneous effects are robust to controlling for correlates of the observability of output (e.g., the distance between the supervisor and the worker) interacted with the treatment dummies (column 2).

Overall, these results are consistent with side payments being larger when worker output is more observable and hence when contractual frictions are likely weaker. This result provides evidence of the likely importance of contractual frictions in preventing transfers from supervisors to workers. Importantly, output observability appears to be limited for most supervisors in our context, thus contracting difficulties may exist even for supervisors in the upper part of our proxy measure of observability.[37]

Results for workers with better outside options than their supervisor The second piece of evidence pointing to the presence of contractual frictions comes from the analysis of worker-supervisor pairs where the worker has a better outside option than her supervisor, as proxied by the worker having a higher hourly wage from her second job than the supervisor. In these pairs, the worker is likely to exert less effort than the supervisor would find optimal, and we thus expect the supervisor to have strong reasons to offer a sizable side payment to the worker.[38] Yet, Table A.13 shows that, even for these pairs, side payments are limited,

[36]Table A.9 shows that these poorly-informed supervisors tend to live further away from the health workers, while they have the same education, age, and wealth score. They also have the same tenure and work the same number of hours.

[37]We do not observe any heterogeneity in side payments with respect to whether the worker is a friend or family member of the supervisor (Table A.11). This suggests that relational contracts have limited ability to attenuate contractual frictions *within* the organization. This contrasts with papers showing that relational contracts attenuate frictions *across* organizations (McMillan and Woodruff 1999; Macchiavello and Morjaria 2015; Adhvaryu et al. 2020; Macchiavello and Morjaria 2021).

[38]In our theoretical framework, it is natural to think of outside options as a key driver of parameters b_1 and b_2, since outside options change the extent to which agents are concerned about losing their job due to underperformance. Agents with strong outside options will have a low value of parameter b and will, all else equal, exert less effort.

and visits are not equalized across treatments. This points to the presence of additional constraints to side payments, such as contractual frictions.

5.3 Alternative Mechanisms

The previous section provides empirical support for our theoretical framework, in which two-sided incentives outperform one-sided incentives due to *both* effort complementarities and limited side payments. This section provides evidence against alternative mechanisms that are not considered in our model but could explain why two-sided incentives outperform one-sided incentives.

Inequality aversion The ineffectiveness of the one-sided incentives treatments could be explained by an aversion to pay inequality. For example, in the supervisor incentives treatment, the health workers may think it is unfair that the supervisor earns money for services provided by the worker, while the worker herself does not earn anything. Similarly, the supervisor may think that it is unfair that she is not paid any incentive in the worker incentives treatment. If this was the case, then one-sided incentives may *reduce* the effort of the non-incentivized person while raising the effort of the incentivized one. This could, in turn, explain why two-sided incentives outperform one-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 seem to have learned it from the supervisor: in T_{supv} (resp., T_{shared}), only 15% (resp., 20%) of workers reported knowing that their supervisor receives an incentive. Of these, only 2% (resp., 10%) were aware of the exact amount earned by the supervisor, while the rest under-estimated this amount. Second, Table A.14 shows that there is no evidence for health workers in T_{supv} and supervisors in T_{worker} to be less satisfied with their payment, the organization, or job in general compared to the control group. If inequality aversion or fairness concerns were the primary mechanisms driving our results, we would instead expect the non-incentivized health workers in T_{supv} and the non-incentivized supervisors in T_{worker} to be less satisfied than workers and supervisors in the control group.[39] 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.

[39]Workers in all treatments are equally likely to find the work environment competitive or to self-identify with their job: columns (7) and (8) in Table A.14. We also find no differential treatment effects on visits depending on the workers' level of inequality aversion, estimated with a set of hypothetical questions (columns 9 to 11).

Finally, we note that the absence of changes in satisfaction with the job and the organization is also inconsistent with the hypothesis that agents increase effort in the shared incentives treatment due to positive reciprocity. Under this story, any incentive payment would elicit a positive effort response that does not depend on the amount of the incentive paid. However, it is unlikely that reciprocal agents would increase effort but not report higher satisfaction with the organization.

Sharp non-linearities in utility, cost or production functions We would expect two-sided incentives to be more effective than one-sided incentives if the returns to offering a piece rate above 1,000 SLL were low due to strongly diminishing returns to individual-level incentives in agents' utility, or strong non-linearities in the cost or production functions. In this section, we provide evidence against this.

Intuitively, in the absence of effort complementarities, shared incentives could outperform the one-sided incentives if there was a sharp discontinuity in agents' utility function for incentive payments above 1,000 SLL. In this case, one-sided incentives would fail to motivate either of the two agents substantially more than the shared incentives treatment. Moreover, such a discontinuity in the utility function of both supervisors and workers would explain why the supervisors provide the same amount of effort in T_{supv} and T_{shared} (as shown in Table 5), and why the shared incentives treatment leads to more visits than the other treatments.

To investigate this, Panel A of Figure A.5 displays non-parametric plots of the treatment effects on output and supervisor effort by worker's and supervisor's wealth score (a proxy for background utility). In the presence of strong non-linearities in utility, treatment effects would decline steeply in wealth, at least for some range of the wealth distribution. The figure shows instead that the treatment effects are reasonably stable over the whole wealth score distribution (if anything, supervisor effort appears to increase with supervisor wealth slightly). This is not surprising since even the wealthiest workers and supervisors in our sample are pretty poor, and doubling the incentive from 1,000 to 2,000 SLL can substantially increase their income.[40]

Alternatively, there may be a similar discontinuity in the cost function. Here, the marginal cost of effort would need to rise sharply at the effort level exerted by the agents in the shared incentives treatment (1,000 SLL incentive).[41] This is an implausible 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.5 presents non-parametric plots of impacts on output and supervisor effort over the distribution of household-worker distance (a proxy for the worker's cost of

[40] Health workers 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 a 15 minutes visit).

[41] This could be the case if, for example, the distribution of household-worker distance was bimodal, with some households living very close to the health workers, and others living so far that an incentive between 1,000 and 2,000 SLL would not cover the cost of reaching them.

visiting a household) and worker-supervisor distance (a proxy for the supervisor's cost of training/monitoring a health worker). Again, we do not find evidence of any discontinuity.

The last possibility is that the results are explained by a discontinuity in the production function, such that the return to the worker's effort increases with supervisor's effort only up to a threshold that coincides with the effort level exerted by the supervisor in the shared incentives treatment. In contrast with this story, Figure A.6 shows that the non-parametric relationship between realized visits and supervisor effort is positive and nearly linear.

6 Structural Model

In this section, we use the exogenous variation generated by the interventions to structurally estimate the model presented in Section 3, allowing for worker and supervisor-specific costs and benefits (as in Appendix D.7). 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 following parameters of the model: complementarity γ, the two costs of effort c_1 and c_2, the baseline incentives b_1 and b_2, the production function parameter α and the contractual friction z. We calibrate z with a regression exercise that is described below. We jointly identify the remaining six parameters using eight empirical moments, i.e., the means of output (household visits) and supervisor effort in the four experimental conditions.[42] Intuitively, the moments capturing supervisor effort are informative about the cost and benefit parameters of the supervisor. Conditional on those parameters, the moments capturing output are informative about the cost and benefit of the worker, the complementarity of effort, and the parameter α.

We calibrate contractual frictions by using data on side payments. In particular, our model shows that $s = k - \frac{z+1}{2z}mp$. This suggests that the slope of a regression line of side

[42] In our model, there is no individual heterogeneity, so we only rely on empirical moments capturing mean outcomes. The specific 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 health worker (as reported in Table A.2, column 2). The measure of supervisor effort we use is the fraction of households that received at least one visit where the health worker was accompanied by the supervisor (as reported in Table 5, column 1). Accompanied visits are a form of training that presumably takes place towards the beginning of the experiment. Using accompanied visits as a measure of supervisor effort is thus compatible with the model assumption that effort choices are sequential. We do not include worker knowledge (Table 5, column 4) as a measure of effort since this variable reflects both supervisor and worker effort. We also do not have a distinct measure of worker effort, since worker effort is hard to distinguish empirically from output (an ideal measure of worker effort would capture the number of attempted household visits and the cost of these attempts, which are very hard to measure in a survey). We thus do not use any moment describing worker effort.

payments s on mp — the product of the piece rate times the share of the piece rate offered to the worker — is informative of the size of contractual frictions z. When there are no frictions ($z = 1$), the slope of the regression line is 1. As frictions grow, the slope drops below 1 and approaches 0.5 from above. This result is intuitive: the stronger the frictions, the less responsive to p the side payment.[43]

To make the model more realistic, we introduce under-reporting by assuming that, for each completed visit, a shock that prevents the worker from reporting the visit occurs with probability $(1 - q)$. This shock occurs after efforts have been exerted, and so its realization is not factored into effort decisions. However, agents know that a shock may occur and hence expect the value of the piece rate to be $m * q$. Except for this change in the expected value of the piece rate, the model remains unchanged. In our headline results, we assume conservatively that supervisors form an expectation about q using the average reporting rate across the experimental conditions. We then show robustness to assuming instead that supervisors realize that q varies with p, and thus that they form separate expectations about the reporting rate in each treatment.

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:

$$\hat{\boldsymbol{\theta}} = \min_{\boldsymbol{\theta}} \left[\boldsymbol{m}_S(\boldsymbol{\theta}) - \boldsymbol{m}\right]' \cdot J(\boldsymbol{m})^{-1} \cdot \left[\boldsymbol{m}_S(\boldsymbol{\theta}) - \boldsymbol{m}\right]. \tag{6}$$

$J(\boldsymbol{m})$ is a diagonal matrix that contains the variance of each moment, ensuring that more precisely estimated moments get a greater weight in estimation. We calculate $J(\boldsymbol{m})$ using a bootstrap with 500 replications. Table 7 presents our main structural results, and Table 8 describes the empirical fit of the simulated moments.

The estimated model tightly fits the empirical moments: it matches both the moments related to supervisor effort and those related to household visits. Crucially, the estimated model is able to reproduce the key result that visits are maximized by the shared incentives treatment.

In contrast, a version of the model based on a production function where efforts are not strategic complements ($y = \alpha e_1 + \beta e_2$) has a much worse fit (see Tables A.15 and A.16). This model version wrongly predicts that worker incentives generate the largest increase in visits, and the value of its minimized loss function is about seven times larger than that of the model that features effort complementarities. Further, in line with the assumption of our

[43] We note two features of this calibration exercise. First, this exercise does not rely on the information on the absolute level of side payments which is contained in the intercept of the regression line ($k = \frac{b_2+m+zb_1}{2z}$), as this is likely to be observed with noise due to misreporting and poor memory. This is also a key reason we calibrate the friction before the main structural estimation procedure. Second, we proxy s, the side payment offered, which we do not observe in the data, with the side payment paid, which we observe in the data.

headline model, the estimate of β — the direct impact of supervisor effort on the number of visits carried out by the health worker — is close to zero. Thus, overall, the findings from this second estimation exercise give support to our original modeling assumptions.

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 (Table 7).

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 of household visits generated by a unit of worker effort is 82% larger when the supervisor exerts the control level of effort, and 116% larger when the supervisor exerts the shared incentives level of effort. Supervisor effort thus plays a crucial role in enabling the worker to carry out household visits, resulting in a strong strategic complementarity between the efforts of the two agents.

The calibrated value of parameter z implies that side payments are 45 percent more costly due to contractual frictions. This constitutes a solid disincentive to offering side transfers, though we are unaware of other estimates of contractual frictions that we can use as a benchmark. A further disincentive against side transfers comes from the fact that the baseline incentive of the supervisor to exert effort (b_2) is lower than that of the worker (b_1). This is not surprising since her role is probably harder to monitor and incentivize. Low supervisor motivation also suggests that reforms that target contractual frictions without also addressing supervisor motivation may backfire, as the supervisor may not necessarily use the greater ability to influence the worker in a way that is consistent with the organization's objectives.

We also find that the overall effort provision from the supervisor is relatively low, as evidenced by the fact that the *total* cost of her effort is lower than that of the worker. The latter is driven by the supervisor having a higher cost of effort parameter than the worker ($c_2 > c_1$) and lower baseline incentive ($b_2 < b_1$).[44] Overall, this indicates that interventions that fail to incentivize the supervisor may be 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.

The results are robust to changing assumptions about the expected reporting rate, as shown in Tables A.17 and A.18. When we assume that supervisors have correct expectations about the reporting rate in each treatment group, we estimate very similar levels of effort complementarity (the worker's marginal product increases to 83% in the control group and

[44] We do not report the unit costs of effort c_2 and c_1 in Table 7 due to the difficulty in interpreting the parameters. We report instead the *total* costs of worker and supervisor effort, as measured by the unit cost of effort times the average effort exerted by the agent in the control group.

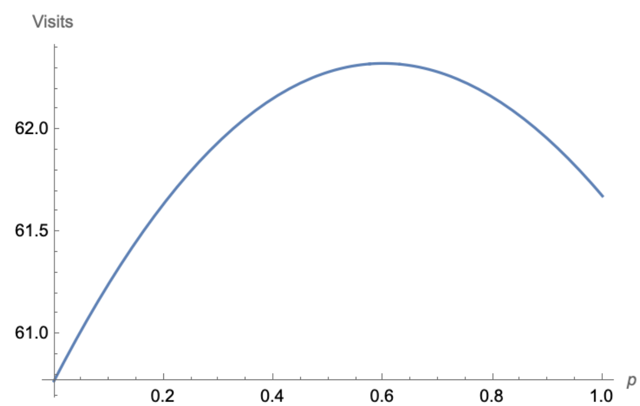
117% in the group incentive condition) and an extremely high contractual friction ($z = 11.74$). This confirms that our core results on the importance of effort complementarity and contractual frictions do not depend on the specific assumption we make on reporting rate expectations.

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 3, we show that, to maximize household visits, the worker should be offered 59% of the overall incentive, which is very close to the equal share 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.[45]

Figure 3: Optimal Incentive p^*

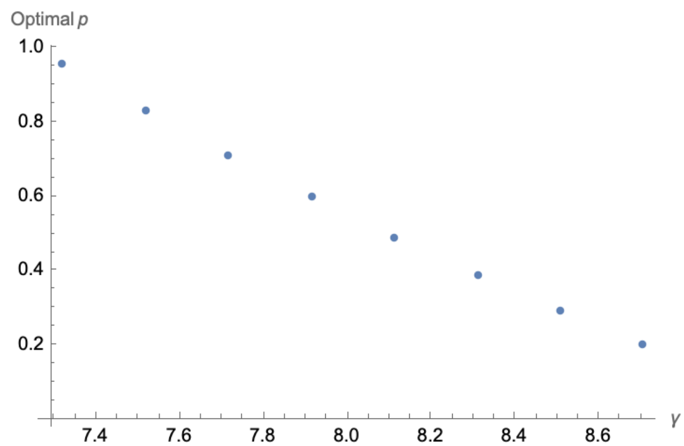


This result, however, depends strongly on the strength of the complementarity between worker and supervisor effort. We illustrate this point with our second counterfactual experiment in Figure 4. Here, we plot the optimal share of the incentive offered to the worker (p^*) for different levels of complementarity. A key result that emerges from this analysis is that, as the complementarity parameter shrinks, the optimal share of the incentive allocated to the worker increases substantially. Quantitatively, if the complementarity parameter was 10 percent *lower* than what we estimate, the optimal incentive scheme would give 80 percent of the piece rate to the worker. If the complementarity parameter was instead 10 percent *higher*

[45]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.

than what we estimate, the optimal incentive scheme would give 60 percent of the piece rate to the supervisor.

Figure 4: Optimal Incentive p^* by Complementarity γ



Thus, these results suggest that in organizations in which effort complementary is weaker than in our settings — e.g., settings in which the role of the supervisor is limited to monitoring, distributing tasks, or making personnel decisions, but not training and advising workers — the optimal split is one that allocates significantly more to the worker. And in organizations where effort complementarity is stronger — e.g., organizations where supervisors are closely involved in the production — the optimal incentive scheme allocates the largest share of the piece rate to the supervisor.

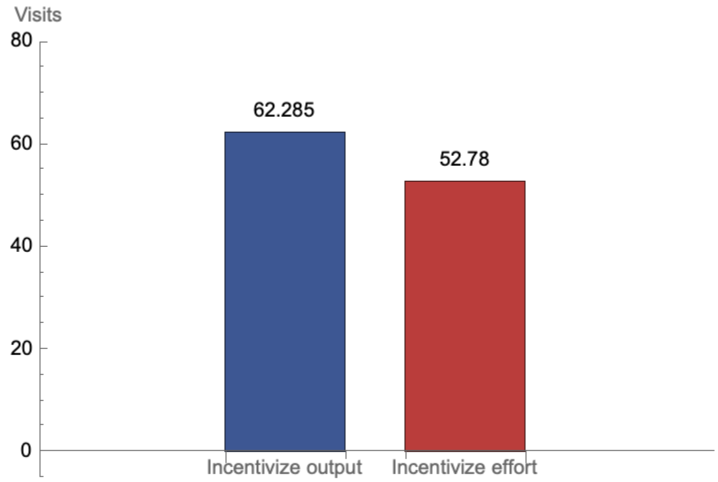
Our final key result highlights that tying incentives to joint output is more effective than directly incentivizing effort (e.g., incentivizing supervisors on the amount of supervision and training, and incentivizing health workers on the number of times they attempt to approach a household, regardless of whether this results in a visit or not). In Figure 5, 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 18% more visits that optimally incentivizing effort, for the same cost. This is because, when efforts are highly complementary, output incentives implicitly help agents internalize their effort's positive external effect on the other player. This makes output incentives particularly effective.[47]

[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. In the effort incentive case, since effort can be observed, the payoff to the worker becomes $\pi_1 = e_1 * (b_1 + mp)$ and the payoff to the supervisor $\pi_2 = e_2 * (b_2 + (1-m) * p)$. In this model, the supervisor always offers zero side transfer since her reward only depends on her own effort.

[47]To see the intuition behind this result, consider a simple case where $\alpha = 1$, $b_1 = b_2 = 0$, $c_1 = c_2 = c$, and

Figure 5: An Alternative Policy that Targets Effort



7 Conclusion

This paper provides novel evidence on the optimal structure of performance incentives in a multi-layered organization. In a field experiment with a large community-health program in Sierra Leone, we show that output is highly sensitive to the allocation of incentives across the hierarchy. Sharing incentives equally between frontline workers and their supervisors generates an increase in health visits that is 61% larger than the increase caused by offering the entire incentive to either of the layers of the organization. These findings are inconsistent with a Coasian view of organizations. They also contradict the priors of most experts who forecasted our results on the Social Science Prediction Platform.[48] And they call in question the common practice in many public sector organizations (including community health programs) around the world to only incentivize frontline workers.

The view of organizations that emerges from our results sees the coexistence of contractual frictions and effort complementarities as a central determinant of organizational performance. This has a number of important policy implications. First, side transfers within the organization are limited and this makes the allocation of incentives across the hierarchy a key lever to

there are no side payments. Let us focus on the worker problem. The principal can either offer them output incentives of the value of $p * m_0$ for each unit of output, or effort incentives of the value of $p * m_e$ for each unit of effort. Under output incentives, the objective function of the worker is $\pi_1 = (e_1 + \gamma e_1 e_2)m_o p - ce_1^2$ and their optimal effort choice is $e_1 = \frac{(1+\gamma e_2)m_o p}{2c}$. Under effort incentives, the objective function of the worker is $\pi_1 = e_1 m_e p - ce_1^2$ and their optimal effort choice is $e_1 = \frac{m_e p}{2c}$. From this, it is apparent that output incentives motivate the worker to choose effort taking into account the strategic complementarity with the supervisor, while effort incentives induce the worker to pick a level of effort that is independent of the strategic complementarity.

[48]Before releasing the results of the field experiment, we invited social scientists to forecast our results on the online Social Science Prediction Platform. Survey participants were told about our context — i.e., the organization, the role of health workers and their supervisors — and were then asked to forecast which of our three treatments they expected to increase output the most. 52% of the respondents indicated the one-sided worker incentives as the most successful ones vs. 28% for the shared incentives and 4% for the supervisor incentives. See Appendix E for more details about the prediction survey.

boost performance. Second, worker-supervisor effort complementarities are strong. Thus, it is optimal to ensure both layers of the organization are properly incentivized. Third, supervisors are poorly motivated. Hence, policies that try to improve Coasian bargaining within the firm may backfire, as supervisors are likely to use side transfers to pursue objectives that are inconsistent with those of the organization.

One final consideration is that, to introduce an incentive scheme such as the one considered in this paper, organizations need to be able to measure output reliably. In our setting, we pay the incentive based on workers' self-reports, while performing extensive back-checks to prevent over-reporting. As digital technologies improve, the costs of monitoring worker self-reports will likely decrease, enabling more organizations to set up incentive schemes like ours (Muralidharan et al. 2021; Dodge et al. 2021; Adhvaryu, Nyshadham, and Tamayo 2022). Further, better information on frontline workers' output can in principle also affect the ability of supervisors and workers to enter into side contracts. Studying how to best allocate access to this information represents a fruitful avenue for future research.

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):512–529.

Adhvaryu, Achyuta, Jean-François Gauthier, Anant Nyshadham, and Jorge Tamayo. 2020. "Absenteeism, Productivity, and Relational Contracts Inside the Firm." NBER Working Paper No. 29581.

Adhvaryu, Achyuta, Anant Nyshadham, and Jorge Tamayo. 2022. "An Anatomy of Performance Monitoring." Working Paper.

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.

Asiedu, Edward, Dean Karlan, Monica P Lambon-Quayefio, and Christopher R Udry. 2021. "A Call for Structured Ethics Appendices in Social Science Papers." *Proceedings of the National Academy of Sciences* 118 (29):e2024570118.

Atkin, David, Azam Chaudhry, Shamyla Chaudry, Amit K Khandelwal, and Eric Verhoogen. 2017. "Organizational Barriers to Technology Adoption: Evidence from Soccer-ball Producers in Pakistan." *The Quarterly Journal of Economics* 132 (3):1101–1164.

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

———. 2011. "Field experiments with firms." *Journal of Economic Perspectives* 25 (3):63–82.

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* 136 (4):2195–2242.

Bandiera, Oriana, Michael Callen, Katherine Casey, Eliana La, Camille Landais Ferrara, and Matthieu Teachout. 2019. "State effectiveness." *International Growth Centre Evidence Paper* .

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 LB Soucat, Jennifer Sturdy, and Christel MJ 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.

Benjamini, Yoav, Abba M. Krieger, and Daniel Yekutieli. 2006. "Adaptive Linear Step-up Procedures that Control the False Discovery Rate." *Biometrika* 93 (3):491–507.

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

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2013. "Does Management Matter? Evidence from India." *The Quarterly Journal of Economics* 128 (1):1–51.

Bloom, Nicholas and John Van Reenen. 2007. "Measuring and Explaining Management Practices across Firms and Countries." *The Quarterly Journal of Economics* 122 (4):1351–1408.

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

Bruhn, Miriam, Dean Karlan, and Antoinette Schoar. 2018. "The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico." *Journal of Political Economy* 126 (2):635–687.

Brynjolfsson, Erik and Paul Milgrom. 2013. *Complementarity in Organizations*, chap. 1. Princeton University Press Princeton, NJ, 11–55.

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

Cai, Jing and Shing-Yi Wang. 2020. "Improving Management through Worker Evaluations: Evidence from Auto Manufacturing." NBER Working Paper No. 27680.

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.

Callen, Michael, Saad Gulzar, Syed Ali Hasanain, and Muhammad Yasir Khan. 2016. "The Political Economy of Public Sector Absence: Experimental Evidence from Pakistan." NBER Working Paper No. 22340.

Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez. 2012. "Inequality at Work: The Effect of Peer Salaries on Job Satisfaction." *American Economic Review* 102 (6):2981–3003.

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.

Casey, Katherine and Rachel Glennerster. 2016. "Reconciliation in Sierra Leone." *Science* 352 (6287):766–767.

Celhay, Pablo A, Paul J Gertler, Paula Giovagnoli, and Christel Vermeersch. 2019. "Long-Run Effects of Temporary Incentives on Medical Care Productivity." *American Economic Journal: Applied Economics* 11 (3):92–127.

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.

———. 1960. "The Problem of Social Cost." *The Journal of Law and Economics* 3:1–44.

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

———. 2022. "How Much Does Your Boss Make? The Effects of Salary Comparisons." *Journal of Political Economy* 130 (3):766–822.

Dal Bó, Ernesto, Frederico Finan, Nicholas Li, and Laura Schechter. 2021. "Information Technology and Government Decentralization: Experimental Evidence from Paraguay." *Econometrica* 89 (2):677–701.

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. 2021. "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.

Dupas, Pascaline and Edward Miguel. 2017. *Impacts and Determinants of Health Levels in Low-Income Countries*. 2. North Holland, 3–94.

Espinosa, Miguel and Christopher T Stanton. 2022. "Training, Communications Patterns, and Spillovers Inside Organizations." *NBER working paper* .

Finan, Frederico, Benjamin A. Olken, and Rohini Pande. 2017. "The Personnel Economics of the Developing State." In *Handbook of Economic Field Experiments*, vol. 2. Elsevier, North-Holland, 467–514.

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.

Garicano, Luis and Luis Rayo. 2016. "Why Organizations Fail: Models and Cases." *Journal of Economic Literature* 54 (1):137–92.

Geng, Tong. 2018. "The Complementarity of Incentive Policies in Education: Evidence from New York City." Working Paper.

Gibbons, Robert. 1999. "Taking Coase Seriously." *Administrative Science Quarterly* 44 (1):145–157.

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

Gibbons, Robert and John Roberts. 2012. *2. Economic Theories of Incentives in Organizations*. Princeton: Princeton University Press, 56–99.

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.

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.

Kala, Namrata. 2019. "The Impacts of Managerial Autonomy on Firm Outcomes." Working Paper.

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

Khan, Muhammad Yasir. 2021. "Mission Motivation and Public Sector Performance: Experimental Evidence from Pakistan." Working Paper.

Kremer, Michael. 1993. "The O-ring Theory of Economic Development." *The Quarterly Journal of Economics* 108 (3):551–575.

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

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

Lee, Robin S., Michael D. Whinston, and Ali Yurukoglu. 2021. "Chapter 9 - Structural Empirical Analysis of Contracting in Vertical Markets." *Handbook of Industrial Organization* 4 (1):673–742.

León-Ciliotta, Gianmarco, Dijana Zejcirovic, and Fernando Fernandez. 2022. "Policy-Making, Trust and the Demand for Public Services: Evidence from a Nationwide Family Planning Program." Working Paper.

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. 2021. "Relational Contracts and Development." *Annual Review of Economics* 14.

Macchiavello, Rocco, Andreas Menzel, Atonu Rabbani, and Christopher Woodruff. 2020. "Challenges of Change: An Experiment Promoting Women to Managerial Roles in the Bangladeshi Garment Sector." Working Paper.

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* 136 (2):1089–1143.

Martinez-Bravo, Monica and Andreas Stegmann. 2022. "In Vaccines We Trust? The Effects of the CIA's Vaccine Ruse on Immunization in Pakistan." *Journal of the European Economic Association* 20 (1):150–186.

McKenzie, David and Christopher Woodruff. 2017. “Business Practices in Small Firms in Developing Countries.” *Management Science* 63 (9):2967–2981.

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

Meyer-Sahling, Jan-Hinrik, Christian Schuster, and Kim Sass Mikkelsen. 2018. “Civil Service Management in Developing Countries: What Works? Evidence from a Survey with 23,000 Civil Servants in Africa, Asia, East Europe and Latin America.” Working Paper.

Milgrom, Paul and John Roberts. 1995. “Complementarities and Fit Strategy, Structure, and Organizational Change in Manufacturing.” *Journal of Accounting and Economics* 19 (2-3):179–208.

Mohanan, Manoj, Katherine Donato, Grant Miller, Yulya Truskinovsky, and Marcos Vera-Hernández. 2021. “Different Strokes for Different Folks? Experimental Evidence on the Effectiveness of Input and Output Incentive Contracts for Health Care Providers with Varying Skills.” *American Economic Journal: Applied Economics* 13 (4):34–69.

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.

Perry, Henry B., editor. 2020. *Health for the People: National Community Health Worker Programs from Afghanistan to Zimbabwe.*

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.

Romano, Joseph P. and Michael Wolf. 2016. “Efficient Computation of Adjusted P-values for Resampling-based Stepdown Multiple Testing.” *Statistics and Probability Letters* 113 (C):38–40.

Roser, Max. 2021. “Child Mortality: An Everyday Tragedy of Enormous Scale that We Can Make Progress Against.”

Shapira, Gil, Ina Kalisa, Jeanine Condo, James Humuza, Cathy Mugeni, Denis Nkunda, and Jeanette Walldorf. 2017. “Effects of Performance Incentives for Community Health Worker Cooperatives in Rwanda.” *World Bank Policy Research Working Paper* (8059).

———. 2018. “Going Beyond Incentivizing Formal Health Providers: Evidence from the Rwanda Community Performance-Based Financing Program.” *Health Economics* 27 (12):2087–2106.

Singh, Prakarsh and William A Masters. 2017. “Impact of Caregiver Incentives on Child Health: Evidence from an Experiment with Anganwadi Workers in India.” *Journal of Health Economics* 55:219–231.

Singh, Prakarsh and Sandip Mitra. 2017. “Incentives, Information and Malnutrition: Evidence from an Experiment in India.” *European Economic Review* 93:24–46.

Soeters, Robert and Fred Griffiths. 2003. “Improving Government Health Services through Contract Management: A Case from Cambodia.” *Health Policy and Planning* 18 (1):74–83.

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.

Wade, Robert. 1985. "The Market for Public Office: Why the Indian state is not Better at Development." *World Development* 13 (4):467–497.

———. 1992. "How to make Street Level Bureaucracies Work Better: India and Korea." *ids bulletin* 23 (4):51–54.

Wagner, Zachary, John Bosco Asiimwe, and David I Levine. 2020. "When Financial Incentives Backfire: Evidence from a Community Health Worker Experiment in Uganda." *Journal of Development Economics* 144:102437.

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."

———. 2021. *Tracking Universal Health Coverage: 2021 Global Monitoring Report.* World Health Organization.

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 (p-value)
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	
A. Characteristics of the supervisors (N=372)											
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)	37.84	8.856	37.91	9.329	37.46	7.869	36.85	8.690	39.13	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 = {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 responsible for	7.984	2.861	7.559	2.799	8.355	2.831	8.011	2.902	8.011	2.899	0.289
B. Characteristics of the health workers (N=2,970)											
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)	37.12	11.47	35.95	11.14	37.79	11.72	37.48	11.72	37.17	11.21	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 = {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 responsible for	55.19	78.59	62.72	120.2	54.08	62.92	53.16	56.37	51.26	60.24	0.375
Distance to supervisor (in km)	3.415	2.945	3.267	2.887	3.815	3.610	3.107	2.141	3.447	2.895	0.190
C. Characteristics of the female household respondent, aggregated to village level (N=2,970)											
Age (in years)	27.79	4.576	28.13	4.741	27.69	4.410	27.56	4.572	27.84	4.586	0.266
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
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
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
Distance to health worker (in km)	1.433	2.630	1.189	2.124	1.591	2.575	1.438	2.894	1.483	2.785	0.506
D. Characteristics of the village (N=372)											
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
Phone network available	0.838	0.368	0.824	0.381	0.845	0.362	0.862	0.345	0.821	0.384	0.490
E. Services provided by the health facility per month (N=372)											
Number of pregnant women services	47.71	45.80	43.76	42.98	51.22	57.17	48.43	37.10	47.34	44.10	0.769
Number of institutional births	13.44	8.266	12.58	6.087	13.67	7.814	13.38	6.845	14.13	11.33	0.509
Number of fully immunized infants	11.41	10.75	10.88	10.10	11.92	12.58	10.67	7.060	12.14	12.39	0.684
Number of malaria cases treated	45.89	32.03	42.29	26.88	51.44	38.88	46.63	31.52	43.23	29.38	0.286
Number of diarrhea cases treated	20.45	17.03	19.62	13.25	21.81	19.47	20.58	21.58	19.76	12.02	0.809

Notes: Each row states the sample mean and standard deviation of a variable, and by treatment group. The last column 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 using standard errors clustered at the PHU level in worker/village level regressions or robust standard errors in PHU/supervisor level regressions. Data source is the supervisor survey in Panel A, the health worker survey in Panel B, the household survey (aggregated to health worker/village level) in Panel C, the health worker's leaflet in Panel D, the facility admin data in Panel E.

Table 2: Household Visits

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	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	Average visit length *(conditional on at least one visit)*	Average number of health topics discussed per visit	Average number of health topics discussed per visit *(conditional on at least one visit)*	% households who trust the health worker as a health provider
Worker incentives	2.090***	0.072***	0.250***	2.024**	0.603	0.164	0.001	0.037
	(0.558)	(0.025)	(0.094)	(0.941)	(1.423)	(0.127)	(0.174)	(0.023)
Supervisor incentives	2.145***	0.082***	0.325***	1.933**	-0.070	0.173	0.003	0.031
	(0.501)	(0.025)	(0.100)	(0.926)	(1.409)	(0.130)	(0.181)	(0.024)
Shared incentives	3.356***	0.127***	0.565***	4.134***	1.590	0.528***	0.425**	0.071***
	(0.492)	(0.023)	(0.092)	(0.927)	(1.290)	(0.134)	(0.185)	(0.024)
Unit	Worker	Worker	Worker	Worker	Worker	Worker	Worker	Worker
Observations	2,926	2,926	2,926	2,926	1,803	2,926	1,803	2,926
Mean dep. var.	7.296	0.709	1.745	14.388	23.404	2.248	2.922	0.745
Mean dep. var. in Control	5.334	0.637	1.448	12.324	22.736	2.015	2.782	0.707
p-value Worker = Supervisor	0.932	0.710	0.492	0.927	0.635	0.946	0.990	0.808
p-value Supervisor = Shared	0.038	0.060	0.026	0.024	0.196	0.017	0.033	0.102
p-value Worker = Shared	0.046	0.026	0.002	0.033	0.451	0.013	0.026	0.147

Notes: Data source is the household survey, aggregated to health worker level. Cols. (5) and (7) restrict the sample to households receiving at least one visit in the past 6 months. All regressions include stratification variables. Standard errors are 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						Diseases incidence			
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)	% children under-5 who had fever	% children under-5 who had cough	% children 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: Data source is the household survey, aggregated to health worker level. The index in col. (1) [resp., col. (7)] estimates an equally weighted average of the z-scores of variables in cols. (2)-(6) [resp., cols. (8)-(10)]. The sample in cols. (1)-(6) is restricted to households with a woman who gave birth in the past year. All regressions include stratification variables. Standard errors are clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Reporting and Incentives Payouts

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Reporting				Cost of the intervention (in 1,000 SLL)				
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]	Incentives payment per health worker	Total incentives payment across all health workers	Incentives payment per supervisor	Total incentives payments in the PHU (cols. 6+7)	Number of visits per worker for each 2,000 SLL spent
Worker incentives	7.182***	0.222***	0.046**	-0.002	16.054***	132.903***	-1.310	131.593***	-
	(0.759)	(0.043)	(0.018)	(0.024)	(1.352)	(12.000)	(1.832)	(11.962)	-
Supervisor incentives	2.181***	0.058*	-0.006	0.041*	-0.047	-1.172	55.280***	54.108***	3.126
	(0.586)	(0.034)	(0.015)	(0.021)	(0.185)	(2.631)	(8.095)	(8.377)	(4.151)
Shared incentives	4.757***	0.091**	0.009	0.046**	5.584***	46.857***	47.097***	93.953***	9.574**
	(0.829)	(0.036)	(0.015)	(0.021)	(0.691)	(6.680)	(6.534)	(13.164)	(4.826)
Unit	Worker	Worker	Worker	Worker	PHU	PHU	PHU	PHU	PHU
Observations	2,970	2,624	2,926	2,926	372	372	372	372	279
Mean dep. var.	5.213	0.177	0.069	0.862	5.431	45.535	25.858	71.392	10.787
Mean dep. var. in Control	1.525	0.078	0.055	0.840	0.000	0.000	0.000	0.000	
Mean dep. var. in Worker inc.									6.508
p-value Worker = Supervisor	<0.001	<0.001	0.004	0.059	<0.001	<0.001	<0.001	<0.001	-
p-value Supervisor = Shared	0.008	0.398	0.333	0.802	<0.001	<0.001	0.428	0.009	0.223
p-value Worker = Shared	0.026	0.005	0.050	0.035	<0.001	<0.001	<0.001	0.032	-

Notes: The reporting rate in col. (2) is measured as the number of reports (from the SMS admin data; col. 1) divided by the total number of visits per month (from the household survey). The latter is measured as the total number of household visits provided by the health worker to sampled households per month divided by the share of households in the village sampled. Under and over reporting in cols. (3)-(4) are indicators for whether the number of reports (from the SMS admin data) is lower or higher than the number of household visits (from the household survey). The outcome variable in col. (9) calculates the number of visits performed by the average worker in the PHU minus the mean number of visits in the control (from the household survey), divided by the total incentive payout in the PHU (from the payroll data; col. 8). We winsorize the top and bottom 1% of the outcome variable due to the presence of outliers and inpute the maximum value of the outcome variable for the few PHUs in which the total incentives payout is zero. In col. (9), the control group is excluded and the omitted group is the worker incentives treatment. All regressions include stratification variables. Standard errors are clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Supervisor Effort

	(1)	(2)	(3)	(4)	(5)
Dep. Var.	% accompanied household visits	% health workers visited in the past 6 months	Supervisor organized training in the past month = {0, 1}	Difference in workers' health knowledge between baseline and endline	Supervisor helped health worker with reporting = {0, 1}
Worker incentives	0.030	-0.050	0.004	0.158	0.002
	(0.022)	(0.137)	(0.005)	(0.116)	(0.024)
Supervisor incentives	0.057**	-0.041	0.006	0.063	0.017
	(0.023)	(0.137)	(0.005)	(0.113)	(0.024)
Shared incentives	0.062***	-0.043	0.003	0.266**	-0.003
	(0.021)	(0.139)	(0.005)	(0.121)	(0.022)
Unit	Worker	Worker	Worker	Worker	Worker
Observations	2,919	2,833	2,864	2,927	2,927
Mean dep. var.	0.204	1.375	0.994	0.313	0.091
Mean dep. var. in Control	0.164	1.417	0.991	0.433	0.086
p-value Worker = Supervisor	0.293	0.950	0.443	0.372	0.499
p-value Supervisor = Shared	0.846	0.987	0.463	0.074	0.341
p-value Worker = Shared	0.181	0.963	0.916	0.350	0.819

Notes: Data source is the household survey in col. (1), aggregated to worker level. Data source is the health worker survey in cols. (2)-(5). The outcome variable in col. (4) calculates the difference in the health knowledge of the health worker between baseline and endline. The health knowledge is measured through a test that counts the number of correct or almost correct answers out of 5 questions asked to the health worker about when and whether to refer a child under-5 to the health clinic. All regressions include stratification variables. Standard errors are clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: 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: Data source is the health worker and supervisor survey. All regressions include stratification variables. Standard errors clustered at the PHU level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 7: Parameter Estimates

	(1)
Complementarity γ	7.9
	(3.3)
Worker baseline incentive b_1	23.5
	(13.7)
Supervisor baseline incentive b_2	16.5
	(12.6)
α	1.5
	(0.7)
Calibrated friction z	1.45
Δ in marginal product of worker effort (shared incentive)	116%
Δ in marginal product of worker effort (control)	82%
Total worker cost of effort (control)	474.8
Total supervisor cost of effort (control)	306.6

Notes: The first panel of the table shows parameter estimates obtained using minimum distance estimation. We use eight empirical moments: supervisor effort in each one of the four treatments, and number of visits per month in each one of the four experimental groups. Supervisor effort is proxied by the proportion of households that receive a visit where the worker is accompanied by the supervisor. Bootstrapped standard errors are reported in parenthesis (we bootstrap the estimation 500 times and truncate the estimated coefficients at the 99th percentile of the distribution). The second panel first shows the calibrated value of contractual frictions. Second, it shows some quantities implied by the parameter estimates. To calculate the change in the marginal product of worker effort we take the derivative of the production function with respect to worker effort (i) with $\gamma = 7.9$ and supervisor effort fixed at the level indicated in parenthesis, and (ii) with $\gamma = 0$. To calculate the total cost of an agent effort we multiply the unit cost of effort by the average effort exerted by the agent in the control group. Costs are expressed in thousand SLL. We report the total cost of efforts rather than the worker and supervisor cost of effort parameters (c_1 and c_2) because the latter are hard to interpret.

Table 8: Moment Fit

Moments	Targeted Real	Simulated
Supervisor effort in worker incentives group	0.198	0.205
Supervisor effort in supervisor incentives group	0.225	0.231
Supervisor effort in shared incentives group	0.228	0.221
Supervisor effort in control group	0.164	0.156
Output in worker incentives group	59.679	61.679
Output in supervisor incentives group	58.896	60.773
Output in shared incentives group	66.895	62.285
Output in control group	41.040	41.156
Value loss function	6.6	

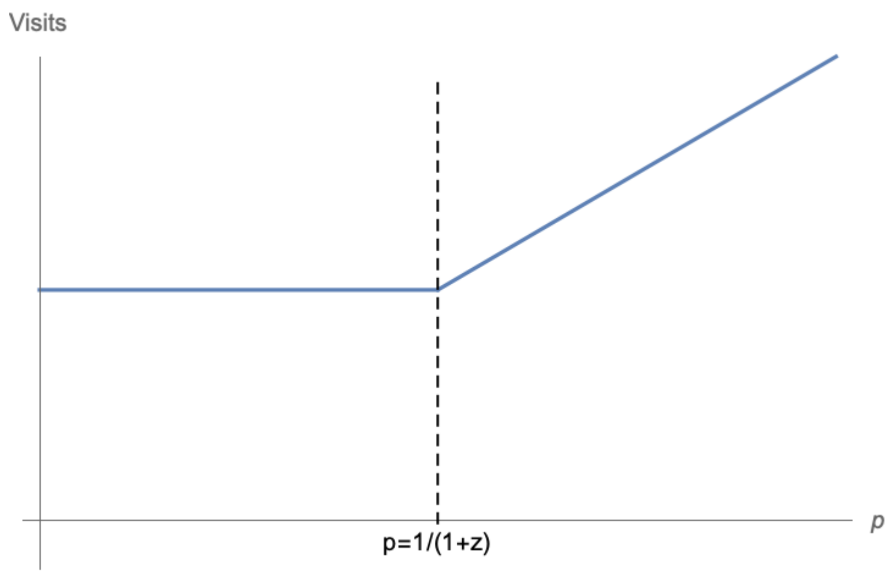
Notes: The table shows the targeted empirical moments used for minimum distance estimation as well as the simulated moments.

Online Appendix (For Online Publication Only)

A Appendix Figures and Tables

Figure A.1: Optimal Incentives (Continued)

(a) Weak Effort Complementarities and No Contractual Frictions ($\gamma < t$, $z = 1$)



(b) Strong Effort Complementarities and No Contractual Frictions ($\gamma \geq t$, $z = 1$)

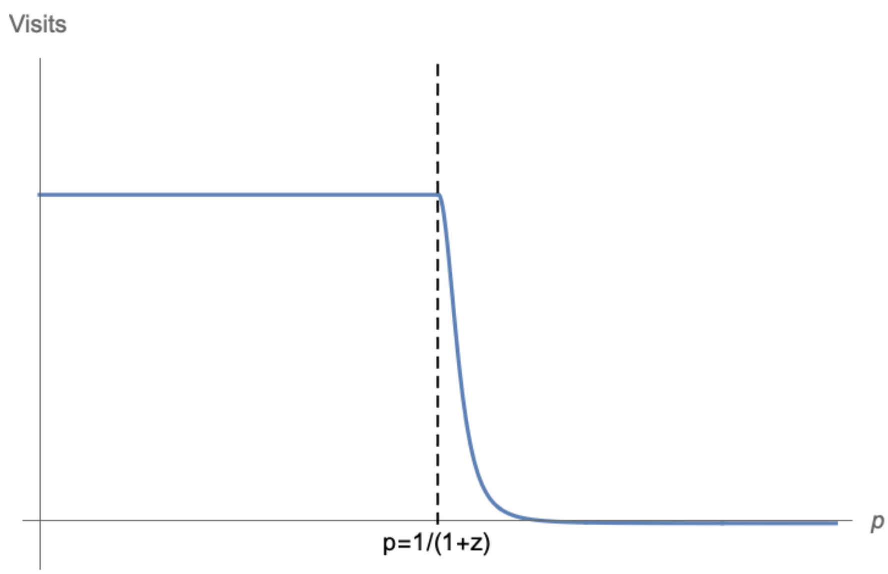
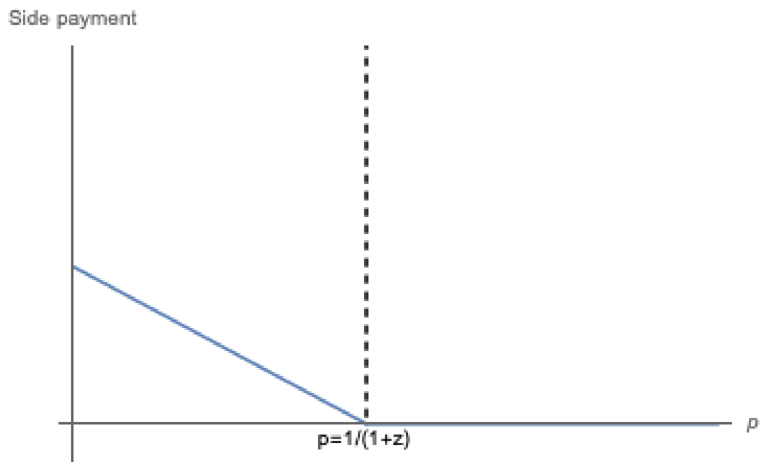
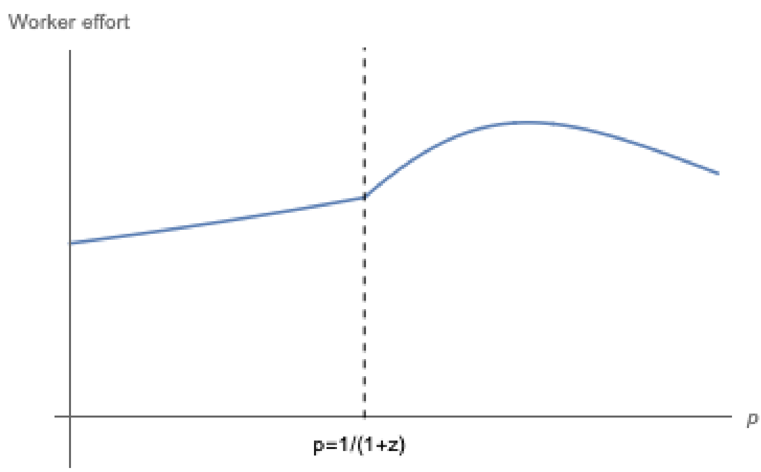


Figure A.2: Side Payment and Efforts as a Function of the Share of the Incentive Offered to the Worker ($\gamma \geq t$, $z > 1$)

(a) Side Payment



(b) Worker Effort



(c) Supervisor Effort

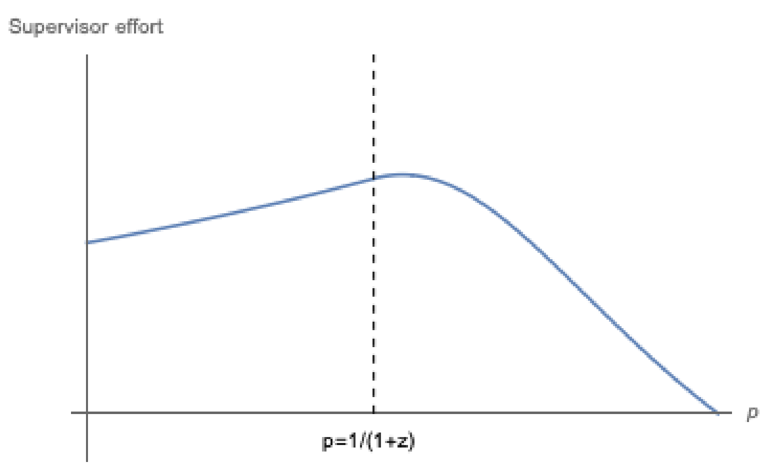
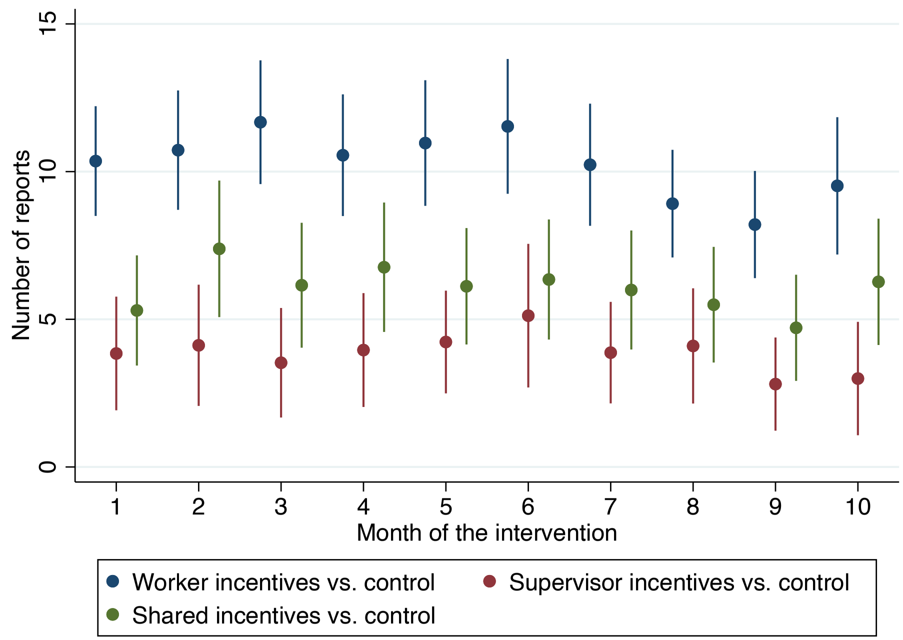


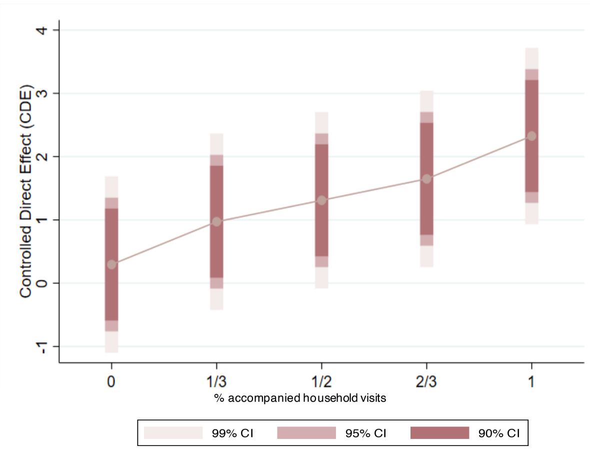
Figure A.3: Time Evolution of SMS Reporting



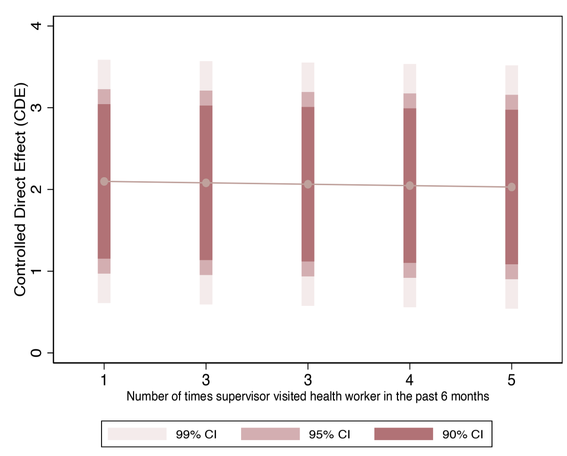
Notes: The figure plots the difference in the number of SMS reports between each treatment group and the control group. The coefficients are estimated from a regression of the number of SMS reports in each single month on the treatment dummies, controlling for the stratification variables and with standard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

Figure A.4: Mediation Analysis

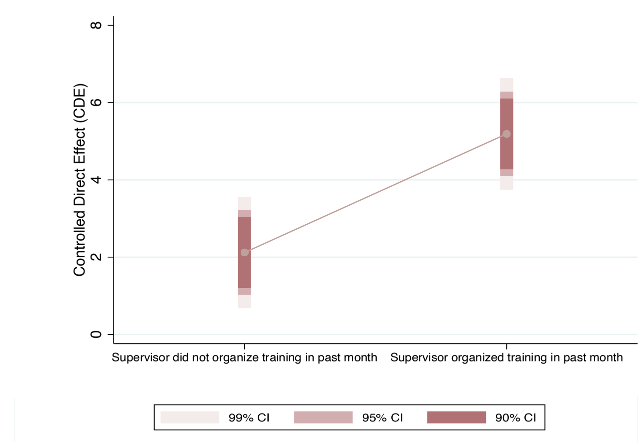
A. In-the-field training



B. One-to-one meetings



C. General training

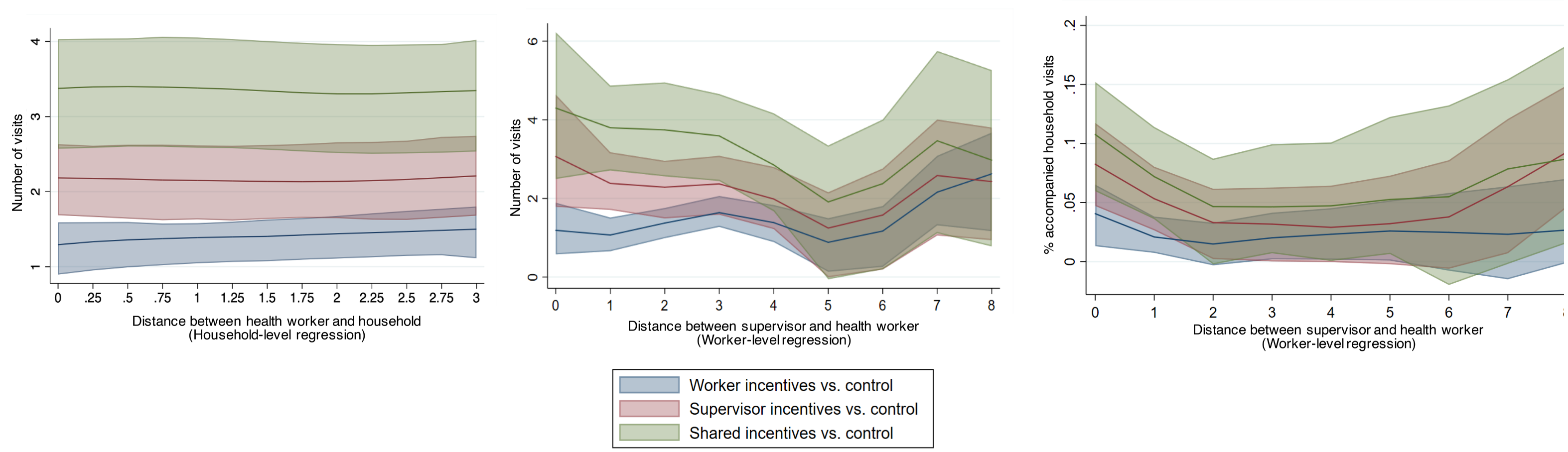


Notes: This figure plots the controlled direct effect (CDE) of the worker incentives treatment on the number of visits provided by a health worker for different values of supervisor's effort, as measured with in-the-field training (Panel A), one-to-one meetings (Panel B), general training (Panel C). Each figure is produced by following the steps outlined in Acharya et al. (2016). First, we regress the number of visits provided by a health worker on the worker incentives treatment, the mediator (supervisor's effort), and their interaction. Second, we obtain a de-mediated outcome, defined as the difference between actual visits and the number of visits predicted by the regression model for a given level of the mediator. Third, we run a regression of the de-mediated outcome on the treatment. This regression identifies the CDE of the intervention, for a given level of the mediator. We repeat this three-step procedure multiple times, changing each time the level at which we fix the mediator. This figure reports the CDE estimates corresponding to these different levels of the mediator.

Figure A.5: Non-Parametric Estimates of Visits and Supervisor Effort by Wealth and Distance

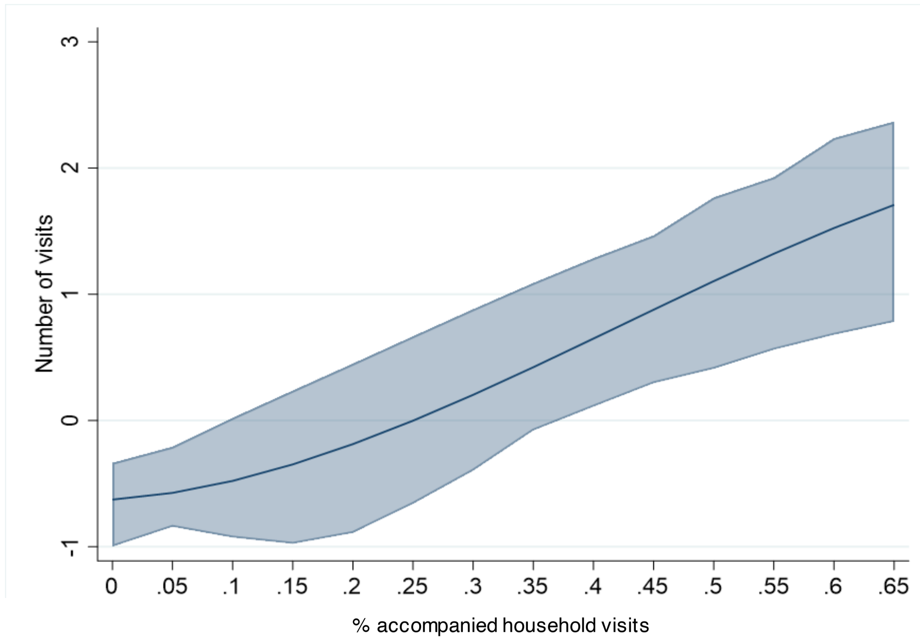
A. Treatment Effects on Visits and Supervisor Effort by Wealth

B. Treatment Effects on Visits and Supervisor Effort by Distance



Notes: This 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. We use a local linear estimator with Epanechnikov kernel function. Standard errors are bootstrapped for each value of the x-axis, with 100 repetitions and the re-sampling is with replacement. 95% confidence intervals presented in the figures. In Panel B, the "distance between health worker and household" is measured at the household level and the analysis is performed at the household level. In all the other figures, the analysis is performed at the health worker level.

Figure A.6: Non-Parametric Estimates of Visits by Supervisor Effort



Notes: This figure plots non-parametric estimates of the number of visits provided by the health worker on the fraction of accompanied household visits (supervisor effort). Standard errors are bootstrapped for each value of the x-axis, with 100 repetitions and the re-sampling is with replacement. 95% confidence intervals presented in the figure.

Table A.1: Balance Checks (Pairwise Treatment Comparisons)

	(1)	(2)	(3)	(4)	(5)	(6)
	P-values from Pairwise Treatment Comparisons					
	Worker Incentives = Supervisor Incentives	Worker Incentives = Shared Incentives	Supervisor Incentives = Shared Incentives	Worker Incentives = Control	Supervisor Incentives = Control	Shared Incentives = Control
A. Characteristics of the supervisors (N=372)						
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 = {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 responsible for	0.375	0.450	0.904	0.054	0.304	0.253
B. Characteristics of the health workers (N= 2,970)						
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 = {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 responsible for	0.532	0.353	0.711	0.291	0.138	0.096
Distance to supervisor (in km)	0.043	0.443	0.228	0.204	0.443	0.636
C. Characteristics of the female household respondent, aggregated to village level (N=2,970)						
Age (in years)	0.851	0.477	0.388	0.099	0.080	0.347
Completed primary education = {0, 1}	0.072	0.257	0.440	0.065	0.923	0.469
Completed secondary education = {0, 1}	0.924	0.776	0.712	0.669	0.755	0.470
Wealth score (0 to 8)	0.785	0.581	0.324	0.122	0.141	0.015
Distance to health worker (in km)	0.727	0.907	0.818	0.184	0.327	0.225
D. Characteristics of the village (N=372)						
Accessible road to health facility= {0, 1}	0.784	0.511	0.361	0.809	0.991	0.400
Phone network available	0.715	0.341	0.210	0.361	0.222	0.955
E. Services provided by local health facilities per month (N=372)						
Number of pregnant women services	0.669	0.539	0.811	0.311	0.467	0.637
Number of institutional births	0.749	0.740	0.565	0.229	0.432	0.206
Number of fully immunized infants	0.358	0.983	0.319	0.520	0.817	0.488
Number of malaria cases treated	0.345	0.102	0.458	0.076	0.368	0.860
Number of diarrhea cases treated	0.674	0.383	0.746	0.379	0.727	0.973

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 are clustered at the PHU level in worker/village level regressions and we use robust standard errors in PHU/supervisor level regressions. Data source is the supervisor survey in Panel A, the health worker survey in Panel B, the household survey in Panel C, the health worker's leaflet in Panel D, and the facility admin data in Panel E.

Table A.2: Household Visits (Aggregated to Entire Village)

	(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***	173.143***
	(0.245)	(4.429)	(40.189)
Supervisor incentives	0.998***	18.007***	154.653***
	(0.229)	(4.139)	(38.853)
Shared incentives	1.414***	25.505***	215.047***
	(0.214)	(3.862)	(35.299)
Unit	Worker	Worker	Supervisor
Observations	2,926	2,926	372
Mean dep. var.	3.154	56.906	448.434
Mean dep. var. in Control	2.275	41.040	303.529
p-value Worker = Supervisor	0.957	0.957	0.709
p-value Supervisor = Shared	0.110	0.110	0.184
p-value Worker = Shared	0.116	0.116	0.369

Notes: Data source is the household survey in cols. (1) and (2), aggregated to health worker level. Col. (2) divides the total number of household visits provided by the health worker to sampled households per month (col. 1) by the share of households in the village sampled. All regressions include stratification variables. Standard errors are clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1.

Table A.3: Household Visits by 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 visits				Disease-related visits			
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	0.037	0.660	0.455	0.154	0.046	0.143	0.046	0.024
p-value Worker = Shared	0.132	0.069	0.151	0.056	0.014	0.020	0.094	0.008

Notes: Data source is the household survey, aggregated to health worker level. 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)]. All regressions include stratification variables. Standard errors clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1.

Table A.4: P-values Corrected for Multiple Hypothesis Testing

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	p-values with MHT correction								
	Romano and Wolf (2016)			Bonferroni correction			Benjamini, Krieger, and Yekutieli (2006)		
	Worker incentives	Supervisor incentives	Shared incentives	Worker incentives	Supervisor incentives	Shared incentives	Worker incentives	Supervisor incentives	Shared incentives
A. Household visits provided by the health worker in the past 6 months [source = household survey]									
Number of visits	0.004	0.004	0.004	0.004	<0.001	<0.001	0.001	0.001	0.001
% households visited	0.004	0.004	0.004	0.080	0.018	<0.001	0.004	0.002	0.001
Number of visit types	0.004	0.004	0.004	0.149	0.023	<0.001	0.006	0.002	0.001
Average visit length	0.004	0.004	0.004	0.579	0.676	<0.001	0.016	0.017	0.001
Number of health topics discussed per visit	0.052	0.052	0.004	1.000	1.000	0.002	0.051	0.051	0.001
% households who trust the health worker as a health provider	0.020	0.052	0.004	1.000	1.000	0.048	0.045	0.051	0.003
B. Pre- and post-natal care, and disease incidence [source= household survey]									
Index on pre- and post-natal care in past 2 years (for next 5 variables)	0.845	0.470	0.004	1.000	1.000	0.038	0.960	0.763	0.040
% women who received at least 4 ante-natal visits before birth	0.972	0.657	0.028	1.000	1.000	0.553	1.000	0.763	0.208
% women with institutional birth	0.861	0.191	0.179	1.000	1.000	1.000	0.960	0.489	0.489
% women who received post-natal visit within 2 days of birth	0.992	0.900	0.984	1.000	1.000	1.000	1.000	1.000	1.000
% women with at least 6 months of breastfeeding	0.992	0.984	0.351	1.000	1.000	1.000	1.000	1.000	0.629
% households with up-to-date infant vaccination	0.992	0.944	0.633	1.000	1.000	1.000	1.000	1.000	0.763
Index on disease incidence (for next 3 variables)	0.992	0.920	0.072	1.000	1.000	1.000	1.000	1.000	0.420
% children under-5 who had fever	0.641	0.992	0.012	1.000	1.000	0.238	0.763	1.000	0.131
% children under-5 who had cough	0.542	0.992	0.984	1.000	1.000	1.000	0.763	1.000	1.000
% children under-5 who had diarrhea	0.972	0.880	0.992	1.000	1.000	1.000	1.000	0.96	1.000
C. Reporting [source = SMS amin data and household survey]									
Number of reports per month	0.004	0.004	0.004	<0.001	0.006	<0.001	0.001	0.001	0.001
Reporting rate = number of reports/ number of visits	0.004	0.135	0.024	<0.001	1.000	0.310	0.001	0.052	0.011
Over-reporting = {0, 1} [number of reports > number of visits]	0.020	0.888	0.880	0.268	1.000	1.000	0.01	0.248	0.218
Under-reporting = {0, 1} [number of reports < number of visits]	0.888	0.080	0.052	1.000	1.000	0.787	0.341	0.035	0.022
D. Cost of the intervention [source = pay admin data]									
Incentives payment per health worker (in 1,000 SLL)	0.004	0.888	0.004	<0.001	1.000	<0.001	0.001	0.289	0.001
Total incentives payment across all health workers (in 1,000 SLL)	0.004	0.888	0.004	<0.001	1.000	<0.001	0.001	0.248	0.001
Incentives payment per supervisor (in 1,000 SLL)	0.880	0.004	0.004	1.000	<0.001	<0.001	0.189	0.001	0.001
Total incentives payments in the PHU (in 1,000 SLL)	0.004	0.004	0.004	<0.001	<0.001	<0.001	0.001	0.001	0.001
Number of visits per worker for each 2,000 SLL spent	-	0.880	0.080	-	1.000	1.000	-	0.189	0.033
E. PS effort [source = household and health worker survey]									
% accompanied household visits	0.151	0.016	0.012	1.000	0.219	0.057	0.523	0.114	0.061
% health workers visited in the past 6 months	0.857	0.857	0.857	1.000	1.000	1.000	1.000	1.000	1.000
Supervisor organized training in the past month = {0, 1}	0.745	0.151	0.857	1.000	1.000	1.000	1.000	0.523	1.000
Difference in health workers' knowledge between baseline and endline	0.151	0.857	0.024	1.000	1.000	0.429	0.523	1.000	0.142
Supervisor ever helped health worker with SMS reporting = {0, 1}	0.964	0.781	0.964	1.000	1.000	1.000	1.000	1.000	1.000
F. Side-payments [source = health worker and supervisor survey]									
Supervisor shared incentive with health worker = {0, 1}	0.805	0.008	0.100	1.000	0.002	0.137	0.277	0.001	0.013
Health worker shared incentive with supervisor = {0, 1}	0.004	0.920	0.088	<0.001	1.000	0.076	0.001	0.331	0.012
Monthly transfer from supervisor to health worker (in 1,000 SLL)	0.319	0.020	0.100	1.000	0.004	0.099	0.08	0.002	0.012
Monthly transfer from health worker to supervisor (in 1,000 SLL)	0.100	0.108	0.159	0.107	0.247	0.771	0.012	0.016	0.029
Net transfer (difference between the last 2 variables, in 1,000 SLL)	0.797	0.064	0.131	1.000	0.028	0.514	0.221	0.006	0.025

Notes: This table presents multiple hypothesis testing corrected p-values for regressions of the row variable on the three treatment indicators, controlling stratification variables and clustering standard errors at the PHU level. Cols. (1) to (6) control for the familywise error rate, i.e. the probability of making any type 1 error. Cols. (7) to (9) control for the false discovery rate, i.e. the expected proportion of rejections that are type 1 errors. The adjusted p-values in cols. (1) to (3) are calculated following the Romano and Wolf (2016) step-down procedure, with 250 bootstrap resampling iterations. In cols. (4) to (6), a Bonferroni adjustement is applied by multiplying the original p-values by the number of outcomes in the table and capping the adjusted p-values at 1.000. Cols (7) to (9) apply the two-stage Benjamini, Krieger, and Yekutieli (2006)'s procedure.

Table A.5: Household Targeting

	(1)	(2)	(3)	(4)	(5)
Dep. Var.	Number of visits				
Definition of covariate X:	Household's wealth score (0 to 8)	Household's distance to health worker (in km)	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 visit accompanied by the supervisor = {0, 1}
Worker incentives	2.118***	2.002***	2.113***	2.080***	1.661*
	(0.555)	(0.673)	(0.568)	(0.587)	(0.912)
Supervisor incentives	2.129***	1.972***	1.889***	1.932***	0.670
	(0.509)	(0.584)	(0.505)	(0.505)	(0.806)
Shared incentives	3.385***	3.632***	3.515***	3.461***	2.227***
	(0.492)	(0.695)	(0.540)	(0.523)	(0.859)
X	0.046	-0.333***	2.583***	0.261	-3.285***
	(0.095)	(0.097)	(0.595)	(0.581)	(0.460)
Worker incentives * X	-0.208	0.206	-0.353	0.244	0.368
	(0.158)	(0.198)	(0.828)	(1.010)	(0.822)
Supervisor incentives * X	0.042	0.199	0.474	1.629	1.594*
	(0.134)	(0.152)	(0.772)	(1.073)	(0.867)
Shared incentives * X	-0.000	0.072	-0.836	-1.066	1.118
	(0.138)	(0.142)	(0.853)	(0.977)	(0.829)
Unit	Household	Household	Household	Household	Household
Observations	8,559	5,538	8,601	8,601	8,459
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	1.465	0.308	0.112	0.793
p-value Worker*X = Supervisor*X	0.151	0.974	0.275	0.261	0.221
p-value Supervisor*X = Shared*X	0.783	0.441	0.096	0.028	0.634
p-value Worker*X = Shared*X	0.232	0.519	0.566	0.246	0.447

Notes: Data source is the household survey. One observation per household. All regressions include stratification variables. Standard errors clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1

Table A.6: More Health Outcomes

	(1)	(2)	(3)
Dep. Var.	Presence of a woman who gave birth in the past year in the household = {0, 1}	% household respondents who know how to prevent malaria and diarrhea	Under-5 mortality rate
Worker incentives	0.036	0.062*	-0.318
	(0.023)	(0.032)	(2.220)
Supervisor incentives	0.035	0.053*	3.083
	(0.023)	(0.032)	(3.838)
Shared incentives	0.010	0.086***	-1.485
	(0.024)	(0.030)	(2.093)
Unit	Worker	Worker	Worker
Observations	2,970	2,970	2,824
Mean dep. var.	0.841	0.563	4.135
Mean dep. var. in Control	0.819	0.511	3.822
p-value Worker = Supervisor	0.976	0.796	0.380
p-value Supervisor = Shared	0.273	0.312	0.218
p-value Worker = Shared	0.263	0.460	0.571

Notes: Data source is the household survey, aggregated to health worker level. In col. (2), a household respondent "knows how to prevent malaria and diarrhea" if she reports that children should (1) sleep under mosquito nets that is sprayed with mosquito repellent/insecticide, (2) use soap, (3) use toilet facility to defecate, and (4) drink clean water. Mortality under-5 is measured as child mortality per 1000 years of exposure to the risk of death (self-reported by households and aggregated to village level). The measure follows the method used Bjorkman Nyqvist et al. (2019). All regressions include stratification variables. Standard errors are clustered at the PHU level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.7: Health Facility Services

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	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. 6-7)	Number of malaria cases treated	Number of diarrhea cases treated
Worker incentives	0.108	5.001	0.895	2.302	0.175*	9.857*	4.309**
	(0.094)	(5.511)	(1.020)	(1.553)	(0.099)	(5.839)	(2.102)
Supervisor incentives	0.097	4.799	1.731	0.852	0.187**	10.104*	2.420
	(0.083)	(4.220)	(1.090)	(1.352)	(0.094)	(5.767)	(2.231)
Shared incentives	0.244*	13.918*	2.552*	3.042*	0.223*	9.455	6.309*
	(0.126)	(7.283)	(1.389)	(1.625)	(0.132)	(6.621)	(3.269)
Unit	PHU	PHU	PHU	PHU	PHU	PHU	PHU
Observations	371	371	371	371	371	371	371
Mean dep. var.	0.000	41.889	13.776	12.406	0.000	57.464	18.936
Mean dep. var. in Control	-0.110	36.063	12.513	10.892	-0.155	49.595	15.582
p-value Worker = Supervisor	0.916	0.971	0.483	0.386	0.908	0.967	0.494
p-value Supervisor = Shared	0.269	0.209	0.584	0.203	0.800	0.923	0.291
p-value Worker = Shared	0.323	0.263	0.246	0.690	0.733	0.953	0.569

Notes: Data source is the admin data from the local health facilities (one per PHU). The number of observations is 371 instead of 372 because of one missing variable. 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)-(7)]. All regressions include stratification variables. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.8: Reporting by Network Availability

	(1)	(2)
Dep. Var.	Reporting rate = number of reports/number of visits	
A. Treatment effects for villages <u>without</u> phone network		
No network * Worker incentives	0.080	0.070
	(0.059)	(0.068)
No network * Supervisor incentives	0.011	0.002
	(0.045)	(0.054)
No network * Shared incentives	0.022	0.022
	(0.044)	(0.050)
B. Treatment effects for villages <u>with</u> phone network		
Network * Worker incentives	0.238***	0.240***
	(0.046)	(0.051)
Network * Supervisor incentives	0.063*	0.068*
	(0.036)	(0.040)
Network * Shared incentives	0.101**	0.089**
	(0.039)	(0.040)
Network	0.014	0.015
	(0.027)	(0.033)
Unit	Worker	Worker
Extra Controls	No	Yes
Observations	2,532	2,227
Mean Dep. Var.	0.177	0.177
Mean Dep. Var. in Control & No Network	0.080	0.078
Treatment comparisons in Panel A (No network)		
p-value Worker=Supv	0.186	0.232
p-value Worker=Shared	0.257	0.378
p-value Supv=Shared	0.748	0.587
Treatment comparisons in Panel B (Network)		
p-value Worker=Supv	0.001	0.002
p-value Worker=Shared	0.010	0.007
p-value Supv=Shared	0.389	0.641
Treatment comparisons across Panels (No network vs. network)		
p-value for Worker incentives	0.016	0.025
p-value for Supervisor incentives	0.239	0.209
p-value for Shared incentives	0.078	0.138

Notes: The table reports coefficients from a fully interacted model in which the treatment dummies are interacted with a dummy for whether the network is available in the village. The outcome variable is measured as the number of reports (from the SMS admin data) divided by the total number of visits per month (from the household survey). The latter is measured as the total number of household visits provided by the health worker to sampled households per month divided by the share of households in the village sampled. Col. (2) controls for the correlates of network availability ($p<0.1$) -- i.e., age and wealth of the health worker, number of households the health worker is responsible for, distance to supervisor -- interacted with the treatment dummies. All regressions include stratification variables. Standard errors are clustered at the PHU level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.9: Correlations

	(1)	(2)	(3)	(4)	(5)	(6)
	Health worker's years of experience	Health worker has high experience (above median of 4 years) = {0, 1}	Supervisor's observability of output [Correlation between actual and perceived ranking]	Supervisor has low observability of output = {0, 1} [Correlation between actual and perceived ranking is in bottom decile]	Difference between supervisor and health worker hourly pay in other job	Supervisor has higher hourly pay than health worker in other job = {0,1}
Unit of Observations	Worker	Worker	Supervisor	Supervisor	Worker	Worker
Mean	5.039	0.473	0.067	0.144	-0.958	0.514
S.D.	4.481	-	0.425	-	18.358	-
A. Characteristics of the health workers						
Male = {0, 1}	0.129***	0.119***			-0.028	0.006
Age (in years)	0.166***	0.178***			0.026	0.045
Completed primary education = {0, 1}	-0.024	-0.051***			-0.002	-0.038
Completed secondary education = {0, 1}	0.022	-0.006			0.027	-0.064
Wealth score (0 to 8)	-0.052***	-0.073***			-0.070*	-0.048
Number of households responsible for	0.017	0.007			0.020	-0.002
Distance to supervisor (in km)	0.069***	0.084***			-0.018	0.013
B. Characteristics of the supervisors						
Male = {0, 1}			-0.013	0.066		
Age (in years)			0.044	-0.061		
Completed primary education = {0, 1}			0.086	-0.008		
Completed secondary education = {0, 1}			-0.098*	0.016		
Wealth score (0 to 8)			0.032	-0.067		
Number of health workers responsible for			0.086	-0.245***		
Average distance to supervisor (in km)			-0.020	0.092*		
% health workers > 1 km away			-0.038	0.133**		

Notes: This table presents summary statistics for each column variable, and pairwise correlations of the column variable with the raw variable. Data source is the health worker or supervisor survey. *** p<0.01, ** p<0.05, * p<0.1.

Table A.10: Heterogeneity by Worker Experience

	(1)	(2)	(3)	(4)
Dep. Var.	Number of visits		Supervisor effort: % accompanied household visits	
A. Treatment effects for workers with experience below the median:				
Low experience * Worker incentives	2.054***	2.395***	0.030	0.030
	(0.628)	(0.722)	(0.025)	(0.027)
Low experience * Supervisor incentives	2.576***	2.661***	0.067**	0.067**
	(0.598)	(0.646)	(0.029)	(0.031)
Low experience * Shared incentives	4.022***	4.335***	0.092***	0.104***
	(0.684)	(0.751)	(0.026)	(0.027)
B. Treatment effects for workers with experience above the median:				
High experience * Worker incentives	2.246***	2.056***	0.030	0.031
	(0.780)	(0.756)	(0.031)	(0.032)
High experience * Supervisor incentives	1.720**	1.657**	0.045	0.045
	(0.669)	(0.643)	(0.030)	(0.031)
High experience * Shared incentives	2.583***	2.638***	0.030	0.022
	(0.608)	(0.670)	(0.030)	(0.032)
High experience	1.057**	1.141*	0.017	0.033
	(0.532)	(0.594)	(0.025)	(0.028)
Unit	Worker	Worker	Worker	Worker
Extra Controls	No	Yes	No	Yes
Observations	2,909	2,552	2,902	2,547
Mean Dep. Var.	7.296	7.296	0.204	0.204
Mean Dep. Var. in Control & Low experience	4.749	4.749	0.131	0.131
Treatment comparisons in Panel A (Low experience)				
p-value Worker=Supv	0.455	0.733	0.226	0.236
p-value Worker=Shared	0.011	0.026	0.029	0.010
p-value Supv=Shared	0.057	0.038	0.431	0.248
Treatment comparisons in Panel B (High experience)				
p-value Worker=Supv	0.551	0.630	0.643	0.676
p-value Worker=Shared	0.684	0.492	0.990	0.781
p-value Supv=Shared	0.234	0.192	0.632	0.482
Treatment comparisons across Panels (Low vs. High experience)				
p-value for Worker incentives	0.824	0.716	0.994	0.973
p-value for Supervisor incentives	0.270	0.218	0.535	0.572
p-value for Shared incentives	0.094	0.077	0.086	0.039

Notes: The table reports the coefficients from a fully interacted model in which the treatment dummies are interacted with a dummy for whether the worker's experience is high or low. "Low experience" is an indicator that takes value one if the health worker has less than the median number of experience (i.e., less than 4 years of experience) as a health worker at baseline. Data source for the outcome variable is the household survey, aggregated to health worker level in cols. (1)-(2) and the health worker survey in cols. (3)-(4). Cols. (2) and (4) control for the health worker characteristics that are significantly correlated (p<.1) with experience -- i.e., gender, age, wealth score, distance to supervisor -- interacted with the treatment dummies. All regressions include stratification variables. Standard errors are clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1.

Table A.11: Heterogeneity by Performance Ranking, Social Distance and Span of Control

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. Var.	Number of visits	Supervisor effort: % accompanied household visits	Side-payment: Supervisor shared incentive with health worker = {0, 1}	Number of visits	Supervisor effort: % accompanied household visits	Side-payment: Supervisor shared incentive with health worker = {0, 1}	Number of visits	Supervisor effort: % accompanied household visits	Side-payment: Supervisor shared incentive with health worker = {0, 1}
Panel A	*Effects for "Low Rank" workers: ranked in the <u>bottom</u> half by the supervisor at baseline*			*Effects for "Low Social Distance" workers: friends/family of the supervisor at baseline*			*Effects for "Low Span of Control" PHU: supervisor responsible for less than 7 health workers*		
Low X * Worker incentives	2.366***	0.038	0.006	1.726***	0.010	0.004	1.828**	-0.038	0.042
	(0.692)	(0.026)	(0.016)	(0.612)	(0.025)	(0.016)	(0.850)	(0.034)	(0.040)
Low X * Supervisor incentives	2.356***	0.050*	0.182***	2.322***	0.042	0.170***	1.730*	0.013	0.077
	(0.583)	(0.028)	(0.047)	(0.616)	(0.030)	(0.044)	(1.003)	(0.041)	(0.049)
Low X * Shared incentives	3.752***	0.078***	0.097**	3.165***	0.052**	0.108**	2.714***	0.004	0.107**
	(0.592)	(0.026)	(0.038)	(0.580)	(0.024)	(0.045)	(0.711)	(0.032)	(0.053)
Panel B	*Effects for "High Rank" workers: ranked in the <u>top</u> half by the supervisor at baseline*			*Effects for "High Social Distance" workers: not friends/family of the supervisor at baseline*			*Effects for "High Span of Control" PHU: supervisor responsible for 8 health workers or more*		
High X * Worker incentives	1.939***	0.020	0.002	2.559***	0.064**	0.007	2.350***	0.055**	-0.007
	(0.626)	(0.026)	(0.016)	(0.788)	(0.032)	(0.019)	(0.700)	(0.027)	(0.019)
High X * Supervisor incentives	1.589**	0.033	0.173***	2.013***	0.086***	0.193***	2.448***	0.076***	0.218***
	(0.722)	(0.027)	(0.048)	(0.646)	(0.030)	(0.054)	(0.589)	(0.028)	(0.062)
High X * Shared incentives	3.301***	0.044*	0.110**	3.592***	0.076**	0.095**	3.722***	0.090***	0.099*
	(0.668)	(0.026)	(0.043)	(0.752)	(0.032)	(0.043)	(0.634)	(0.027)	(0.052)
High X	0.515	0.009	0.004*	-0.244	-0.065***	0.004	-1.120*	-0.016	0.012
	(0.481)	(0.019)	(0.002)	(0.519)	(0.024)	(0.010)	(0.578)	(0.031)	(0.036)
Meaning of X	**Performance Ranking**			**Social Distance**			**Span of Control**		
Unit	Worker	Worker	Worker	Worker	Worker	Worker	Worker	Worker	Worker
Observations	2,696	2,689	2,685	2,915	2,908	2,904	2,926	2,919	2,915
Mean Dep. Var.	7.296	0.204	0.084	7.296	0.204	0.084	7.296	0.204	0.084
Mean Dep. Var. in Control & Low X	5.118	0.161	0.012	5.410	0.145	0.014	5.619	0.181	0.000
Treatment comparisons in Panel A (Low X)									
p-value Worker=Supv	0.989	0.696	0.000	0.420	0.312	0.000	0.935	0.224	0.574
p-value Worker=Shared	0.072	0.145	0.017	0.045	0.107	0.022	0.364	0.201	0.311
p-value Supv=Shared	0.038	0.331	0.145	0.237	0.746	0.303	0.376	0.837	0.665
Treatment comparisons in Panel B (High X)									
p-value Worker=Supv	0.647	0.648	0.000	0.494	0.455	0.001	0.894	0.505	0.000
p-value Worker=Shared	0.059	0.405	0.012	0.242	0.711	0.040	0.081	0.238	0.028
p-value Supv=Shared	0.033	0.717	0.313	0.040	0.736	0.138	0.066	0.643	0.120
Treatment comparisons across Panels (Low vs. High X)									
p-value for Worker incentives	0.523	0.505	0.214	0.312	0.121	0.807	0.637	0.033	0.268
p-value for Shared incentives	0.538	0.246	0.161	0.627	0.490	0.757	0.288	0.041	0.921
p-value for Supervisor incentives	0.341	0.576	0.281	0.681	0.210	0.473	0.536	0.211	0.082

Notes: The table reports the coefficients from a fully interacted model in which the treatment dummies are interacted with a dummy for whether the worker is ranked in the bottom/top half by the supervisor in cols (1)-(3), whether the worker is a friend/family member of the supervisor in cols (4)-(6), whether the span of control is high or low in cols. (7)-(9). All regressions include stratification variables. Standard errors are clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1.

Table A.12: Heterogeneity by Output Observability

	(1)	(2)
Dep. Var.	Side-payment: Supervisor shared incentive with health worker = {0, 1}	
A. Treatment effects for supervisors with low observability of output:		
Low observability * Worker incentives	-0.028	-0.027
	(0.032)	(0.031)
Low observability * Supervisor incentives	0.032	0.044
	(0.065)	(0.064)
Low observability * Shared incentives	0.067	0.062
	(0.089)	(0.087)
B. Treatment effects for supervisors with high observability of output:		
High observability * Worker incentives	0.009	0.007
	(0.018)	(0.019)
High observability * Supervisor incentives	0.205***	0.200***
	(0.052)	(0.052)
High observability * Shared incentives	0.107**	0.110**
	(0.042)	(0.043)
High observability	-0.008	-0.006
	(0.023)	(0.025)
Unit	Worker	Worker
Extra Controls	No	Yes
Observations	2,915	2,915
Mean Dep. Var.	0.084	0.084
Mean Dep. Var. in Control & Low observability	0.000	0.000
Treatment comparisons in Panel A (Low observability)		
p-value Worker=Supv	0.370	0.293
p-value Worker=Shared	0.305	0.335
p-value Supv=Shared	0.750	0.871
Treatment comparisons in Panel B (High observability)		
p-value Worker=Supv	<0.001	<0.001
p-value Worker=Shared	0.021	0.017
p-value Supv=Shared	0.135	0.160
Treatment comparisons across Panels (Low vs. High observability)		
p-value for Worker incentives	0.315	0.380
p-value for Supervisor incentives	0.040	0.061
p-value for Shared incentives	0.680	0.619

Notes: The table reports the coefficients from a fully interacted model in which the treatment dummies are interacted with a dummy for whether the supervisor has high/low observability of output. "Low observability" is an indicator that takes value one if the correlation between the actual worker ranking (based on endline household data on visit) and the supervisor's perceived worker ranking at endline (based on the supervisor survey) is in the bottom decile (i.e. is negative). Data source for the outcome variable is the health worker survey. Col. (2) also controls for correlates of observability (i.e., supervisor completed secondary school), interacted with the treatment dummies. All regressions include stratification variables. Standard errors are clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1.

Table A.13: Sample of Workers with Higher Outside Option than Supervisors

	(1)	(2)
Dep. Var.	Number of visits	Side-payment: Supervisor shared incentive with health worker = {0, 1}
Sample:	Workers with higher outside option than their supervisor	
Worker incentives	0.453	-0.033
	(1.344)	(0.065)
Supervisor incentives	2.303	0.248**
	(1.396)	(0.103)
Shared incentives	3.286**	0.029
	(1.322)	(0.080)
Unit	Worker	Worker
Observations	291	293
Mean dep. var.	7.641	0.116
Mean dep. var. in Control	5.848	0.044
p-value Worker = Supervisor	0.184	0.020
p-value Supervisor = Shared	0.498	0.088
p-value Worker = Shared	0.033	0.323

Notes: Sample restricted to workers with higher outside option than supervisor. These are workers with an average hourly earnings from any outside (secondary) job which is higher than the one of their supervisor, conditional on both the worker and the supervisor being engaged in an outside job with a positive income (data from health worker and supervisor surveys). Data source for the outcome variable is the household survey, aggregated to health worker level in col. (1) and the health worker survey in col. (2). All regressions include stratification variables. Standard errors are clustered at the PHU level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.14: Job Satisfaction

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Health worker is satisfied with […] = {0, 1}			Supervisor is satisfied with […] = {0, 1}					Number of visits		
Dep. Var.	Incentives payments	Organization paying the incentives	Job	Incentives payments	Organization paying the incentives	Job	Health worker perceives the work environment as non-competitive	Health worker self-identifies herself through her job = {0, 1}	Sample: Workers with "low" inequality aversion	Sample: Workers with "medium" inequality aversion	Sample: Workers with "high" inequality aversion
Worker incentives	0.344***	0.074***	-0.039	0.086	0.012	-0.067	-0.046	0.006	1.982*	2.359***	0.922
	(0.032)	(0.024)	(0.036)	(0.056)	(0.049)	(0.061)	(0.044)	(0.026)	(1.063)	(0.571)	(1.257)
Supervisor incentives	-0.003	0.014	-0.040	0.305***	0.082*	-0.073	-0.023	-0.019	1.823***	2.419***	1.757
	(0.030)	(0.028)	(0.038)	(0.062)	(0.042)	(0.062)	(0.043)	(0.030)	(0.679)	(0.598)	(2.064)
Shared incentives	0.198***	0.048**	-0.054	0.329***	0.092**	-0.040	-0.024	-0.018	2.803***	3.675***	4.369**
	(0.036)	(0.024)	(0.036)	(0.063)	(0.041)	(0.059)	(0.043)	(0.029)	(0.835)	(0.572)	(1.831)
Unit	Worker	Worker	Worker	Supervisor	Supervisor	Supervisor	Worker	Worker	Worker	Worker	Worker
Observations	2,709	2,825	2,876	360	364	359	2,923	2,923	772	1,913	236
Mean dep. var.	0.357	0.870	0.793	0.311	0.909	0.738	0.727	0.829	7.150	7.442	6.532
Mean dep. var. in Control	0.219	0.837	0.828	0.132	0.860	0.787	0.746	0.838	5.795	5.017	5.478
p-value Worker = Supervisor	<0.001	0.023	0.986	0.001	0.098	0.929	0.598	0.379	0.884	0.929	0.662
p-value Supervisor = Shared	<0.001	0.186	0.709	0.738	0.757	0.605	0.987	0.963	0.252	0.062	0.266
p-value Worker = Shared	<0.001	0.254	0.679	<0.001	0.048	0.664	0.609	0.394	0.487	0.040	0.074

Notes: Data source is the health worker survey in cols. (1)-(3) and (7)-(11), and the supervisor survey in cols. (4)-(6). A worker/supervisor is defined as unsatisfied with the incentive payment if she reports that the incentive she is paid per valid SMS report is "not fair" (too little). A worker/supervisor is defined as unsatisfied with the environment if she reports that the environment is competitive rather than cooperative. ``Health worker self-identifies herself through her job" is a dummy variable that takes value one if the health worker answers "my job as a community health worker" to the following question: "We have spoken with many people in Sierra Leone and they identify themselves to different groups. Some people self identify themselves as belong to an ethnic group, a language, a religion, etc. Others identify themselves describe themselves in terms of their job. Besides being a citizen of Sierra Leone, which specific group do you feel you belong to first and foremost?". Questions in cols. (7) and (8) were not asked to the supervisor. The sample size changes across columns because a number of health workers and supervisors answered "don't know" to the questions. Cols. (9) to (11) are restricted to workers indicated in the panel headings. 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 ("low") if the worker answers that Sheka would always show up to work, regardless of the amount; value 1 ("medium") if worker reports that Sheka would not show up only if amount=120,000 and would show up otherwise; value 2 ("high") if the worker reports that Sheka would not show up if amount > 30,000. All regressions include stratification variables. Standard errors clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1

Table A.15: Parameter Estimates (No Effort Complementarities)

	(1)
Marginal product of worker effort α	75.41
Marginal product of supervisor effort β	0.00
Worker cost of effort parameter c_1	457.45
Supervisor cost of effort parameter c_2	0.16
Worker baseline incentive b_1	6.84
Supervisor baseline incentive b_2	13.76
Calibrated friction z	1.45

Notes: The first panel of the table shows parameter estimates obtained using minimum distance estimation for the version of the model where the supervisor correctly expects the reporting rate to differ by treatment. The second panel first shows the calibrated value of contractual frictions. Second, it shows some quantities implied by the parameter estimates.

Table A.16: Moment Fit (No Effort Complementarities)

Moments	Targeted Real	Simulated
Supervisor effort in worker incentive group	0.198	0.206
Supervisor effort in supervisor incentive group	0.225	0.204
Supervisor effort in shared incentive group	0.228	0.191
Supervisor effort in control group	0.164	0.206
Output in worker incentive group	59.679	64.476
Output in supervisor incentive group	58.896	58.054
Output in shared incentive group	66.895	59.785
Output in control group	41.040	42.500
Value loss function	43.0	

Notes: The table shows the targeted empirical moments used for minimum distance estimation as well as the simulated moments. In this version of the model the supervisor correctly expects the reporting rate to differ by treatment.

Table A.17: Parameter Estimates (Alternative Assumption on Expected Reporting Rate)

	(1)
Complementarity γ	28.5
Worker baseline incentive b_1	133.1
Supervisor baseline incentive b_2	93.2
α	5.4
Calibrated friction z	11.74
Δ in marginal product of worker effort (shared incentive)	117%
Δ in marginal product of worker effort (control)	83%
Total worker cost of effort (control)	2,739.8
Total supervisor cost of effort (control)	1,733.6

Notes: The first panel of the table shows parameter estimates obtained using minimum distance estimation for the version of the model where the supervisor correctly expects the reporting rate to differ by treatment. The second panel first shows the calibrated value of contractual frictions. Second, it shows some quantities implied by the parameter estimates. We report the total cost of efforts rather than the worker and supervisor cost of effort parameters (c_1 and c_2) because the latter are hard to interpret.

Table A.18: Moment Fit (Alternative Assumption on Expected Reporting Rate)

Moments	Targeted Real	Simulated
Supervisor effort in worker incentive group	0.198	0.205
Supervisor effort in supervisor incentive group	0.225	0.231
Supervisor effort in shared incentive group	0.228	0.221
Supervisor effort in control group	0.164	0.156
Output in worker incentive group	59.679	61.679
Output in supervisor incentive group	58.896	60.773
Output in shared incentive group	66.895	62.285
Output in control group	41.040	41.157
Value loss function	6.6	

Notes: The table shows the targeted empirical moments used for minimum distance estimation as well as the simulated moments. In this version of the model the supervisor correctly expects the reporting rate to differ by treatment.

Table A.19: Heterogeneity by Promotion Incentives

	(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 per visit	
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. fixed salary	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. fixed salary	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. fixed salary	-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. fixed salary	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. fixed salary	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. fixed salary	-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. fixed salary	-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: Data source is the household survey, aggregated to health worker level. All regressions include stratification variables. Standard errors clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1

Table A.20: Sample of Health Workers without Promotion Incentives

	(1)	(2)	(3)	(4)	(5)	(6)
	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	Number of health topics discussed per visit	% households who trust the health worker as a health provider
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: Data source is the household survey, aggregated to health worker level. Sample restricted to health workers who did not experience any change in promotion incentives. All regressions include stratification variables. Standard errors clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1

B Additional Material on the Context and Intervention

B.1 Context: Checklists

Workers are expected to follow a checklist when they visit a household. The checklist differs depending on the type of visit the health worker conducts:

(i) Prenatal visits to a pregnant woman: Health workers are asked to visit expecting mothers at least four times over the course of a pregnancy. During these visits, health workers should first make sure not only the pregnant woman but also her husband or other decision-makers in the family are present. Second, they assess the pregnant woman for danger signs (e.g., convulsion or fever) that would require an immediate referral to the PHU. Third, they use the Mother, Newborn, and Child Health Card to assess previously agreed actions and current health practices related to the pregnancy with the family. Fourth, health workers present new visit-specific information to the family (e.g., helping with planning for the birth including arranging transportation so the woman can give birth at the PHU). Fifth, health workers and families identify barriers together and agree on an action plan until the next visit. Finally, health workers must fill a register that documents what they have done during the visit.

(ii) Accompanying a pregnant woman to the PHU for childbirth: The health workers should accompany pregnant women to the PHU to give birth. At the PHU, the health worker should help the family to obtain all necessary drugs and other supplies. In case a woman delivers at her home rather than at the PHU, the health worker should assist during the birth, communicate the birth to the head of the PHU, and accompany the woman for a post-natal visit at the PHU as soon as possible after the birth.

(iii) Postnatal visits within one month of birth: Health workers are asked to visit mothers with newborn babies at least four times during the first month after birth. During these visits, health workers first assess the mother and baby for the presence of any danger signs (e.g., fever or convulsions) that would require a referral to the PHU. Second, they discuss with the family how well they were able to implement health practices agreed upon with the health worker during the previous visit. Third, health workers present new visit-specific information about health behaviors relevant to the mother and baby (e.g., telling the mother to keep the baby warm and only breastfeed the baby). Fourth, they go over a checklist of recommended health behaviors and check whether or not the family knows about and follows them. Fifth, for the items on the checklist that the family does not follow yet, health workers discuss barriers and possible solutions with the family and make a new action plan to be discussed during the next visit. Finally, health workers fill out a register that documents what they have done during the visit.

(iv) Child health checkup visits: Health workers are asked to visit mothers and their young children five times between the age of 1 - 15 months. During these visits, health workers first assess the child for danger signs (e.g., convulsions or being unable to breastfeed) that would require an immediate referral to the PHU. Second, they use the Mother, Newborn, and Child Health Card to assess previously agreed actions and current health practices related to the pregnancy with the family. Third, health workers present visit-specific information to the mother (e.g., advising the mother on how to transition from exclusive breastfeeding to other foods after the age of 6 months or reminding the mother of scheduled vaccinations for the child). Fourth, health workers and families identify barriers together and agree on an action plan until the next visit. Finally, health workers must fill a register that documents what they have done during the visit.

(v) Visits in which a disease is diagnosed and the patient is either treated or referred to the health facility: The main focus of health workers is on children who are younger than 5 years. They are trained to identify whether a child has diarrhea, malaria, or pneumonia and to decide whether or not the child can be treated by the health worker or whether it needs to be referred to the PHU. First, health workers assess the child for general dangers signs (e.g., convulsions or the child being unable to breastfeed or drink) which would require an immediate referral to the PHU. Second, they assess the child for the three conditions above (e.g., they count the breaths per minute and compare this to age-specific threshold values in order to assess a child for pneumonia) and decide whether or not the child requires treatment and whether or not the child needs to be referred to the PHU. Health workers also should always assess children for malnutrition.

(vi) Follow-up visits of sick patients: For sick children that were not referred to the PHU, health workers are supposed to do at least two follow-up visits at the child's home on the third and sixth days after the start of the treatment. During these follow-up visits, health workers re-assess the sick child following the same steps as during the initial visit. They also should discuss the condition of the child with the caregiver and counsel the caregiver on disease-specific steps they need to undertake as well as general recommended health behaviors (e.g., hand washing or bed net use).

(vii) Routine household visits: First, health workers introduce themselves and the purpose of the visit. Second, they use the Family Health Card and assess previously agreed upon actions as well as current household health practices with the family. Third, health workers present new health information (e.g., on topics like hand-washing and sanitation, bed net use, or family planning) to the family. Finally, health workers and families identify barriers together and agree on an action plan until the next routine household visit by the health worker.

B.2 Context: Supervision

Supervisors have three main tools to train and advise health workers:

(i) Monthly trainings: Supervisors host a monthly meeting at the PHU which all health workers under their supervision are supposed to attend. During these trainings, supervisors provide information on health knowledge (how to prevent diseases, recognize dangerous signs). Central to these monthly meetings is the facilitation of mutual learning among health workers. They are asked to share both successes and barriers they experienced during their work in the previous month. Depending on the number of affected health workers, supervisors help them individually or collectively find solutions for the barriers that have been identified. This often involves re-training health workers on the checklists mentioned above or advising them on effective communication strategies health workers can use with households.

(ii) One-to-one trainings: Supervisors are asked to visit each health worker under their supervision in their village once per month. During these field visits, supervisors go through the records of health workers and randomly select three recent households the health worker provided a service to. For each of these three cases, supervisors ask the health worker about the detailed actions the health worker took and validate whether the steps on the checklists mentioned above have been followed. Supervisors then provide detailed feedback in which they identify gaps in the health worker's knowledge and explain again in detail how to provide the health services correctly.

(iii) In-the-field supervision / direct observation: Supervisors are asked to accompany the health worker to household visits and directly observe how the health worker conducts the visit. During these household visits, supervisors identify both the strengths and weaknesses of the health worker and raise awareness about the importance of her work with the family. After the household visit, supervisors provide personal feedback to the health worker in private.

B.3 Intervention: Choice of the Treatments

Theoretically, the set of possible splits an organization can select from is larger 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 could not test the effect of a wider set of 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 this split was the most natural in our setting. More precisely, we asked these informants 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 75% of the incentive (1,500 SLL), and the remaining 8% chose another split. In line with this, our structural model confirms that the optimal split is indeed very close to the 50%-50% one: see Section 6.

B.4 Intervention: Location of the Experiment and Randomization

This section discusses the location of the experiment, the randomization procedure and provides details on promotion incentives.

Location. Our experiment takes place in 372 PHUs across six districts of Sierra Leone. One district is located in the south (Bo), one in the east (Kenema), three in the north (Bombali, Tonkolili and Kambia) and one in the west (Western Area Rural). Out of the existing 823 PHUs across the six districts, we excluded half because no up-to-date and verified list of health workers was available, and selected 372 PHUs from the remaining eligible PHUs to be part of the experiment.

Randomization. The 372 PHUs were randomized into four groups of equal size: the worker incentives treatment, the supervisor incentives treatment, the shared incentives treatment, and the control group. The randomization was stratified by: (1) district, (2) average distance between the residence of the supervisor and the health workers in the PHU being above or below the median, and (3) the number of health workers in the PHU being above or below the median. We stratify by these variables because these are key predictors of our main outcome variables. These variables were measured at baseline by surveying the supervisor and the health workers before the randomization took place.

Promotion Incentives. A random sample of 2,081 health workers out of the 2,970 health workers in this study experienced a change in the promotion system. More specifically, six months after the start of the experiment which is the focus of this paper, the promotion system became meritocratic in 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). See Deserranno, Kastrau, and León-Ciliotta (2021) for more details.

Table A.19 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 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's fixed wage. Table A.20 moreover shows that the effects of our incentives treatments persist if we restrict the analysis to the sub-sample of health workers that did not take part in this separate study.

B.5 Intervention: The Reporting System

The reporting system works in 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.[49] All health workers in our study (including those in the control group) are asked to report their visits. The incentive was only paid for household visits that fall in one of these categories: (i) prenatal visits to a pregnant woman, (ii) accompanying a pregnant woman to the PHU for childbirth, (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).

(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.

(iii) The SMS information is continuously back-checked by a team of 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.

All health workers were promised a fixed bonus of SLL 10,000 conditional on truthful reporting at the end of the experiment. Despite this, we show in the paper that the reporting rate is low in all treatments.

C Research Ethics

Following Asiedu et al. (2021), we detail key aspects of research ethics.

C.1 IRB

The project received IRB from the Universitat Pompeu Fabra (Parc de Salut MAR: 2018/7834/I), Northwestern University (ID: STU00207110) and from the Sierra Leone Ethics and Scientific Review Committee (no IRB number was assigned by this local institution).

[49] When the patient is a child, the health worker reports the name and phone number of the primary care giver. When the household does not have a phone, the health worker reports the phone number of a neighbor.

We obtained informed consent from all participants prior to the study. The consent form described the participants' risks and rights, confidentiality, and contact information. Research staff and enumerator teams were not subject to additional risks in the data collection process. None of the researchers have financial or reputation conflicts of interest with regard to the research results. No contractual restrictions were imposed on the researchers limiting their ability to report the study findings.

The interventions under study did not pose any potential harm to participants and non-participants. The intervention rollout took place according to the evaluation protocol. Our data collection and research procedures adhered to protocols around privacy, confidentiality, risk-management, and informed consent. Participants were not considered particularly vulnerable (beyond some households residing in poverty). Besides individual consent from study participants, consultations were conducted with local representatives at the district levels. All the enumerators involved in data collection were aware about implicit social norms in these communities.

The findings from this project were presented to the MoHS in Sierra Leone. No activity for sharing results with participants in each study village is planned due to resource constraints. We do not foresee risks of the misuse of research findings.

C.2 AEA RCT Registry

The study was pre-registered in the AEA RCT Registry with the number AEARCTR-0003345. This paper follows the pre-analysis plan closely. The outcomes variables we use in the paper were all pre-registered.

The only deviation from the pre-analysis plan is that we do not study treatment effects on the "number of hours that the workers/supervisors self-report dedicating to their job", as a measure of their effort. This is because self-reported hours exhibit limited variability in the data, likely because of a self-reporting bias (as mentioned by enumerators in the field). Moreover, self-reported hours worked by the health workers do not correlate with the average number of hours households report having been visited by the worker (number of visits × average visit length): the correlation is -0.019 and is not statistically significant. Similarly, self-reported hours worked by the supervisors only weakly correlates with the likelihood that a household reports having received a joint visit in which the supervisor was present. For these reasons, we refrain from using self-reported hours as outcome variables in our analysis and use instead an objective measure of output reported by a third party (health visits, as reported by households).

D Model Appendix

D.1 Set Up

This section solves the model under the assumption that $b_1 = b_2 = 0$, $c_1 = c_2 = c$, $m = 1$ and $\alpha = 1$. We will later relax these assumptions.

We first quickly summarize the simplified set-up. A supervisor (player 2) and a worker (player 1) exert efforts e_1 and e_2 to produce output y, where $y = e_1 + \gamma e_1 e_2$. Thus, output depends on the efforts of players 1 and 2 and on the level of effort complementarity (γ).

Effort is costly to both the worker and the supervisor, and we assume that the cost of effort is quadratic: ce_i^2 (with $c > 0$). Before the start of the game, a principal offers to pay p to the worker and $1-p$ to the supervisor for every unit of output produced, where $p \in [0,1]$. There are two time periods. In period 1, the supervisor chooses effort e_1 and offers a side transfer s to the worker for every unit of output produced. Contractual frictions increase the cost of the side transfer to the principal by a factor of z ($z > 1$). Transfers can only go from the supervisor to the worker: $s \geq 0$. In period 2, the worker observes e_1 and s, and chooses e_2.

The payoff of the worker is as follows:

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

And the payoff of the supervisor:

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

D.2 A Key Assumption

In what follows, we will make the following assumption about the strength of the effort complementarity:

Assumption 1: $\frac{8c^2}{z} > \gamma^2$; $c, \gamma \in \mathbb{R}^+$.

As it will become clear in the next section, this assumption guarantees that both agents exert positive efforts. We can show that the following claim is true.

Claim 0: If assumption 1 ($\frac{8c^2}{z} > \gamma^2$) holds; then, it is also true that:

a) $2c^2 - \gamma^2 p(1-p) > 0$

b) $8zc^2 - \gamma^2(1 + p(z-1))^2 > 0$

Proof:

The proof will be divided in two parts. First, we show that assumption 1 implies a). Then, we 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: $\frac{2c^2}{z} > \frac{\gamma^2}{4}$. Thus, by the above and the transitivity of the inequality this also implies that $\frac{2c^2}{z} > \gamma^2 p(1-p)$, and by $2c^2 \geq \frac{2c^2}{z}$ implies $2c^2 > \gamma^2 p(1-p)$ (what we wanted to show).

Part 2: First note that:

$$8zc^2 - \gamma^2(1 + p(z-1))^2 > 0 \iff \frac{8zc^2}{(1 + p(z-1))^2} > \gamma^2$$

Therefore, we want to show $\frac{8zc^2}{(1+p(z-1))^2} \geq \frac{8c^2}{z}$ since it is sufficient to show that Assumption 1 implies b):

$$\frac{8zc^2}{(1 + p(z-1))^2} \geq \frac{8c^2}{z} \iff z^2 \geq 1 + 2p(z-1) + p^2(z-1)^2$$

$$\iff z^2(1-p)(1+p) \geq 2zp(1-p) + (1-p)^2 \iff z^2(1+p) - 2zp - (1-p) \geq 0$$

The quadratic function $z^2(1+p) - 2zp - (1-p)$ has roots $z_1 = 1$ and $z_2 = \frac{p-1}{p+1} < 0$, taking negative values between the two (in $(\frac{p-1}{p+1}, 1)$) and weakly positive elsewhere. Since $z \geq 1$, this means that for all values of z, $z^2(1+p) - 2zp - (1-p) \geq 0$ and so $\frac{8zc^2}{(1+p(z-1))^2} \geq \frac{8c^2}{z}$.

D.3 The Model: Main Analysis

We solve the model by backward induction:

Period 2:
The maximization problem of the worker in the second period is:

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

Thus, her optimal level of effort is:

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

Period 1:
Player 2 anticipates the optimal action of player 1 in period 2. Thus the maximization problem of player 2 is:

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

Thus, the optimal effort and side transfer are:

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

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

Let us first focus on the case where $p \leq \frac{1}{1+z}$. In this case, the side transfer is strictly positive and the optimal effort of the supervisor is given by::

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

Plugging e_2^* into e_1^*, we get:

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

In this case, the output y is given by:

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

We then consider the case where $p > \frac{1}{1+z}$. Now the side transfer is censored at zero. Optimal efforts are given by:

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

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

And so output is given by:

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

What level of p maximises output?

Suppose the principal wants to set p to the level that maximizes output. This maximization problem is divided in two parts: first, we maximize y assuming that $p \leq \frac{1}{1+z}$; then, we maximize y assuming that $p > \frac{1}{1+z}$.. We will refer to the first part of the problem as the "left-hand side" problem (or LHS problem for brevity), and to the second part of the problem as the "right-hand side" problem (or RHS problem for brevity). Also, we will use $p_a^* = p^*(p \leq \frac{1}{1+z})$ to denote the level of p that maximizes output in the LHS problem and $y(p_a^*)$ as the level of output when $p = p_a^*$. We will use p_b^* and $y(p_b^*)$ symmetrically to denote the level of p that maximizes output in the RHS problem, and the corresponding level of output. After solving the two problems, we compare $y(p_a^*)$ to $y(p_b^*)$. If $y_a^* > y_b^*$ ($y_a^* < y_b^*$), the solution to the overall problem is given by p_a^* (p_b^*).

We now solve the LHS problem:

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

The derivative of the objective function with respect to p is given by:

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

Assumption 1 implies that this derivate is positive for any value of p. To see this, note that (i) $c > 0$ and $z > 1$ (which guarantee that the numerator is positive), and (ii) the second part of Claim 0 shows that Assumption 1 implies that $8zc^2 - \gamma^2(1+p(z-1))^2 > 0$ for any

p, such that the denominator is positive for any level of p.

This shows that, as long as $p \leq \frac{1}{1+z}$, output grows in p. Thus, the LHS problem is solved by choosing the largest possible value for p: $p_a^* = \frac{1}{1+z}$.

To find the solution to the RHS problem, we solve:

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

In this case, the optimal p is given by the solution to:

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

Claim 0 shows that Assumption 1 implies that $2c^2 - \gamma^2 p(1-p) > 0$,. Thus, in the RHS problem, the optimal p is 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:

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

Interestingly, p_b decreases with γ, as can be seen from the derivative of p_b^* with respect to γ:

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

In order for p_b to be the global maximum of the RHS problem, we need to ensure that (i) $\frac{d^2y}{d^2p} < 0$ (the second derivative is negative), (ii) 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 (iii) that $\frac{1}{6} + \frac{\sqrt{\gamma^2+24c^2}}{6\gamma} \leq 1$. We tackle each one of these requirements in turn:

- 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 and $8c^2 \geq \frac{8c^2}{z}$.

- 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 $\gamma^2 < 8c^2$ (again, implied by Assumption 1 and $8c^2 \geq \frac{8c^2}{z}$).

- The condition $\frac{1}{6} + \frac{\sqrt{\gamma^2+24c^2}}{6\gamma} \leq 1$ is equivalent to $c^2 \leq \gamma^2$. That is, the complementarity has to be high enough for a two-sided incentive to be generate higher output compared to a one-sided incentive paid to the worker.

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

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

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

And the corresponding levels of output are:

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

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

.

The optimal p is found by comparing $y(p_a^*)$ to $y(p_b^*)$.

D.4 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 $p = 1$:

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

We have that:

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

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 $p = 0$:

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

We have that:

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

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 c and the contracting cost of the supervisor z. 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 responsive to γ.

D.5 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 ($s \geq 0$) as long as $p \leq \frac{1}{1+z}$. As $z = 1$, this condition simplifies to $p \leq \frac{1}{2}$.

On the other hand, the worker exerts effort:

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

Let us then analyze the maximization problem of the principal:

- If $p \leq \frac{1}{2}$ and so $s = \frac{1-2p}{2}$, then $y = \frac{1}{4c}$. This is independent of p; that 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:

$$\max_p \frac{p}{2c}$$

  The 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 does not contribute directly to production.

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

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

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 candidates:

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

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

Again, $p^* = 1$ 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 is 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 $s > 0$:

$$e_2^* = \frac{\gamma}{8c^2 - \gamma^2}$$

$$e_1^* = \frac{2c}{8c^2 - \gamma^2}$$

$$y = \frac{16c^3}{(8c^2 - \gamma^2)^2}$$

- When $p > \frac{1}{2}$ 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 is 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 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.6 Proof of Result 1

As before, we assume that Assumption 1 ($\frac{8c^2}{z} > \gamma^2$; $c, \gamma \in \mathbb{R}^+$) holds.

Result 1: When effort complementarity is lower than a threshold 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 ($p^* \in (0,1)$). 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.

Proof: To prove this statement we will first separately prove the following claims (given assumption 1):

Claim 1. The interior solution to the left-hand side problem ($\max_{p \leq \frac{1}{1+z}} y$) is strictly optimal when there are contractual frictions ($z > 1$). Otherwise, any $p \leq \frac{1}{1+z}$ leads to the same level of output.

Claim 2. When $\gamma^2 > c^2$, the principal's maximization problem always has an interior solution.

Claim 3. There exists a point $t = \frac{2c^2((1+z)^2-(1+z)^{\frac{3}{2}})}{z}$ such that for all γ such that $c^2 > \gamma^2 > 0$, $y(1) < y(\frac{1}{1+z})$ i.f.f. $\gamma^2 > t$; while $y(1) > y(\frac{1}{1+z})$ i.f.f. $\gamma^2 < t$.

Proof of Claim 1: When solving the model, we showed that the solution to the principal's left-hand side (LHS) problem, that is, $\max_{p \leq \frac{1}{1+z}} y$ has a unique global solution $p^* = \frac{1}{1+z}$ when $z > 1$ and multiple solutions, namely any $p \leq \frac{1}{1+z}$ when $z = 1$. This follows from the derivative of the objective function (y) with respect to p, which is increasing in p whenever $z > 1$ and is flat and equal to 0 whenever $z = 1$:

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

Proof of Claim 2: As explained above, $p^* = \frac{1}{1+z}$ is a global (not necessarily strict) solution to the principal's LHS maximization problem regardless the value of z. For the right-hand side (RHS) problem ($\max_{p > \frac{1}{1+z}} y$) we found that there is an interior solution (which is also the global solution to the RHS problem) whenever $\gamma^2 > c^2$. Therefore, there will always be an interior value $p^* \in (0,1)$ that solves the principal's problem (since the overall solution follows from the comparison of the value of output achieved under the solution to the LHS and RHS maximization problems).

Proof of Claim 3: First, note: $y(p=1) = \frac{1}{2c}$ and $y\left(p = \frac{1}{1+z}\right) = \frac{2c^3(1+z)^3}{(2c^2(1+z)^2 - \gamma^2 z)^2}$. Now we want to analyze when the following inequality is true:

$$y(p=1) > y\left(p = \frac{1}{1+z}\right) \iff \frac{1}{2c} > \frac{2c^3(1+z)^3}{(2c^2(1+z)^2 - \gamma^2 z)^2}$$

$$\iff (2c^2(1+z)^2 - \gamma^2 z)^2 > 4c^4(1+z)^3 \iff 4c^4(1+z)^3 - 4c^2(1+z)^2\gamma^2 + \gamma^4 z > 0$$

The LHS of the above inequality is a quadratic function in γ^2. Therefore, we solve for its roots to understand when it takes positive or negative values (that is, when the inequality holds) and we find the following two roots:

$$\gamma_1^2 = \frac{2c^2}{z}((1+z)^2 - (1+z)^{\frac{3}{2}})$$

$$\gamma_2^2 = \frac{2c^2}{z}((1+z)^2 + (1+z)^{\frac{3}{2}})$$

Then plugging in for some value of γ^2 in the middle of the two roots, e.g., $\frac{2c^2(1+z)^2}{z}$, we see that the quadratic function takes negative values:

$$4c^4(1+z)^3 - 4c^2(1+z)^2 \frac{2c^2(1+z)^2}{z} + \left(\frac{2c^2(1+z)^2}{z}\right)^2 z = -\frac{4c^4(1+z)^3}{z} < 0$$

This means that $4c^4(1+z)^3 - 4c^2(1+z)^2\gamma^2 + \gamma^4 z > 0$ i.f.f. $\gamma^2 \in (-\infty, \gamma_1^2) \cup (\gamma_2^2, \infty)$ and, conversely, $4c^4(1+z)^3 - 4c^2(1+z)^2\gamma^2 + \gamma^4 z < 0$ i.f.f. $\gamma^2 \in (\gamma_1^2, \gamma_2^2)$.

Finally, note that $c^2 \leq \gamma_2^2$, which is equivalent to $1 \leq \frac{2}{z}((1+z)^2 + (1+z)^{\frac{3}{2}})$, that is true for all $z \geq 1$ since $1 < \frac{(1+z)^2 + (1+z)^{\frac{3}{2}}}{z}$. This implies that $\forall \gamma^2 < c^2$ it is true that $4c^4(1+z)^3 - 4c^2(1+z)^2\gamma^2 + \gamma^4 z > 0$ (and so $y(1) > y\left(\frac{1}{1+z}\right)$) i.f.f. $\gamma^2 \in (-\infty, \gamma_1^2)$. And by analogy, $4c^4(1+z)^3 - 4c^2(1+z)^2\gamma^2 + \gamma^4 z < 0$ (and so $y(1) < y\left(\frac{1}{1+z}\right)$) i.f.f. $\gamma^2 \in (\gamma_1^2, c^2)$. Noting that $\gamma_1^2 = t$ completes the proof of Claim 3.

We showed that if $c^2 > t > \gamma^2$, then $y(1) > y(\frac{1}{1+z})$. Since the only two candidates for being the global optimum of y with respect to p when $c^2 > \gamma^2$ and $z > 1$ are precisely $p = 1$ and $p = \frac{1}{1+z}$, under contractual frictions ($z > 1$) the global optimum is attained when $p = 1$. In addition, since under $z = 1$ $y(\frac{1}{1+z}) = y(0)$, as shown in the special case in Section D.5; $y(1) > y(\frac{1}{1+z})$ also implies that $y(1) > y(0)$, such that when $c^2 > t > \gamma^2$ and $z = 1$, $p = 1$ is still the global maximum. This shows: "When effort complementarity is lower than a threshold 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 ($p^* \in (0, 1)$)." follows from Claim 2 when $\gamma^2 > c^2 > t$ and from Claim 3 when $c^2 > \gamma^2 > t$. On the other side, "If there are contractual frictions, this optimal two-sided scheme is the unique optimal scheme." follows from the previous discussion together with Claim 1.

Finally, the last statement: "If there are no contractual frictions $p^* = 0$ may also be optimal." is directly proved in the special case in Section D.5 where $z = 1$ and $\gamma > 0$.

D.7 The Model with Heterogeneity

In this final section we extend the model to allow workers and supervisors to have different costs and benefits. Output is now given by: $\alpha e_1 + \gamma e_1 e_2$. Further, we assume that the cost of effort is given by: $c(e_1) = c_1 e_1^2$, $c(e_2) = c_2 e_2^2$. Moreover, both players get a different benefit (b_1 and b_2) for each unit of production. Finally, the payment per unit of output is given by m.

The payoff of the worker will look as follows:

$$\pi_1 = (\alpha e_1 + \gamma e_1 e_2)(b_1 + mp + s) - c_1 e_1^2$$

And the payoff of the supervisor:

$$\pi_2 = (\alpha e_1 + \gamma e_1 e_2)(b_2 + m(1-p) - sz) - c_2 e_2^2$$

Let us solve the model by backward induction:

Period 2:
The maximization problem of the worker in the second period is:

$$\max_{e_1} (\alpha e_1 + \gamma e_1 e_2)(b_1 + mp + s) - c_1 e_1^2$$

Thus, the worker's optimal level of effort is:

$$e_1^* = \frac{(b_1 + s + mp)(\alpha + \gamma e_2)}{2c_1}$$

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

$$\max_{e_2, s} \frac{(b_1 + s + mp)(b_2 + m(1-p) - sz)(\alpha + \gamma e_2)^2}{2c_1} - c_2 e_2^2$$

Thus, the optimal effort of player 2 and the optimal side transfer are:

$$e_2^* = \frac{\gamma\alpha(b_1 + s + mp)(b_2 + m(1-p) - sz)}{2c_1c_2 - \gamma^2(b_1 + s + mp)(b_2 + m(1-p) - sz)}$$

$$s^* = \begin{cases} \frac{(b_2+m) - zb_1 - mp(z+1)}{2z}, & p \leq \frac{b_2+m-zb_1}{m(z+1)} \\ 0, & p > \frac{b_2+m-zb_1}{m(z+1)} \end{cases}$$

Let us first focus on the case where $p \leq \frac{b_2+m-zb_1}{m(z+1)}$. In this situation:

$$e_2^* = \frac{\gamma\alpha\eta^2}{8zc_1c_2 - \gamma^2\eta^2}$$

where $\eta = b_1 z + b_2 + m(1 + p(z-1))$.

And plugging e_2 into e_1:

$$e_1^* = \frac{2\alpha\eta c_2}{8zc_1c_2 - \gamma^2\eta^2}$$

In this case, the output y as a function of p is:

$$y = \frac{16\alpha^2 c_1 c_2^2 z\eta}{(8zc_1c_2 - \gamma^2\eta^2)^2}$$

In the case in which $p > \frac{b_2+m-zb_1}{m(z+1)}$, we will assume that $s = 0$:

$$e_2^* = \frac{\gamma\alpha(b_1+mp)(b_2+m(1-p))}{2c_1c_2 - \gamma^2(b_1+mp)(b_2+m(1-p))}$$

$$e_1^* = \frac{\alpha(b_1+mp)c_2}{2c_1c_2 - \gamma^2(b_1+mp)(b_2+m(1-p))}$$

And so the output is:

$$y = \frac{2\alpha^2 c_1 c_2^2(b_1+mp)}{(2c_1c_2 - \gamma^2(b_1+mp)(b_2+m(1-p)))^2}$$

Implications

There are at least two implications of this model's extension. First, the condition for positive side payments is now $p \leq \frac{b_2+m-zb_1}{m(z+1)}$. This condition becomes harder to satisfy as z grows and as $b_2 - b_1$ shrinks. Second, as long as side payments are positive, output is $y = \frac{16\alpha^2 c_1 c_2^2 z\eta}{(8zc_1c_2-\gamma^2\eta^2)^2}$. When $z = 1$, output is not a function of p: all levels of p result in the same level of output. On the other hand, when $z > 1$, output is a function of p.

E Prediction Survey Appendix

In collaboration with the Social Science Prediction Platform,[50] we invited social scientists to forecast how our treatments affect household visits compared to the control group. The participants made their forecasts before the results of this study were made public. Participants were paid to participate in the survey. 90% of the participants are economists; 41% of whom are faculty members and 45% are graduate students.

Participants were asked to forecast the average number of household visits health workers conduct in T_{worker}, T_{supv}, and T_{shared} after giving them a 700-words description of the study — i.e., the organization, the role of health workers, the role of supervisors — and informing them about the average number of household visits and its standard deviation for control group workers:

"We are interested to hear your predictions about the effects of the different incentive schemes on the main outcome variable, the number of household visits conducted by the community health worker in the previous 6 months as reported by the household's female primary caregiver during the endline household survey. Control Group Reference: As a

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

reference point, community health workers in the control group conducted on average 5.3 visits per household in the 6 months preceding the endline survey, with a standard deviation of 5.6. We would like you to predict the number of visits that the health workers conducted in the other three experimental conditions: How many visits do you think the health workers carried out when the 2,000 incentive was paid in full to the community health worker? How many visits do you think the health workers carried out when the 2,000 incentive was paid in full to the supervisor? How many visits do you think the health workers carried out when the 2,000 incentive was shared equally between the community health worker and the supervisor?"

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.