The Allocation of Incentives in Multi-Layered Organizations: Evidence from a Community Health Program in Sierra Leone*

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

January 12, 2024

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 healthcare provision is highly affected by how incentives are allocated between frontline workers and their supervisors. Sharing incentives equally between these two layers raises completed health visits by 61% compared to the unilateral allocations that are typical in public-health organizations. Also, the shared incentives uniquely improve overall health service provision and health outcomes. We provide reduced form and structural evidence that these results are driven by a combination of effort complementarities and contractual frictions, and we explore the implications of these forces for the optimal design of incentive policies in multi-layered organizations.

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

*We thank Michael Callen, Ernesto Dal Bó, Frederico Finan, and Rachel Glennester for insightful suggestions. We are also grateful to Oriana Bandiera, Michael Best, Katherine Casey, Zoe Cullen, Jonas Hjort, Paul Niehaus, Yusuf Neggers, and Alessandro Tarozzi. 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. Luz Azlor, Andre Cazor Katz, Margaux Jutant, Mustapha Adedolapo Kokumo, and Raquel Lorenzo Vidal provided outstanding research assistance. This project would not have 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 (Barcelona School of Economics CEX2019-000915-S), funded by MCIN/AEI/10.13039/501100011033 and grants RYC2017-23172 and PID2020-120118GB-I00 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: Bocconi and 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. Understanding how incentives affect the inner workings of hierarchies and the interaction between their different layers are first-order questions in economics since the emergence of complex, hierarchical organizations plays a key role in economic growth and in the development of an efficient state (Caliendo, Monte, and Rossi-Hansberg 2015; Wilson 2019; Bandiera et al. 2022). Indeed, extensive research has studied whether raising the *level* of financial incentives — either at the upper or lower tier of the hierarchy — strengthen organizational performance (Finan, Olken, and Pande 2017; Bandiera et al. 2019). However, surprisingly little attention has been devoted to understanding the optimal *allocation* of incentives across the various layers of an organization. In other words, suppose an organization wants to introduce or raise financial incentives. How should these incentives be distributed across the levels of the hierarchy? Answering this question requires a deep understanding of the functioning of vertical organizations. Most importantly, in a hierarchy, workers' and managers' efforts often complement each other (Alchian and Demsetz 1972; Kremer 1993; Brynjolfsson and Milgrom 2013), and hence, no layer of the organization can be considered in isolation from the other layers. Additionally, individuals may partly or fully offset the initial allocation of incentives through side transfers, but we lack good evidence of whether this happens in practice or not. Whether the allocation of incentives is a first-order driver of organizational performance thus remains an open question in the literature.

In this paper, we show that the allocation of financial incentives across the hierarchy of a large public-health organization substantially affects the provision of healthcare 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 — measured here as 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. These findings might seem surprising under a Coasian view of organizations, which postulates that any incentive allocation should result in the same output level (Coase 1937, 1960). However, we argue that these results can be reconciled within an alternative framework, emphasizing (i) the strong complementarity between worker and supervisor effort, and (ii) the limited redistribution of the incentive due to contractual frictions. We leverage the experiment to estimate a structural model that quantifies the importance of these structural forces for the optimal design of incentive policies in multi-layered organizations.

The program we study is a large community-based health initiative designed to improve health service provision in Sierra Leone. It plays a critical role in the government's effort to boost health outcomes in a country with some of the world's highest infant and maternal mortality rates. The program is composed of health units, each of which comprises an average of eight health workers responsible for conducting health visits in households within their community, and one supervisor. Supervisors play a pivotal role in our context: they train and advise the workers, provide them with the necessary skills to perform health visits, and help them build trust in the community. Their support is crucial to boosting the quantity and the quality of health visits provided by the worker.[1]

The field experiment we design creates random variation in the recipient of a new piece-rate scheme across 372 health units of Sierra Leone, covering roughly half of the country. 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. To measure the impacts of these interventions, we rely on extensive survey data on completed health visits (quantity and quality) and health outcomes collected from a random sample of households in each village where the program operates.

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 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 64% 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 offset by a decrease in visit duration, nor by less pro-poor targeting. On the contrary, we find that health workers in the shared incentives treatment receive more extensive training from their supervisors, leading to enhanced knowledge of how to provide health services effectively (a result we revisit below). Moreover, a higher share of households report trusting the health worker in the shared incentives treatment than in the other two treatments. 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).

Our findings also indicate that shared incentives yield the most significant improvements in health outcomes. Pregnant or expecting women are more likely to report having received a minimum of four pre-natal visits from any healthcare provider and delivering in a health facility (as opposed to at home) in the shared incentives treatment than in the one-sided

[1] Such potent complementarities would not exist if the supervisor's role were merely confined to monitoring. Therefore, our study diverges from recent literature that focuses on the surveillance 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. 2022).

incentives treatments or the control. Households in this treatment also report a reduced occurrence of fever among children under the age of five, coupled with an enhanced understanding of disease prevention measures. These findings are substantiated by administrative records from local health facilities, which reveal a higher count of services for pregnant women, institutional births, and fully immunized infants in the shared incentives treatment.

In the second part of the paper, we study the mechanisms explaining the large boost in output generated by shared incentives. To guide the 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 first period of the model, 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 visits delivered at the end of the game. In the second period, the worker chooses how much effort to exert to provide visits. The key intuition of the model is that two-sided incentives schemes such as our shared incentive intervention are optimal when (i) worker and supervisor efforts are strategic complements, (ii) the supervisor is relatively ineffective at motivating the worker through side payments due to contractual frictions. We will provide qualitative evidence on the drivers of effort complementarity and contractual frictions in Section 2.

We provide empirical evidence in support of these two mechanisms. We first focus on the strategic complementarity of effort. We show that shared incentives generate the same increase in supervisor effort as supervisor incentives. This might appear surprising since the *direct* incentive offered to the supervisor in the shared incentives treatment is half the magnitude of that in the supervisor incentives treatment. However, when efforts are strategic complements, the supervisor is *indirectly* incentivized by schemes that raise worker effort. Second, we find that shared incentives generate a larger increase in visits when effort complementarity is plausibly higher due to the lower level of experience of the worker. Third, we carry out a formal mediation analysis which shows the important mediating role of supervisor effort in the worker incentives treatment.

We then explore contractual frictions. Using descriptive data, we document that in our context, contractual frictions arise from the difficulty of making binding commitments (as in Casaburi and Macchiavello 2019), and the prevalence of social norms constraining redistribution, effectively maintaining payments within the organizational layer to which they were initially allocated (as in Hines and Thaler 1995). Furthermore, we leverage detailed survey data on inter-personal 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. Our key piece of evidence indicating that the use of transfers is indeed limited because of contractual frictions is that supervisor welfare, which we calculate using our model and a choice task, is higher under shared incentives than under supervisor incentives — a finding that is difficult to reconcile with alternative explanations for the limited use of transfers.

We rule out three alternative explanations of our results. First, that fairness concerns limit the effectiveness of the worker and supervisor incentives treatments (Breza, Kaur, and Shamdasani 2018; Cullen and Perez-Truglia 2022). This is contradicted by the substantial supervisor effort we observe in the worker incentives treatment, the lack of workers' awareness about supervisor pay, and the absence of treatment effects on job satisfaction. Second, the notion that positive reciprocity or a sense of team spirit is the key driver of the effectiveness of shared incentives. Again, this explanation is not consistent with our data on job satisfaction and job perceptions. Third, the argument that shared incentives are effective because the returns to additional incentives fall rapidly due to sharply diminishing marginal utility or increasing marginal costs around the 1,000 SLL cutoff. This explanation contradicts our analysis of treatment effect heterogeneity based on proxies of utility and costs, as this analysis does not uncover any evidence supporting the hypothesized sharp fall in incentive effectiveness.

In the final 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 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. We estimate that the marginal return to worker effort is up to 114% higher due to the complementarity with supervisor effort. Further, our calibrated contractual friction parameter implies that difficulties in contracting increase the cost of side transfers by more than threefold. We derive three lessons on optimal policy based on counterfactual policy analysis. First, we calculate that the optimal policy would offer 61% of the value of the incentive to the worker, and 39% to the supervisor. Second, we find that the optimal allocation of the incentive is quite sensitive to the exact value of the complementarity parameter, which emphasizes the importance of re-calibrating the policy in new contexts. Third, we show that, due to the complementarity of agents' efforts, interventions that tie incentives to joint output are more effective than interventions that incentivize effort directly, since they help agents internalize the impact of their effort on the productivity of their partner.

This paper contributes to four strands of the literature. First, we show that the allocation of incentives in hierarchies 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 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.[2] Our findings reveal that agents engage in very limited fine-tuning of the allocation of incentives through transfers, due to the presence of contractual frictions.[3] 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 deepen our understanding of hierarchical organizational structures and pinpoint an overlooked policy approach to enhance public-sector effectiveness in developing countries (Callen et al. 2023; Bandiera et al. 2019; Finan, Olken, and Pande 2017).

Second, we provide evidence on the *productive* role of middle managers in hierarchical organizations. This adds to the literature demonstrating the importance of management practices (e.g., Bloom et al. 2013; McKenzie and Woodruff 2017; Macchiavello et al. 2020; Adhvaryu, Murathanoglu, and Nyshadham 2023). Notably, our paper relates to but differs from the well-established literature focusing on the monitoring responsibilities of managers. This body of work, 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. 2022; Rasul and Rogger 2018; Kala 2019), examines how to optimally delegate authority and prevent harmful collusion between workers and supervisors. However, this literature remains mostly silent on how supervisor effort can directly increase the returns to worker effort. In our experiment, we intentionally 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 the implications of this complementarity for the design of incentives. Recent findings on the ripple effects of training interventions throughout the organizational hierarchy echo our focus on the productive role of public-sector managers (Espinosa and Stanton 2022).

Third, we extend the literature on effort complementarities within organizations. Seminal theoretical work by Alchian and Demsetz (1972); Itoh (1991); Milgrom and Roberts (1995); Ray, Baland, and Dagnelie (2007); Brynjolfsson and Milgrom (2013) has explored the implications of complementarities for designing incentives. Empirically, several studies

[2] 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). Behrman et al. (2015) study who should be incentivized for students' test scores in Mexican schools (students, teachers or administrators) and which structure should these incentive have, but do not shed light on the mechanisms. Geng (2018) complements Behrman et al. (2015) by providing evidence that supports the presence of effort complementarities between students and teachers.

[3] 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).

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 they encourage free-riding (Muralidharan and Sundararaman 2011; Babcock et al. 2015; Friebel et al. 2017). However, unlike our paper, this literature does not address the optimal allocation of such incentives. This is partly due to the challenge of offering asymmetric incentives to workers performing similar tasks within "horizontal" teams, often due to fairness concerns (Card et al. 2012; Breza, Kaur, and Shamdasani 2018) or inflexible contractual arrangements. In contrast, within "vertical" teams, asymmetric incentives are more viable as workers across different layers of the organization possess distinct responsibilities and experience levels.

Our findings carry significant policy implications. They contribute to the understanding of how to expand access to healthcare in low-income countries, a pivotal objective of global public policy (Dupas and Miguel 2017; Roser 2021). The World Health Organization (WHO) estimates that half of the world's population lacks coverage for essential health services, and the burden of health expenses is so severe that it plunges over 100 million people into extreme poverty (WHO 2021). In the context of a low-income, post-conflict nation with one of the world's highest infant mortality rates (Casey and Glennerster 2016), our study demonstrates that healthcare access and health outcomes can be substantially improved through adjustments in the allocation of incentives, without altering their overall level.

Our results also hold policy implications for the design of decentralized development programs. These programs employ "last-mile" frontline workers to provide essential public services to their communities, and have become a primary mode of service delivery in low-income countries since the 1990s (Bandiera et al. 2023). A majority of these programs predominantly incentivize frontline workers in a unilateral manner, often overlooking the importance of supervisors. Perry (2020) demonstrates, for instance, that merely 1.7% of community health worker programs (similar to the one examined in this study) extend incentives to supervisors, while most concentrate on community health workers. In line with this, many of the social scientists invited to forecast our results on the Social Science Prediction Platform expected worker incentives to maximize output, despite being informed of the key role played by supervisors in our context.[4] Contrary to to conventional policy approach and expert predictions, our study highlights that last-mile service delivery can be enhanced by reallocating a portion of the incentives from frontline workers to supervisors.

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

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 (WHO 2017). Such elevated mortality rates have been attributed to a critical shortage of health workers, together with limited access to health facilities throughout the country (WHO 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. Attrition is limited in our setting, both for health workers and supervisors. The median health worker has 4 years of experience and has known the supervisor for almost the same duration.

Role of the health workers (bottom layer) The role of the health workers is to provide a package of basic healthcare services in their community. They do so by making home visits to expecting mothers or mothers who recently gave birth, during which they provide: (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 (described in Appendix B.1) each time they provide a service.

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 by organizing monthly trainings, which cover vital health topics, such as diagnosing, treating, and recognizing danger signs for referral to health facilities.[5] Importantly, they also provide "in-the-field" training and guidance 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 "on-going" on-site training. 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.

Complementarities across layers Supervisors stimulate demand for health services by building trust towards the health workers in the community. This is particularly important because community members may initially have doubts about the expertise of the health worker — who is typically known by the community as a farmer or a shopkeeper — and this may hinder the utilization of the worker's services. The supervisor plays a key role in transferring health knowledge to the worker and legitimizing her position in the eyes of the community, which can boost both the quantity and the quality of the household visits provided by the worker. 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. Likewise, a supervisor's effort yields higher returns when the worker is motivated and fully leverages the increased demand for their services, which the supervisor helped create.

It is worth noting that such strong complementarities would likely not exist if the supervisor's role was merely confined to monitoring. This distinctive aspect differentiates our paper from recent literature that predominantly 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. 2022).

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 2,000 SLL incentive is paid entirely to the health worker who provides the visit.[6] In the supervisor incentives treatment (T_{supv}), the 2,000 SLL

[5] Supervisors offer two types of trainings: general trainings, which are provided to all health workers at the local health facility one day per month, and one-to-one trainings, which are given to each specific health worker in their respective village as needed. We provide details on the content, frequency and objectives of the trainings organized by the supervisors in Appendix B.2.

[6] 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.

incentive 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. Importantly, the three treatments diverge in terms of which layer of the organization receives the incentives while maintaining a constant incentive amount per reported visit. This allows us to answer the question of how a given incentive should be allocated across the layers of an organization.[7]

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, as detailed in Appendix B.4, 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, the average distance between the residence of the supervisor and the health workers in the PHU, and the number of health workers in the PHU. 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). We describe the change in the promotion system and show that our results are orthogonal to this variation in Appendix B.5.[8]

It is worth noting 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.[9] 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 their partner.

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.

[7] Appendix B.3 explains our rationale for equally dividing the incentive across layers in the shared incentives treatment (1,000 SLL each) and the process for deciding on the incentive amount.

[8] 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.

[9] 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.

This enabled us to establish the credibility of the new incentive scheme in the eyes of all experimental participants.[10]

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. For payment initiation, the SMS must include the service date and the patient's contact number, and must be sent from the worker's registered phone number. This condition ensures that neither supervisors nor households have the capability to report services on behalf of the workers.

All health workers participating in our study, including those in the control group, were instructed to report their visits and underwent training on using the reporting system. The training sessions were conducted at the PHU, involved the participation of all workers and their respective supervisors, and lasted two days. To maintain consistent focus on the training material across treatments, workers in T_{worker} and T_{shared} were informed of the incentives paid based on SMS reporting only after receiving information about the reporting system. By the end of the training, all workers were able to report a visit without the need for supervisor assistance. As a result, supervisors were not involved in assisting workers with the reporting process. See Appendix B.6 for more details on these trainings.

Our set-up discourages over-reporting through rigorous back-checks and strong penalties. A randomized 25% of reports undergo verification by contacting the household mentioned in the report. Should a worker be discovered reporting a non-existent visit, they would be disqualified from any further incentive payments and would also be reported to the Ministry of Health and Sanitation (MoHS). All workers, including those in the control group, were subjected to the same number of back-checks to ensure comparability across experimental groups.

We will later show that the threat of being caught "cheating" was credible, nearly eliminated over-reporting and no worker was disqualified from the incentive payments. However, our design does not entirely mitigate the risk of under-reporting. Despite the SMS reporting tool being cost-free, reporting inherently entails expenses in terms of time and the necessity to gather information such as patients' names and phone numbers — details they might not always be willing to share. The unreliability and unpredictability of mobile phone coverage in rural areas of Sierra Leone further exacerbates the challenges health workers face in promptly sending the SMS reports.[11]

[10] The external organization is a multi-service consulting firm that provides expert advice and services to the government, and which has worked extensively with them in the past.

[11] Low reporting rates are common in 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 presents evidence 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 2023, 2022), we informed all supervisors in the study about the worker incentives but did not inform the workers about the supervisor incentives.[12] This limits negative morale concerns resulting from pay inequality because 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 4.4).

We intentionally withheld from supervisors information regarding the quantity of SMSs sent by individual workers or information about workers' earnings under 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 and contractual frictions After supervisors were informed about the incentives allocated to them (if any), we clearly communicated to them that they were allowed to transfer their incentives, either wholly or partially, to workers. While we did not suggest a specific amount to be transferred, we informed the supervisors that these transfers could be viewed as potential incentives to stimulate worker effort (see Appendix B.7 for more details). Nonetheless, we will later show that, due to the presence of contractual frictions, less than 10% of the supervisors opted for such transfers.

In a recent follow-up survey, we asked supervisors why they opted for such low transfers. 75% of them indicated that "paying workers in exchange for more visits is inappropriate." This suggests the prevalence of social norms constraining redistribution and effectively maintaining payments within the organizational layer to which they were initially allocated — a phenomenon akin to the flypaper effects in Hines and Thaler (1995). Our context is moreover one where side agreements are informal, leaving the worker with limited recourse to penalize the supervisor for non-compliance with a side payment arrangement (for instance, the worker's threat to reduce future effort lacks credibility as the organization can penalize the worker for insufficient effort). As in many other contexts, making binding commitments is thus difficult in our setting (Bubb, Kaur, and Mullainathan 2018), and the supervisor may have to compensate the worker for the perceived risk of default In line with this, 55% of the

[12] Information about supervisor's incentives was provided privately to the supervisor at the end of the training on using the reporting system. See Appendix B.7 for more details.

supervisors we surveyed indicated that "health workers would lack trust in their payment."[13]

2.3 Data and Balance Checks

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 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 utilize two administrative data sources. Firstly, we record the number of valid SMS reports sent by each health worker during the experiment, linked to their corresponding incentive payments. This facilitates the tracking of reporting over time. Secondly, the MoHS provided us with data detailing the monthly number of health services rendered at each local health facility, encompassing institutional births, children immunizations and fever/malaria/diarrhea cases treated at the facility.

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).[14] 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 wealthy

[13] One supervisor stated, "I did not think about it," as the explanation for not transferring any funds, while no supervisor said that they did not transfer because "they deserve the money more than the health workers."

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

and educated than health workers and supervisors, with only 25% having completed primary school. 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 often unreliable. Panel E shows that health facilities record 47 pregnant women visits per month, 13 institutional births, 11 infants immunized, and 66 cases of malaria/diarrhea among under-five children.

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 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 Main Results

This section presents the main results of this paper. In Section 3.1, we present the results on household visits and health outcomes, our measures of output. In Section 3.2, we present the results on reported (as opposed to actual) visits and incentives payments.

We estimate the following regression equation:

$$Y_{ij} = \alpha + \beta_1 T_{worker,j} + \beta_2 T_{supv,j} + \beta_3 T_{shared,j} + Z_j + \varepsilon_{ij}, \tag{1}$$

where Y_{ij} is the outcome of interest for health worker i in PHU j (e.g., the number of household visits provided by the health worker, health outcomes in the community served by the health worker). $T_{worker,j}$, $T_{supv,j}$, 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. Z_j are the stratification variables discussed in Section 2.2. We estimate standard errors clustered by PHU (level of the randomization).

We report p-values corrected for multiple hypothesis testing using three alternative procedures in Table A.3.[15] For the sake of brevity, we do not discuss the corrected p-values in the main text as these are notably similar to the non-corrected ones we discuss below. Following Asiedu et al. (2021), we detail key aspects of research ethics in Appendix C. The outcome

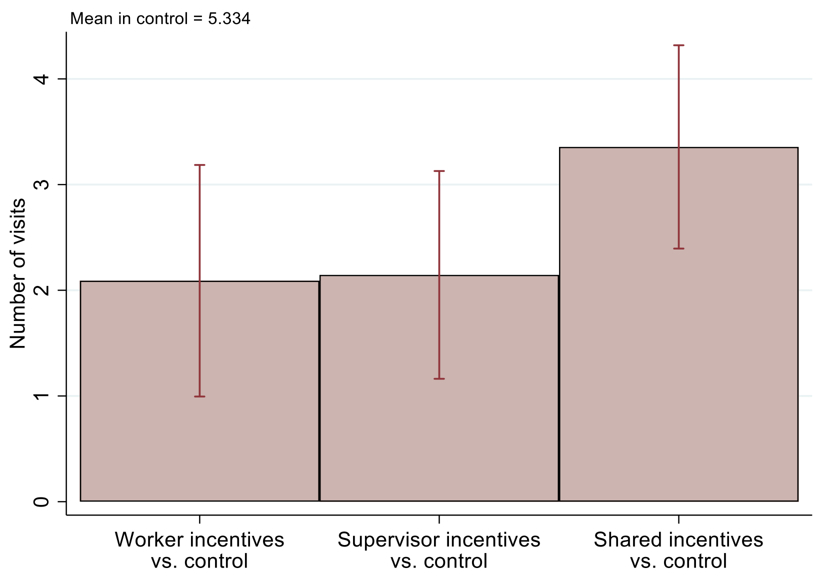
[15] The three procedures are: Bonferroni, Romano and Wolf (2016), and Benjamini, Krieger, and Yekutieli (2006). The Bonferroni procedure controls for the familywise error rate. This procedure is conservative, as it assumes that test statistics are independent. Therefore, we also present corrected p-values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

variables of interest and the research design were specified in the AEA RCT Registry.

3.1 Household Visits and Health Outcomes

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 may differ from the actual number of visits due to under-reporting, as discussed in Section 3.2. 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.[16] For each health worker, we then calculate the mean number of visits received by a household (mean of 7.3). The results are reported in Table 2 column (1) and Figure 1 here below.

Figure 1: 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.

We find that introducing performance-based incentives significantly boosts the number of household visits provided by the health worker, regardless of whether the incentives are one- or two-sided. The mean number of visits per household in the control group is 5.334. This number increases by 2.090 (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.

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

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 shared incentives scheme, which achieves 17% more visits overall. Relative to the control group, the boost in visits generated by the shared incentive schemes is 61% larger than the boost in either of the one-sided schemes. When we break down household visits by their type, we find that, compared to the one-sided treatments, shared incentives generate significant gains over both natal-related and disease-related visits (Table A.2).

Quality of the visits 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 a 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 3) 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 5). The latter is consistent with the health workers receiving more training from the supervisor, as further discussed in Section 4.2. 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 2) and the number of health topics discussed by 26% (column 4). This captures both the intensive and the extensive margin of effort. Importantly, the shared incentives also maximize trust: the fraction 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).

Targeting of the visits 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). Leveraging household-level regressions, we show in Table A.4 that households who are geographically and socially close to the health workers are visited more often, but that this is equally true in all treatments. This further alleviates concerns related to quantity-quality trade-offs and 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.2), and does not reduce the visit length (Table 2, column 2). Alternatively, health workers in the shared incentives treatment may increase the intensity to which they visit the same household rather than increasing coverage – i.e., the share of households ever visited by the health worker. Table 2 (columns 6 and 7) rejects this possibility. The shared incentives treatment increases both the coverage and the number of visits received by a given household.

Access to natal-care services and disease incidence We now test whether the increase in 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 WHO framework (four pre-natal visits, institutional birth, post-natal care within two days of delivery, up-to-date vaccination, breastfeeding).[17] 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). 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.[18] 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.5 (column 2). We find no significant effects on under-five mortality rates (Table A.5, column

[17] Pre-natal care is measured by asking women who gave birth in the year preceding the endline survey whether they underwent a minimum of four pre-natal visits from any healthcare provider. Post-natal care is gauged by querying if they delivered their child in a healthcare facility (as opposed to at home), whether they received a post-natal visit within two days of delivery, whether they breastfed their infant for a duration of at least six months, and whether their infants are in accordance with the vaccination schedule. In Table A.5 (column 1), we show that our treatments do not affect the likelihood that a household is composed of a woman who gave birth in the year preceding the endline survey.

[18] 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).

3), presumably due to the relatively short timeframe of the experiment. The results are robust to multiple hypothesis testing corrections (Table A.3).

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.6. 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 incentives 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.

3.2 Reporting and Incentives Payments

Reporting Recall that the incentive scheme pays 2,000 SLL per health visit *reported* by the health worker. The health worker receives the full amount in T_{worker}, half of it in T_{shared}, and none of it in T_{supv}. Therefore, workers in T_{worker} have the highest incentive to report the visits they carry out. In line with this, Table 4 (column 1) shows that workers in T_{worker} report the highest number of visits, even though our previous analysis established that the number of visits actually carried out is highest in T_{shared}. More precisely, in the six months before the endline survey, workers sent an average of 26.1 SMS reports per month in T_{worker}, 18.9 in T_{shared}, 11.2 in T_{supv}, and 4.6 in the control group.[19]

In Table 4 (column 2), we present results on the reporting rate, that is, the ratio between the number of SMS reports per month (column 1) and the actual number of visits per month that we compute using our household survey.[20] Visits are under-reported in all experimental conditions. Health workers in the worker incentives treatment have the highest reporting rate (90%). In the other treatments, health workers report a smaller share of the visits (going from 23.4% in the control group to 50.7% in the shared incentives treatment). This is consistent with the notion that reporting is costly, leading some workers to under-report when not sufficiently incentivized. Over-reporting is however minimal. Only 3.6% of the

[19] A previous version of the paper employed an incorrect method to calculate the number of SMS reports, deviating from standard field practice and yielding an erroneously low count. This issue has been rectified in the current version.

[20] The actual number of visits per month is calculated by scaling up the number of actual visits among the random sample of households we interviewed for the total number of households in the community. The resulting reporting rate might be over- or under-estimated for a single health worker due to sampling error, but 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.

health workers in the worker incentives treatment ever reported a visit that the recipient household did not confirm during the back-check, compared to 2.7% in the shared incentives treatment, and 2% in the supervisor incentives treatment.

Contrary to other variables in this study, we are able to track the progression of reporting over time. As demonstrated in Figure A.1, differences in reported visits across treatments remain relatively consistent over time. Provided that the reporting rate did not vary over time, this suggests that differences in actual visits across treatments remained stable during the course of the experiment.

Incentives payments The treatments also generated clear differences in incentive payments (Table 4, columns 3-5). Workers receive the largest incentive payments in T_{worker} (53 thousand SLL per month), followed by T_{shared} (18 thousand SLL), and by T_{supv} (0 SLL). Supervisors receive similar and statistically indistinguishable payments in T_{supv} and T_{shared} (22 and 19 thousand SLL per health worker they supervise).[21] Overall, the new incentive scheme costs the organization a total of 52 thousand SLL per health worker in T_{worker}, 38 thousand SLL in T_{shared}, and 22 thousand SLL in T_{supv}.

Cost-effectiveness We assess cost-effectiveness by calculating the number of actual visits generated for every 2,000 SLL spent on the incentive scheme. Our estimates suggest that, in comparison to the control group, the organization gains an additional 2.24 monthly visits per worker for each 2,000 SLL spent on T_{worker}. This figure rises to 3.2 visits for every 2,000 SLL spent on T_{supv} and increases further to 5.35 visits for each 2,000 SLL spent on T_{shared}.[22] Consequently, the shared incentives treatment is the most cost-effective.

Two channels account for the superior cost-effectiveness of the shared incentives. Firstly, it maximizes the actual number of visits rendered, thereby allocating a larger portion of the total payout to marginal visits instead of subsidizing infra-marginal visits that would have occurred even without incentives. Secondly, it is less costly than the worker incentives treatment as it incites a lower reporting rate, leading the organization to save on incentive payouts that would have otherwise been awarded to workers and supervisors. The latter might be an unwarranted driver of cost-effectiveness, and the organization might want to find a way to compensate workers for unclaimed benefits.

[21] This is because the higher number of reports in T_{shared} offsets the larger incentives paid out in T_{supv}. Note that supervisors receive payments that are considerably larger than those of workers, since each supervisor is responsible for several workers. On average, supervisors earn 182 and 151 thousand per month from the incentive scheme in T_{worker} and T_{shared}, respectively.

[22] We calculate the number of visits performed by the average worker in the treatment 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). We winsorize the top and bottom 1% of the outcome variable due to the presence of outliers and input the maximum value of the outcome variable for the few PHUs in which the total incentives payment is zero.

Overall, if the organization aims to maximize household visits or minimize cost per visit, it should choose shared incentives. Note, however, that shared incentives impose a larger total cost compared to supervisor incentives. If this cost exceeds the organization's budget constraint, the organization could either opt for supervisor incentives, which offer a similar increase in visits as worker incentives but at a lower cost, or they could reduce the incentive amount in the shared incentives scheme.

4 Mechanisms: Effort Complementarities and Contractual Frictions

The previous section showed that health workers provide significantly more household visits under shared incentives than under the one-sided incentives schemes, with no reduction in visit quality. We now study the mechanisms underlying this result. In Section 4.1, we present a theoretical framework that illustrates how the combination of complementarities and contractual frictions can make shared incentives uniquely effective. In Sections 4.2 and 4.3, we use the framework to motivate a number of empirical tests to assess the role of effort complementarity and contractual frictions in our context, respectively. Finally, in Section 4.4, we present empirical evidence against three alternative mechanisms that are not considered in our model but could explain why two-sided incentives outperform one-sided incentives: inequality aversion, reciprocity and sharp non-linearities in the utility or cost function.

4.1 Theoretical Framework

We consider the case of a single frontline worker (player 1), a single supervisor (player 2), and a principal.[23] 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 by training and advising the worker. At the beginning of the game, the principal sets the incentive scheme faced by the worker and the supervisor in order to maximize the number of visits. The worker and supervisor then make two sequential decisions. First, 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 the worker completes. Second, 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 training support offered to the worker is given in advance of the worker's choice of effort.

The supervisor may face a different cost of inducing worker effort compared to the principal. On the one hand, the supervisor's close proximity to the worker enables her to better

[23] In our empirical setting supervisors are responsible for multiple workers. We abstract from this feature since it would not change the basic intuitions discussed below.

observe worker effort compared to the principal, provide more personalized training and feedback, more effectively boost worker morale, and ultimately be more effective in motivating the worker than the principal. On the other hand, the supervisor may find it hard to establish side-contracts with the workers due to contractual frictions such as the difficulty of making binding commitments (e.g., the supervisor may need to compensate the worker for the perceived risk of default) and social norms against monetary transfers within an organization. To model these opposing forces in a simple way, we assume that a side payment of s costs the supervisor zs, with $z > 0$. z is a reduced form parameter that measures how costly the supervisor finds it to incentivize the worker relative to the principal. If contractual frictions dominate, then $z > 1$. If the motivation effect dominates, $0 < z < 1$. Finally, if contractual frictions and the motivation effect fully offset each other, $z = 1$.

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

$$y = \alpha e_1 + \beta e_2 + \gamma e_1 e_2 \tag{2}$$

where α is positive, β weakly positive, and $\alpha > \beta$. In what follows, without loss of generality, we set $\beta = 0$, since it is plausible that, when $e_1 = 0$, the supervisor cannot generate any visit no matter how much effort she spends training and advising the worker. 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.

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. Each agent i pays a convex cost of effort: $c(e_i) = c_i e_i^2$. Further, each player 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).[24] 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., $p = 1$ in T_{worker}, $p = 0.5$ in T_{shared}, and $p = 0$ in T_{supv}. The supervisor, on the other hand, is paid an incentive of $(1 - p)m$ per visit completed by the worker. 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.[25] In the basic version of the model, we assume that all visits are reported. We can easily incorporate under-reporting by introducing a random reporting shock. Under reasonable assumptions, this extension does

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

[25]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.

not qualitatively change the analysis of the model.[26]

In sum, the payoffs of the worker and of the supervisor are given by:

$$\pi_1 = (b_1 + pm + s) * y(e_1, e_2) - c(e_1) \tag{3}$$

$$\pi_2 = (b_2 + (1-p)m - zs) * y(e_1, e_2) - c(e_2). \tag{4}$$

The principal's problem is to choose the incentive scheme — the level of p — that maximizes household visits. We will refer to this scheme as the *"optimal"* scheme. Also, 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"*. In Appendix D, we formally solve the model and provide both effort functions and a side payment function.

The optimal incentive scheme We now derive three key insights on how contractual frictions and effort complementarities determine the optimal incentive scheme.

First, when the contractual frictions between the worker and supervisor are limited ($z < 1$), it may be optimal to assign the entire incentive to the supervisor ($p = 0$). In this scenario, the principal leverages the supervisor's ability to motivate the worker and transfers the entire incentive to the supervisor, anticipating that a portion of the incentive will be passed on to the worker as a side payment. This arrangement improves efficiency since $z < 1$, making it more cost-effective for the supervisor to incentivize the worker compared to the principal. We illustrate this case in Figure A.2.

Second, when contractual frictions are large ($z > 1$) and effort complementarities are limited ($\gamma \leq t$, where t is a threshold defined in the Appendix), it is optimal to allocate the entire payment to the worker ($p = 1$). In this scenario, the principal capitalizes on the worker's unique position to conduct household visits and utilizes the incentive payment to directly motivate worker effort. Providing any incentive to the supervisor would be inefficient, as the supervisor is both less productive and less effective in motivating the worker compared to the principal. We illustrate this case in Figure A.3.

Third, when contractual frictions are large ($z > 1$) and effort complementarities exceed a threshold ($\gamma > t$), the optimal scheme is two-sided. In this final scenario, the presence of

[26] Specifically, to model under-reporting, we assume that, after visits are carried out, each visit suffers from a random, idiosyncratic reporting shock with probability $1 - q$. This shock captures factors such as the temporary absence of mobile phone network, the loss of the household details required in the reporting form, or the worker forgetting to report that visit. When a shock occurs, the visit is not reported, and hence the incentive is not paid. Thus, in expectation, each visit is paid an incentive of mq. If we assume that the reporting rate is fixed across treatments, then the model is essentially unaffected, except for a change in the level of the incentive paid. Alternatively, we can assume that the reporting rate increases in p. This would raise the relative attractiveness of the worker incentives 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. Since, in the experiment, visits are maximized by the shared incentives treatment, it is likely that this threshold is not exceeded in the context we study.

large effort complementarities makes it optimal to elicit substantial effort from both players. To achieve this, the principal cannot rely on the supervisor due to the presence of contractual frictions. The optimal incentive scheme is thus two-sided, as illustrated in Figure A.4.

4.2 Empirical Evidence of Effort Complementarities

The model motivates a number of empirical tests to evaluate the existence of effort complementarities and contractual frictions. This section examines three pieces of evidence that support the presence of effort complementarities. The subsequent section presents the evidence on contractual frictions.

Evidence 1: Supervisor effort does not monotonically increase with supervisor incentives. Our first test aims to evaluate how supervisor effort varies with the level of supervisor incentives. Finding that supervisor effort does *not* increase monotonically with supervisor incentives would be consistent with strategic complementarity of efforts, since, when efforts are complements, a decrease in direct supervisor incentives can be compensated by the indirect incentive arising from increased worker effort (see the expressions for optimal supervisor effort in Appendix D).

In Table 5 (column 1), we confirm that one crucial aspect of supervisor effort, namely providing in-the-field training and guidance by accompanying health workers during household visits, does not display a monotonic increase with the incentive payment received by supervisors. Our findings reveal that the share of households who report having received a visit in which the health worker was accompanied by the supervisor (henceforth, "accompanied visit") rises by 6.2 percentage points (38%) in T_{shared}, and by 5.7 percentage points (35%) in T_{supv}, in comparison to the control group. Notably, the increase in supervisor effort is statistically indistinguishable in the two treatments (and, if anything, is higher in T_{shared}) despite the fact that the incentive per visit received by the supervisor is twice as large in T_{supv} compared to T_{shared}. This result aligns with the existence of sizable effort complementarities.[27]

Additionally, we show that workers improve their health knowledge only in the shared incentives treatment (Table 5, column 4).[28] Neither of the one-sided incentives treatments yields a significant improvement in worker's health knowledge, as this presumably requires a combination of worker and supervisor effort, which is only achieved by offering shared incentives. This further supports the presence of effort complementarities. Gaining health

[27] Columns (2) and (3) of Table 5 show that our treatments do not affect the number of one-to-one meetings the health workers report having received from the supervisor, or the likelihood that the supervisor invited them to attend a general training. The latter is unsurprising since supervisors are mandated to organize such trainings on a monthly basis, and 99.4% of them comply with this requirement.

[28] We administered a quiz on broad health knowledge to all health workers at baseline and endline, and estimate the difference in health knowledge over time. See Appendix B.8 for details on how we measure health knowledge.

knowledge is very important in our setting, since, as we have shown in Section 3, it enables workers to discuss a broader range of health topics during household visits, gain more trust within the community, increase demand for their services, and ultimately provide more households visits.[29]

Evidence 2: Shared incentives are less effective for highly-experienced workers. As a second test, we study whether the difference in output between the shared and worker incentives treatments is larger for workers who are less experienced. These workers are more likely to rely on the support of their supervisors to perform at their best, as they have limited training and are less recognized as health workers in the community. We would thus expect the strategic complementarity of effort to be higher for these workers. And, as we show formally in Appendix D, the stronger the effort complementarity, the larger the output difference between the shared incentives treatment (which we will show to be quantitatively close to the optimal two-sided incentives scheme in Section 5) and the worker incentives treatment.

We confirm that shared incentives outperform worker incentives by a larger margin when workers have less experience, and hence, effort complementarities are likely to be higher. Table A.7 presents treatment effects for workers with experience below and above the median (4 years). For less experienced workers, the treatment effect of shared incentives on household visits is almost double the size than the treatment effect of worker or supervisor incentives (column 1). Additionally, shared incentives stimulate a significantly higher increase in supervisor effort. For more experienced workers, the treatment effect gap between shared incentives and one-sided incentives is much smaller. Adding controls for worker characteristics correlated with experience does not qualitatively affect these results (columns 2 and 4 of Table A.7). Overall, these findings further substantiate effort complementarities as a pivotal mechanism in our experiment.

Evidence 3: Important mediating role of supervisor effort in the worker incentives treatment. To strengthen the evidence for effort complementarities, we utilize a mediation analysis to demonstrate how increased supervisor effort stimulates higher worker effort and increases worker productivity. We do not directly observe worker effort, but we can exploit the fact that we measure supervisor effort and output (visits). Following Acharya, Blackwell, and Sen (2016)'s methodology, we calculate the Controlled Direct Effect (CDE) of the worker incentives treatment on household visits. This gauges the impact of worker incentives on

[29]Note that the upswing in accompanied visits in T_{shared} and T_{supv} explains only a small portion (16%) of the increase in total household visits shown in Table 2 (column 1). The latter mostly originates from a greater number of unaccompanied visits (health worker visits without supervisor presence). No household in our data reports having received a visit from the supervisor without the presence of the health worker.

visits while keeping the level of a mediator (here, supervisor's effort) fixed.[30] In Figure A.5, we show that this de-mediated impact of worker incentives increases when supervisor effort increases. This is consistent with higher supervisor effort stimulating higher worker effort and increasing their productivity, and thus consistent with effort complementarity (if efforts were substitutes, higher supervisor effort would have resulted in smaller de-mediated treatment effects).

4.3 Empirical Evidence of Contractual Frictions

Table 6 shows that side payments are limited in our context. 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).[31] Column (1) shows that the share of supervisors who made positive side payments is 1.1% in the control group, 1.6% in T_{worker}, 11.3% in T_{shared}, and 19.4% in T_{supv}. Column (3) shows that the average amount that a supervisor transferred each month to a worker is 702 SLL (resp., 432 SLL) in T_{supv} (resp., T_{shared}), which is less than 3% of the supervisor's incentive payouts. Columns (2) and (4) show that workers also occasionally made side payments to their supervisor in T_{worker}, but the amount of such transfers is negligible (average of 151 SLL, which is less than 0.3% of their incentive payouts). Overall, this evidence shows that side payments are minimal in our context.

In our theoretical framework, side payments can be limited for two reasons. First, the value that the supervisor attaches to a household visit — absent the incentive scheme — may be small compared to that of the health worker (a small or negative value of $b_2 - b_1$). In this case, the worker would exert more effort than what the supervisor would want to induce with a side payment, and so no side payment is necessary. Second, contractual frictions ($z > 1$) can restrict the side payment offered. In this case, the supervisor would want the worker to exert more effort but refrains from using side payments due to their high costs.[32] To empirically distinguish between these two reasons, we propose two tests, both of which

[30]This quantity captures the treatment effect that would be observed if supervisor effort was fixed at an exogenous level, while worker's effort was allowed to change in response to the treatment. For additional information regarding the mediation analysis, please refer to the notes of Figure A.5.

[31]This was asked to health workers rather than supervisors to limit recall bias. Supervisors and workers have no incentive to misreport transfers because these were allowed in our setting. See Section 2.2 for details.

[32]Formally, the side payment s^* offered by the supervisor is given by:

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

This expression shows that s^* decreases both in contractual frictions z and in the term $b_2 - b_1$.

indicate that contractual frictions are a key reason why side payments are limited.

Evidence 1: Side payments are limited even when workers have better outside options than their supervisor ($b_2 > b_1$). As a first test, we examine side payments when the worker has a better outside option than her supervisor, indicated by higher education levels or higher hourly wages from a secondary job. These are cases where $b_2 - b_1$ is likely positive.[33] We expect the worker to exert less effort than the supervisor would find optimal in these cases, and the supervisor to have strong reasons to offer a sizable side payment to the worker in the absence of contractual frictions. Yet, columns 1 and 2 of Table A.8 show that, even in these (relatively rare) cases, side payments are modest across all treatments, suggesting that contractual frictions must play a role in explaining the limited use of side payments. In line with this, we have documented in Section 2.2 that 75% of the supervisors mentioned contractual frictions (arising from social norms against redistribution or the difficulty of making binding commitments) as the main reason for not making side payments.

Evidence 2: Supervisor welfare is higher with shared incentives. As a second test, we estimate whether shared incentives generate higher supervisor welfare than supervisor incentives. In the absence of contractual frictions, the supervisor always prefers to receive a larger share of the incentive. In the presence of contractual frictions, the supervisor may instead find it costly to motivate the worker with side payments and hence can be better off under a scheme that offers a larger share of the incentive to the worker. Thus, finding that supervisor welfare is higher under shared incentives would constitute strong evidence that contractual frictions constrain transfers.

We present two pieces of evidence showing that supervisor welfare is higher under shared incentives than under supervisor incentives. First, we provide direct evidence on the three components of supervisor welfare highlighted in our theoretical framework. These elements are: (i) the net financial gain from the scheme, which is given by incentive payments minus side payments (this is the term: $((1-p)m - sz)y$), (ii) the cost of effort ($c(e_2)$), and (iii) the direct benefit generated by household visits ($b_2 y$). In Table A.8 (column 3), we show that the financial gain from the scheme is quantitatively similar, and statistically indistinguishable in the shared and supervisor incentives treatments. Further, in Table 5, we have shown that supervisor effort is equalized in the two treatments. Finally, we know from Table 2 that visits are significantly higher with shared compared to supervisor incentives, and hence that $b_2 y$ is higher with shared incentives. To sum up, components (i) and (ii) are approximately equal in the two treatments, while component (iii) is significantly higher under shared incentives.

[33] 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.

This suggests that supervisor welfare is higher under shared incentives than under supervisor incentives.

Second, we demonstrate a clear preference among supervisors for the shared incentives treatment over the supervisor incentives treatment. When prompted to indicate their preferred distribution of incentives at endline, 63% of supervisors opted for an equal division of incentives between themselves and the workers (1,000 SLL each). Meanwhile, 23% selected 1,500 SLL for themselves and 500 SLL for the worker, and only one supervisor chose to retain the entire incentive. These findings reinforce the notion that supervisors' welfare is likely higher under the shared incentives than the supervisor incentives treatment.[34]

In summary, both tests offer compelling evidence that supervisors are better off under shared incentives, pointing to the presence of contractual frictions that prevent supervisors from fully capitalizing from the supervisor incentives treatment.

4.4 Alternative Mechanisms

The previous two sub-sections provide empirical support for our theoretical framework, in which two-sided incentives outperform one-sided incentives due to the coexistence of effort complementarities and contractual frictions. We now provide evidence against alternative mechanisms that are not considered in our model but could explain why two-sided incentives outperform one-sided incentives.

Fairness concerns We first consider fairness concerns, such as those based on aversion to pay inequality. These concerns could lower the morale of the unincentivized agents in the one-sided incentives schemes (workers in T_{supv} and supervisors in T_{worker}) and thus reduce their effort. We provide three pieces of evidence against this mechanism. First, we observe that, contrary to the prediction of this mechanism, supervisor's effort is, if anything, higher in T_{worker} relative to the control group (though the two are statistically indistinguishable). This is inconsistent with supervisors being demotivated by the unequal allocation of incentives. Second, we show that health workers are largely unaware of supervisor pay, which means that fairness concerns are also likely to be of limited relevance for workers. Only 15% (resp., 20%) of workers in T_{supv} (resp., T_{shared}) reported knowing that their supervisor received an incentive, and in the vast majority of these cases they underestimated the size of the incentive. Finally, we show that one-sided incentives do not decrease the job satisfaction of unincentivized agents. Table A.9 shows that there is no evidence that unincentivized agents

[34]When choosing between different incentive schemes, supervisors may care about fairness — a point we explore in more depth in the next section. However, when $z < 1$, fairness-minded supervisors would allocate the entire incentive to themselves and then redistribute a fair share of the incentive to their workers. Thus, the fact that supervisors want the organization to implement an equal split of the incentive upfront is evidence of contractual frictions regardless of whether supervisors have fairness concerns or not.

are less satisfied with their payment, the organization, or the job compared to the control group. If fairness concerns were the primary mechanisms driving our results, we would instead expect the unincentivized agents to be less satisfied than their control peers. All in all, these three pieces of evidence imply that it is unlikely that fairness concerns are a first-order driver of our results.

Positive reciprocity We next consider positive reciprocity towards the organization paying the incentive. Under this mechanism, incentivized agents increase effort to reciprocate what they perceive to be a kind act of the organization. However, Table A.9 shows that incentivized agents (workers in T_{worker} and supervisors in T_{supv}) do not report higher satisfaction with the job and the organization. Additionally, workers in all treatments are equally likely to "self-identify" through their job (rather than an ethnic group, a language, a religion) and to find the work environment competitive: columns (7) and (8) in Table A.9. In general, these results are inconsistent with the positive reciprocity mechanism, as it is unlikely that reciprocal agents would increase effort but not report a change in their satisfaction with or perception of the organization.

Sharp non-linearities in the utility or cost function We finally consider non-linearities in the utility or cost function. Two-sided incentives could outperform one-sided incentives if the utility or cost functions were highly non-linear so that the returns to additional incentives above 1,000 SLL fell sharply with the size of the piece rate offered. For example, with a cubic cost function and no effort complementarities, shared incentives ($p = .5$) would maximize output in a model where the supervisor and the worker contemporaneously choose efforts. In this model, marginal costs increase steeply with effort, making it suboptimal for the principal to try to elicit too much effort from a single agent.

We first note that, while these non-linearities could explain why two-sided incentives outperform one-sided incentives, they would not be able to account for several of our additional findings presented above. For example, the results of the mediation analysis and the finding that supervisor welfare is higher under shared incentives than under supervisor incentives.

Next, we provide suggestive empirical evidence against strong non-linearities. Suppose that marginal utility fell steeply in consumption — an hypothesis that is not particularly plausible in our context given the fact that even the wealthiest workers and supervisors in our sample are relatively poor, and that incentive payments are a nontrivial fraction of their earnings.[35] We would then expect incentives to have weaker effects for wealthier workers and supervisors, since these agents would derive less utility from the financial rewards offered by the intervention. Contrary to this hypothesis, in Panel A of Figure A.6 we show

[35] 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).

non-parametrically that the impacts of incentives on household visits are approximately constant over the distribution of worker and supervisor wealth. Further, suppose that marginal costs increased steeply in effort, e.g., because additional units of effort require traveling an ever greater distance — again a hypothesis that does not match our context well since workers can increase output by visiting the same households more often. In this case, we would expect a weaker impact of worker (supervisor) incentives when the worker (supervisor) lives far away from the households she serves (the worker she supervises), since these agents face a higher marginal costs of effort and will thus increase effort by a lower amount when incentivized. Contrary to this hypothesis, in Panel B of Figure A.6, we show fairly homogeneous impacts of the two treatments across the distribution of household-worker distance and worker-supervisor distance. All in all, our results do not appear consistent with strong non-linearities in the utility or cost function.

5 Structural Model

In this section, we use the exogenous variation generated by the interventions to structurally estimate the model presented in Section 4.1, 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.

5.1 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.[36] 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

[36] In our model, there is no individual heterogeneity, so we only rely on empirical moments capturing mean outcomes. For the structural analysis, we measure visits with the "total monthly visits", adjusted by the inverse of the sampling probability, and supervisor effort with the share of accompanied visits. We do not directly observe worker effort (as this is challenging to accurately measure in a survey), and 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.[37]

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.[38]

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:

$$\hat{\boldsymbol{\theta}} = \min_{\boldsymbol{\theta}} \left[\boldsymbol{m}_S(\boldsymbol{\theta}) - \boldsymbol{m}\right]' \cdot J(\boldsymbol{m})^{-1} \cdot \left[\boldsymbol{m}_S(\boldsymbol{\theta}) - \boldsymbol{m}\right]. \tag{5}$$

$J(\boldsymbol{m})$ 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.10 and A.11). 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

[37] For this calibration, we do not rely on 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.

[38] Alternatively, we could assume that supervisors realize that q varies with p, and thus that they form separate expectations about the reporting rate in each treatment. However, under this assumption, side payments would be unrealistically unresponsive to p (a slope coefficient lower than 0.5). Additionally, the fit of this model is worse than that of the model that assumes a common reporting rate.

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.

5.2 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 80% larger when the supervisor exerts the control level of effort, and 114% 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 complementarity between the efforts of the two agents.

The calibrated value of parameter z implies that side payment costs increase more than threefold due to contractual frictions. This constitutes a strong 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 ($b_2 < b_1$).[39] Overall, this indicates that interventions that fail to incentivize the supervisor may be ineffective: the contribution of the supervisor is key to ensuring the worker can be productive, but absent additional incentive, the supervisor will under-provide her key support to the worker.

5.3 Counterfactual Policies

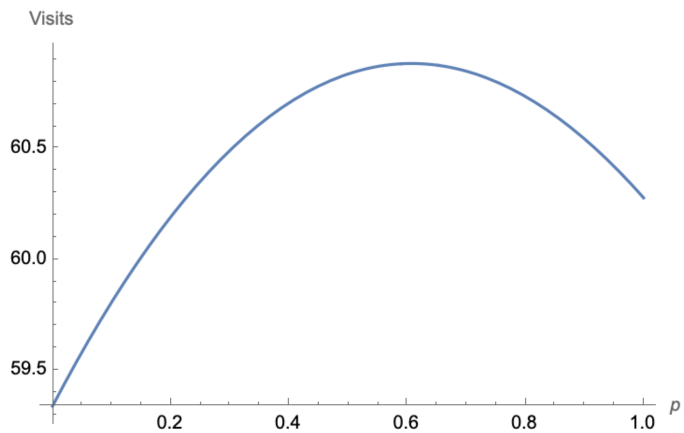
We conduct three counterfactual policy experiments that explore, in turn, how to optimally share the incentive between the two agents, how the optimal incentive changes as key struc-

[39] 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.

tural 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 2, we show that to maximize household visits, the worker should be offered 61% 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.

Figure 2: Optimal Incentive p^*

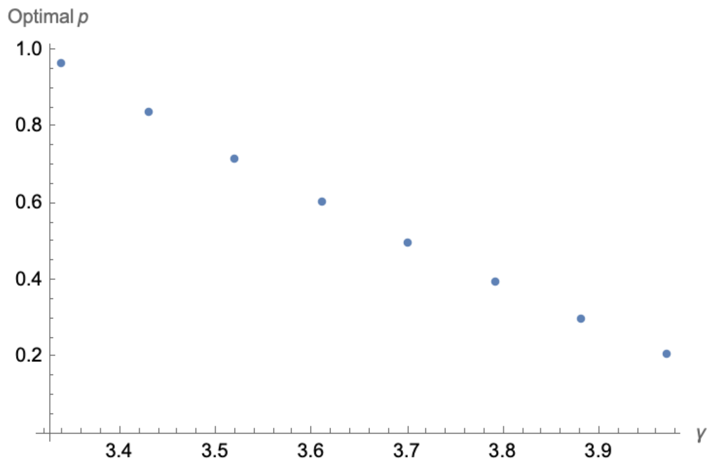


This result, however, depends strongly on the strength of the complementarity between worker and supervisor effort. We illustrate this point with our second counterfactual experiment in Figure 3. Here, we plot the optimal share of the incentive offered to the worker (p^*) for different levels of complementarity. 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% higher than what we estimate, the optimal incentive scheme would give 80% of the piece rate to the supervisor.

Thus, these results suggest that in organizations in which effort complementarity 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

Figure 3: Optimal Incentive p^* by Complementarity γ



a household, regardless of whether this results in a visit or not). 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.[40] What emerges is that, at the current level of complementarity, incentivizing output through an equally-shared piece rate generates 7.2% more visits than optimally incentivizing effort (60.8 visits instead of 56.7 visits) 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.[41]

6 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

[40]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.

[41]For insight into this result, consider a simplified scenario where $\alpha = 1$, $b_1 = b_2 = 0$, $c_1 = c_2 = c$, and there are no side payments. Focusing on the worker's problem, the principal can offer either output incentives worth $p * m_0$ for each unit of output, or effort incentives worth $p * m_e$ for each unit of effort. With output incentives, the worker's objective is $\pi_1 = (e_1 + \gamma e_1 e_2) m_o p - c e_1^2$ and her optimal effort choice is $e_1 = \frac{(1+\gamma e_2) m_o p}{2c}$. With effort incentives, the worker's objective function is $\pi_1 = e_1 m_e p - c e_1^2$ and her optimal effort choice is $e_1 = \frac{m_e p}{2c}$. Evidently, output incentives drive the worker to select effort that considers the strategic complementarity with the supervisor, while effort incentives prompt the worker to opt for effort levels independent of the strategic complementarity.

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, which postulates that any incentive allocation should result in the same output level (Coase 1937, 1960). They also contradict the priors of most experts who forecasted our results on the Social Science Prediction Platform.[42] 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 (Perry 2020).

The view of organizations that emerges from our results sees the coexistence of effort complementarities and contractual frictions as a central determinant of organizational performance. This has a number of important policy implications. First, the allocation of incentives across the hierarchy is a key lever to boost performance in organizations where vertical side transfers are limited. Second, it is optimal to ensure both layers of the organization are properly incentivized when worker-supervisor effort complementarities are strong. Third, policies that try to improve Coasian bargaining within the firm may backfire if supervisors are poorly motivated (as in our context) and may use side transfers to pursue objectives that are inconsistent with those of the organization.

The estimates should not be extrapolated to other organizations without an understanding of their context. In the organization we study, the complementarity between worker and supervisor effort is strong, as we expect to be the case in several other organizations. However, we would expect weaker complementarities, and hence a less pronounced impact of shared incentives, in contexts in which the role of the supervisor is limited to monitoring. Our structural results emphasize the importance of calibrating incentives in each organization based on the specific level of effort complementarities that prevails in the hierarchy.

It would also be important to ascertain whether our results persist in the longer run. Our evidence suggests that treatment effects are fairly stable for a period of one year between baseline and endline.[43] However, our findings also reveal that shared incentives are less potent for workers with over four years of tenure, which suggests that, depending on the level of turnover of the organization, the treatment effects of shared incentives may eventually decline. Further, supervisor's knowledge may gradually get depleted (Garicano and Rayo 2016), potentially making worker incentives more attractive over a longer time horizon than what we measured in our experiment. Exploring the dynamic impacts of incentive allocation

[42] 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.

[43] This conclusion is drawn from Figure A.1, which charts the trajectory of reported visits. Assuming that the reporting rate neither escalates nor varies considerably across treatments, the figure indicates that the impact of the shared incentives on output remains constant over time.

across organizational layers offers an exciting opportunity for future research.

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. 2022; 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." Working Paper.

Adhvaryu, Achyuta, Emir Murathanoglu, and Anant Nyshadham. 2023. "On the Allocation and Impacts of Managerial Training." Working Paper.

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.

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.

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, Robin Burgess, Erika Deserranno, Ricardo Morel, Munshi Sulaiman, and Imran Rasul. 2023. "Social Incentives, Delivery Agents, and the Effectiveness of Development Interventions." *Journal of Political Economy Microeconomics* 1 (1):162–224.

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

Bandiera, Oriana, Ahmed Elsayed, Anton Heil, and Andrea Smurra. 2022. "Economic Development and the Organisation of Labour: Evidence from the Jobs of the World Project." *Journal of the European Economic Association* 20 (6):2226–2270.

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.

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

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.

Caliendo, Lorenzo, Ferdinando Monte, and Esteban Rossi-Hansberg. 2015. "The Anatomy of French Production Hierarchies." *Journal of Political Economy* 123 (4):809–852.

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.

———. 2023. "The Political Economy of Public Sector Absence." *Journal of Public Economics* 218:104787.

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.

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 and Ricardo Perez-Truglia. 2023. "The Old Boys' Club: Schmoozing and the Gender Gap." *American Economic Review* 113 (7):1703–1740.

Cullen, Zoë B. and Ricardo Perez-Truglia. 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.

Dodge, Eric, Yusuf Neggers, Rohini Pande, and Charity Troyer Moore. 2022. "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." 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. 2005. "Four Formal(izable) Theories of the Firm?" *Journal of Economic Behavior & Organization* 58 (2):200–245.

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

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

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

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.

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 Mass Sterilizations Campaign." 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 Manage in the Bangladeshi Garment Sector." Working Paper.

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.

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.

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.

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." Tech. rep., Ourworldindata.

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." Working Paper.

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.

WHO. 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.

Wilson, James Q. 2019. *Bureaucracy: What Government Agencies Do and Why They Do It*. Hachette UK.

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

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	0.534
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, aggregate at 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, 1}	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 derived from a regression of each variable on the three 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 (aggregate at the health worker/village level) in Panel C, the health worker's leaflet in Panel D, the facility admin data in Panel E. The wealth score counts the number of items owned on a list of 8 household items (e.g., clothes, pair of shoes, cooking pots).

Table 2: Household Visits

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Household visits provided by the health worker in the past 6 months							**Trust**
Dep. Var.	Number of visits	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 visited	Number of visit types	% households who trust the health worker as a health provider
Worker incentives	2.090***	2.024**	0.603	0.164	0.001	0.072***	0.250***	0.037
	(0.558)	(0.941)	(1.423)	(0.127)	(0.174)	(0.025)	(0.094)	(0.023)
Supervisor incentives	2.145***	1.933**	-0.070	0.173	0.003	0.082***	0.325***	0.031
	(0.501)	(0.926)	(1.409)	(0.130)	(0.181)	(0.025)	(0.100)	(0.024)
Shared incentives	3.356***	4.134***	1.590	0.528***	0.425**	0.127***	0.565***	0.071***
	(0.490)	(0.927)	(1.290)	(0.134)	(0.185)	(0.023)	(0.092)	(0.024)
Unit	Worker	Worker	Worker	Worker	Worker	Worker	Worker	Worker
Observations	2,926	2,926	1,803	2,926	1,803	2,926	2,926	2,926
Mean dep. var.	7.296	14.388	23.404	2.248	2.922	0.709	1.745	0.745
Mean dep. var. in Control	5.334	12.324	22.736	2.015	2.782	0.637	1.448	0.707
p-value Worker = Supervisor	0.932	0.927	0.635	0.946	0.990	0.710	0.492	0.808
p-value Supervisor = Shared	0.038	0.024	0.196	0.017	0.033	0.060	0.026	0.102
p-value Worker = Shared	0.046	0.033	0.451	0.013	0.026	0.026	0.002	0.147

Notes: Data source is the household survey, aggregate at the health worker level. Cols. (3) and (5) limit the sample to households that have received 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, aggregate at the 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 Payments

	(1)	(2)	(3)	(4)	(5)
	Reporting		Incentives payments per month (in 1,000 SLL)		
Dep. Var.	Number of reports per month	Reporting rate = number of reports/ number of visits	Incentives payments to health worker	Incentives payments to supervisor (per health worker)	Total incentives payments per health worker (col. 3 +4)
Worker incentives	21.547***	0.666***	52.506***	-0.239	52.267***
	(2.278)	(0.128)	(4.385)	(0.475)	(4.397)
Supervisor incentives	6.543***	0.175*	-0.424	22.685***	22.261***
	(1.757)	(0.103)	(0.674)	(3.265)	(3.302)
Shared incentives	14.272***	0.273**	18.816***	18.879***	37.695***
	(2.487)	(0.108)	(2.447)	(2.405)	(4.840)
Unit	Worker	Worker	Worker	Worker	Worker
Observations	2,970	2,624	2,970	2,970	2,970
Mean dep. var.	15.639	0.530	18.575	10.538	29.113
Mean dep. var. in Control	4.574	0.234	0.000	0.000	0.000
p-value Worker = Supervisor	<0.001	<0.001	<0.001	<0.001	<0.001
p-value Supervisor = Shared	0.008	0.398	<0.001	0.348	0.008
p-value Worker = Shared	0.026	0.005	<0.001	<0.001	0.025

Notes: The reporting rate in col. (2) is calculated by dividing the number of SMS reports (from the SMS admin data; col. 1) by the total monthly visit count, adjusted by the inverse of the sampling probability (obtained from the household survey) 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)
Dep. Var.	% accompanied household visits	% health workers visited in the past 6 months	Supervisor organized general training in the past month = {0, 1}	Difference in health workers' knowledge between baseline and endline
Worker incentives	0.030	-0.050	0.004	0.158
	(0.022)	(0.137)	(0.005)	(0.116)
Supervisor incentives	0.057**	-0.041	0.006	0.063
	(0.023)	(0.137)	(0.005)	(0.113)
Shared incentives	0.062***	-0.043	0.003	0.266**
	(0.021)	(0.139)	(0.005)	(0.121)
Unit	Worker	Worker	Worker	Worker
Observations	2,919	2,833	2,864	2,927
Mean dep. var.	0.204	1.375	0.994	0.313
Mean dep. var. in Control	0.164	1.417	0.991	0.433
p-value Worker = Supervisor	0.293	0.950	0.443	0.372
p-value Supervisor = Shared	0.846	0.987	0.463	0.074
p-value Worker = Shared	0.181	0.963	0.916	0.350

Notes: Data source is the household survey in col. (1), aggregate at the health worker level. Data source is the health worker survey in cols. (2)-(5). The dependent variable in col. (1) is equal to the share of household visits accompanied by the supervisor. The dependent variable in col. (4) calculates the difference in health knowledge of the health worker between baseline and endline. Health knowledge is gauged by a test that counts the number of accurate or nearly accurate responses out of 5 questions asked to the health worker concerning the appropriate steps to refer a child under 5 to a 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	0.048	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	8.0
	(2.4)
Worker baseline incentive b_1	70.5
	(27.0)
Supervisor baseline incentive b_2	49.3
	(192.3)
α	1.5
	(5.5)
Calibrated friction z	3.61
Δ in marginal product of worker effort (shared incentives)	118%
Δ in marginal product of worker effort (control)	83%
Total worker cost of effort (control)	1461.7
Total supervisor cost of effort (control)	918.2

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. Costs are expressed in thousand SLL. Boostrapped 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 $\gamma = 8.0$ 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.

Table 8: Moment Fit

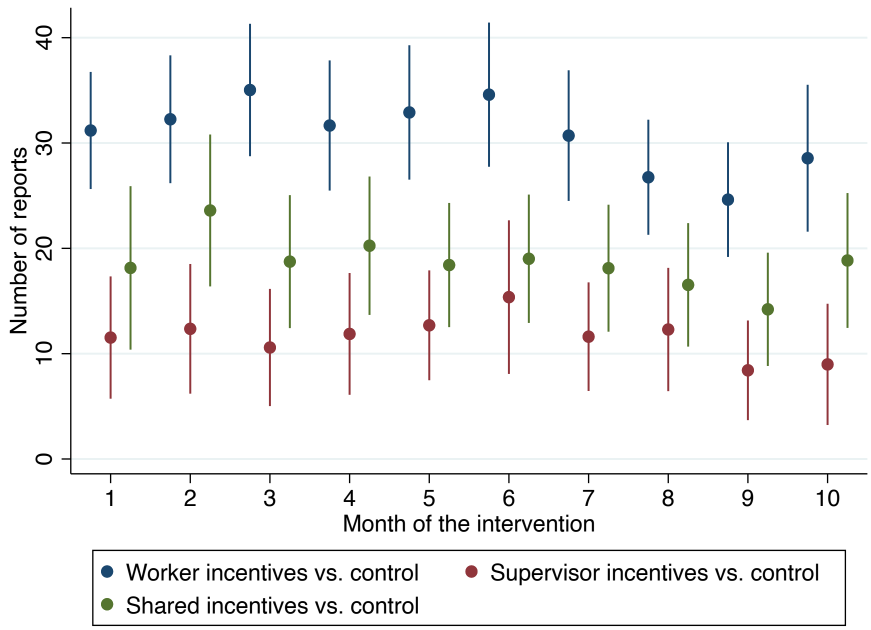
Moments	Targeted Real	Simulated
Supervisor effort in worker incentives treatment	0.198	0.205
Supervisor effort in supervisor incentives treatment	0.225	0.231
Supervisor effort in shared incentives treatment	0.228	0.221
Supervisor effort in control group	0.164	0.156
Output in worker incentives treatment	59.679	61.678
Output in supervisor incentives treatment	58.896	60.774
Output in shared incentives treatment	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: SMS Reporting over Time



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

Figure A.2: Model Predictions With Weak Contractual Frictions ($z < 1$)

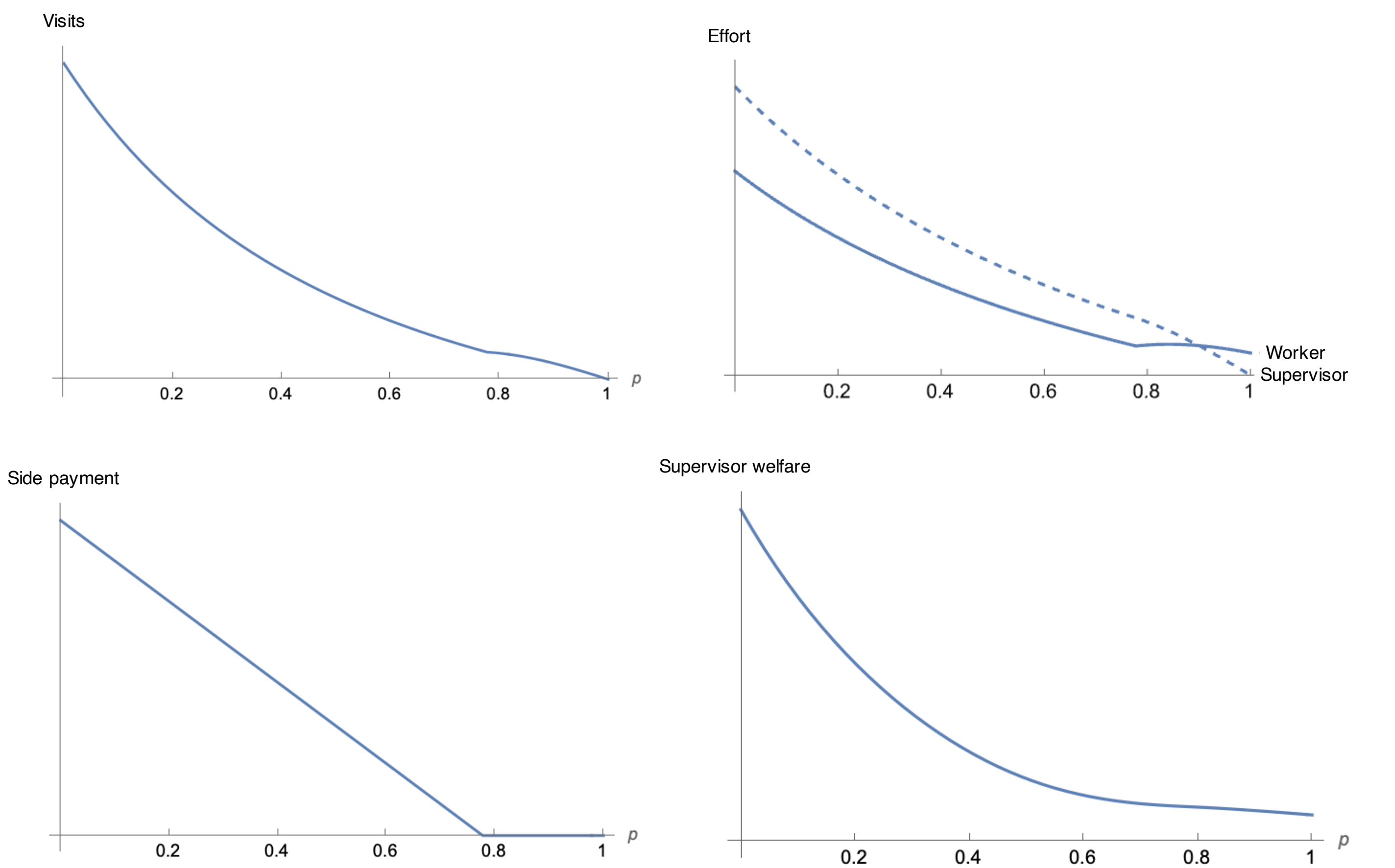


Figure A.3: Model Predictions With Strong Contractual Frictions and Weak Effort Complementarities ($z > 1; \gamma \leq t$)

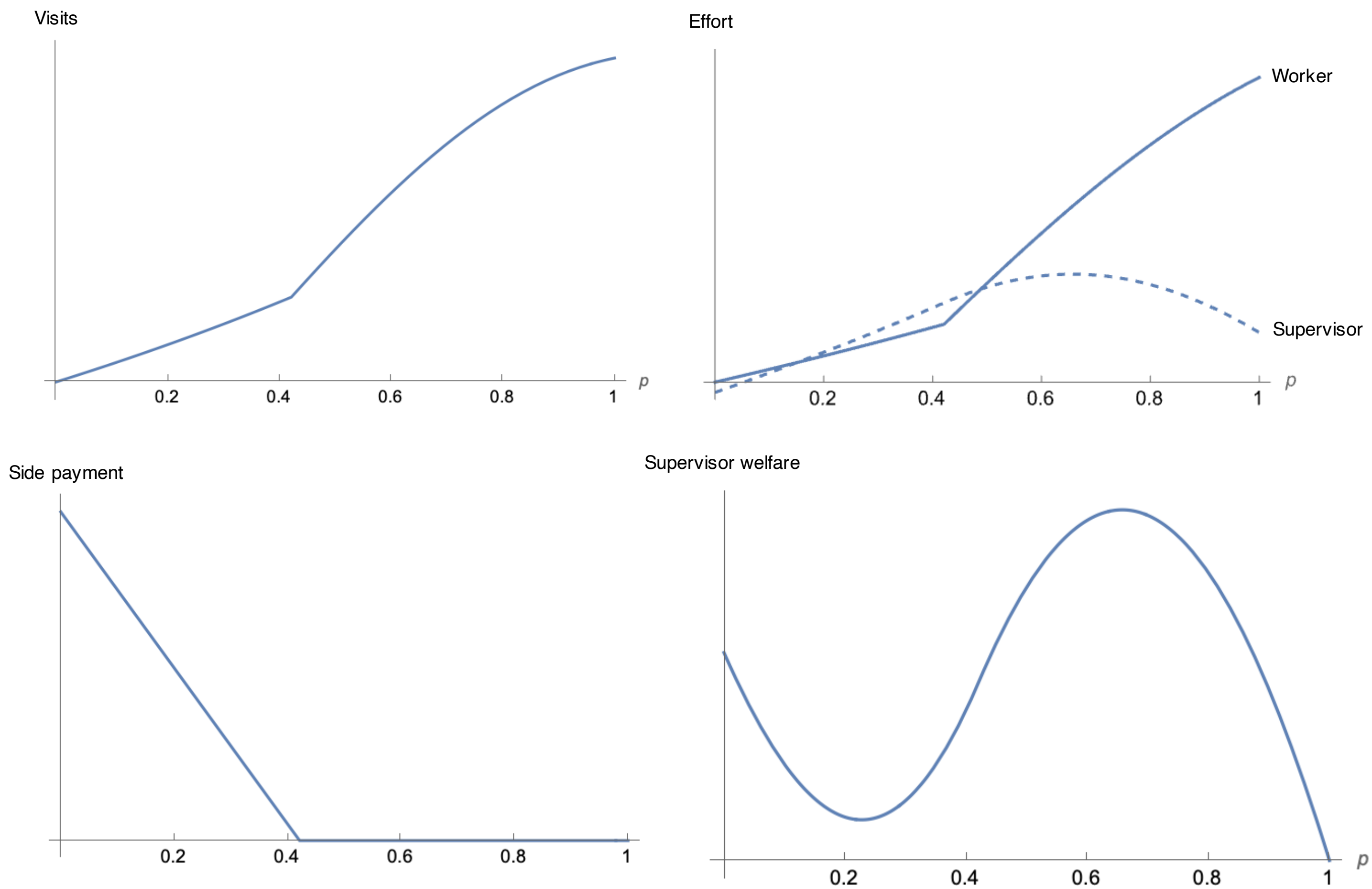


Figure A.4: Model Predictions With Strong Contractual Frictions and Strong Effort Complementarities ($z > 1; \gamma > t$)

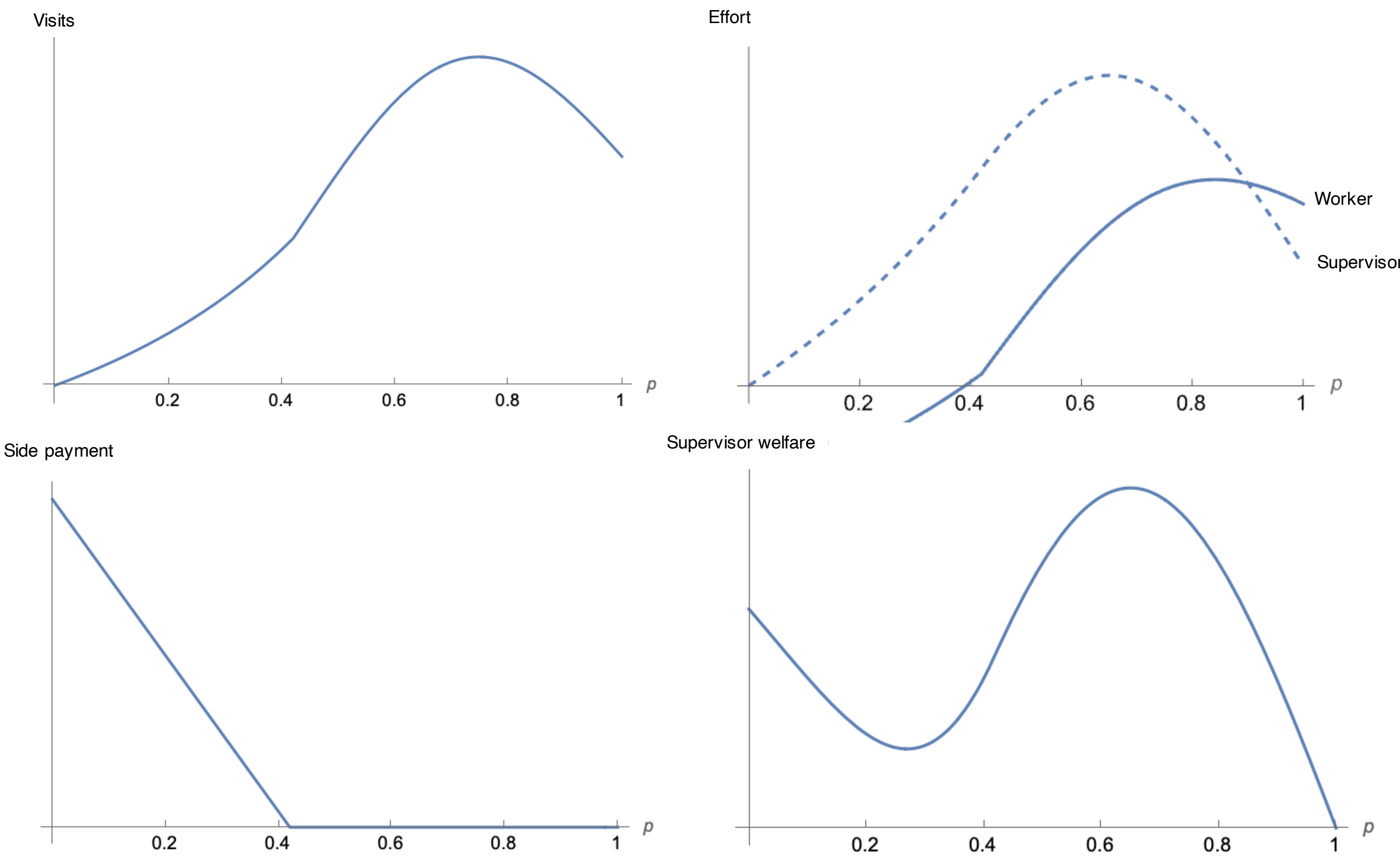
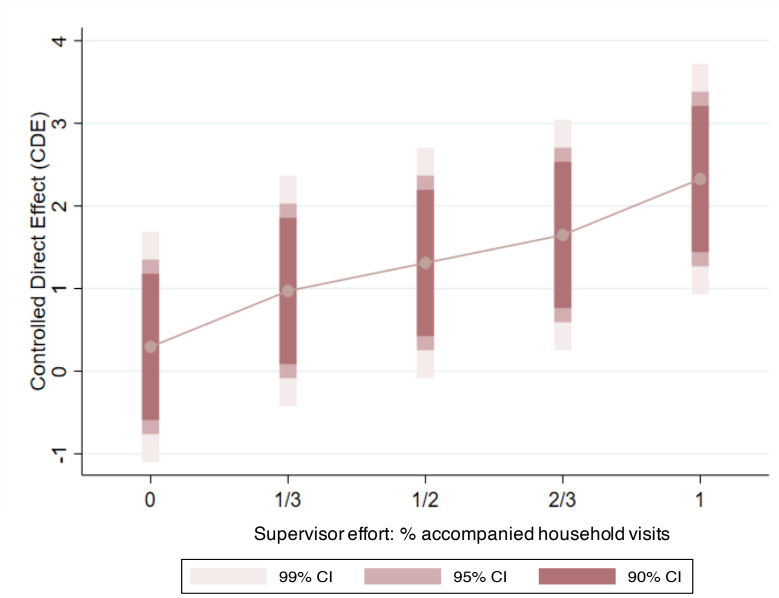


Figure A.5: Mediation Analysis

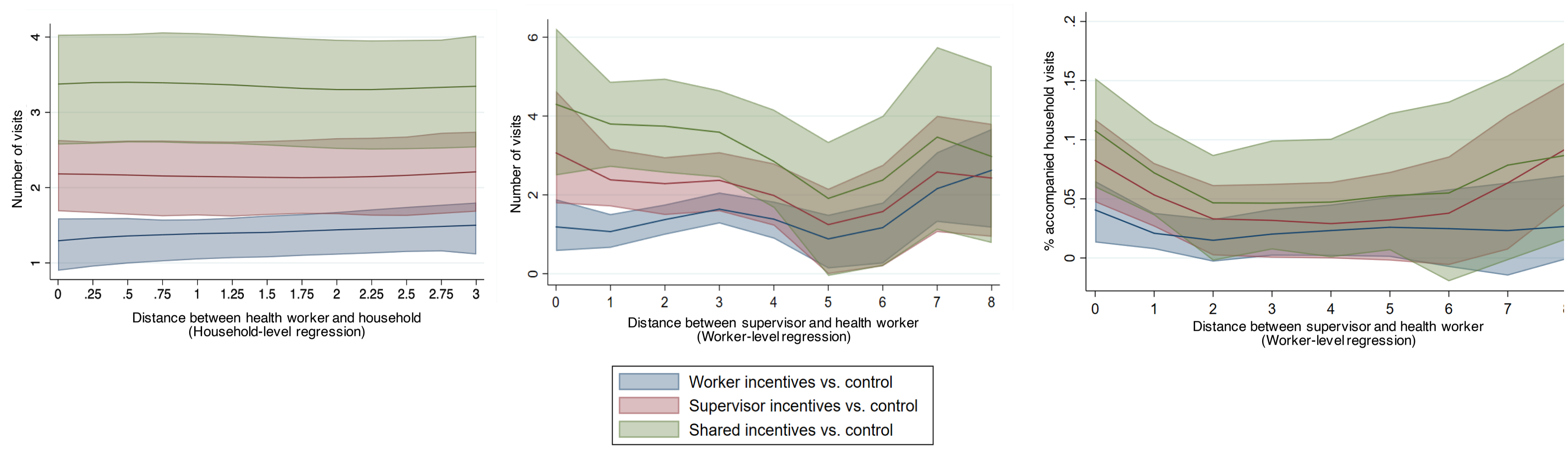


Notes: This figure plots the controlled direct effect (CDE) of the worker incentives treatment on the number of visits provided by a health worker for different values of supervisor's effort, as measured with the share of visits accompanied by the supervisor. We focus on the comparison between the worker 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 et al. (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.

Figure A.6: 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, with 95% confidence intervals. 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. 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.

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, aggregate at 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, 1}	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. The wealth score counts the number of items owned on a list of 8 household items (e.g., clothes, pair of shoes, cooking pots).

Table A.2: Household Visits by Type

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	% households who received [visit type] from the health worker in the past 6 months							
	Natal-related visits				**Disease-related visits**			
Dep. Var.	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)	(0.021)	(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, aggregate at the 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.3: 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 & Yekutieli (2006)		
	Worker incentives	Supervisor incentives	Shared incentives	Worker incentives	Supervisor incentives	Shared incentives	Worker incentives	Supervisor incentives	Shared incentives
Quantity of visits									
Index on quantity of visits (for next 2 variables)	0.004	0.004	0.004	0.002	0.000	0.000	-	-	-
Number of visits	0.004	0.004	0.004	0.002	0.000	0.000	0.001	0.001	0.001
% households visited	0.004	0.004	0.004	0.040	0.009	0.000	0.004	0.002	0.001
Quantity of visits									
Index on quality of visits (for next 4 variables)	0.004	0.004	0.004	0.239	0.161	0.000	-	-	-
Number of visit types	0.004	0.004	0.004	0.124	0.019	0.000	0.006	0.002	0.001
Average visit length	0.008	0.008	0.004	0.483	0.563	0.000	0.016	0.017	0.001
Number of health topics discussed per visit	0.032	0.032	0.004	1.000	1.000	0.001	0.051	0.051	0.001
% households who trust the health worker as a health provider	0.008	0.032	0.004	1.000	1.000	0.040	0.045	0.051	0.003
Health outcomes									
Index on pre- and post-natal care in past 2 years (for next 5 variables)	0.801	0.422	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	0.944	0.606	0.032	1.000	1.000	0.553	1.000	0.763	0.208
% women with institutional birth	0.817	0.171	0.171	1.000	1.000	1.000	0.960	0.489	0.489
% women who received post-natal visit within 2 days of birth	0.980	0.829	0.964	1.000	1.000	1.000	1.000	1.000	1000
% women with at least 6 months of breastfeeding	0.980	0.956	0.295	1.000	1.000	1.000	1.000	1.000	0.629
% households with up-to-date infant vaccination	0.980	0.880	0.586	1.000	1.000	1.000	1.000	1.000	0.763
Index on disease incidence (for next 3 variables)	0.980	0.857	0.092	1.000	1.000	1.000	1.000	1.000	0.420
% children under-5 who had fever	0.598	0.976	0.008	1.000	1.000	0.238	0.763	1.000	0.131
% children under-5 who had cough	0.470	0.980	0.968	1.000	1.000	1.000	0.763	1.000	1.000
% children under-5 who had diarrhea	0.944	0.817	0.980	1.000	1.000	1.000	1.000	0.960	1.000

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.4: 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. The wealth score counts the number of items owned on a list of 8 household items (e.g., clothes, pair of shoes, cooking pots). All regressions include stratification variables. Standard errors clustered at the PHU level. *** p<0.01, ** p<0.05, * p<0.1

Table A.5: 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, aggregate at the 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 aggregate at 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.6: 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 observation. 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.7: 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 dependent variable is the household survey, aggregate at the 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.8: Evidence of Contractual Frictions

	(1)	(2)	(3)
Dep. Var.	Side-payment: Supervisor shared incentive with health worker = {0, 1}		Net financial gain for the supervisor per health worker (incentives payments - net transfer to the worker*z)
Sample:	Workers with higher education than their supervisor	Workers with higher earnings from outside job than their supervisor	All workers
Worker incentives	0.032	-0.033	-0.092
	(0.028)	(0.065)	(0.327)
Supervisor incentives	0.216***	0.248**	20.699***
	(0.066)	(0.103)	(1.800)
Shared incentives	0.161***	0.029	17.579***
	(0.062)	(0.080)	(1.279)
Unit	Worker	Worker	Worker
Observations	512	293	2,488
Mean dep. var.	0.096	0.116	9.732
Mean dep. var. in Control	0.000	0.044	0.057
p-value Worker = Supervisor	0.004	0.020	0.000
p-value Supervisor = Shared	0.532	0.088	0.158
p-value Worker = Shared	0.036	0.323	0.000

Notes: The sample in cols. (1) to (4) is restricted to workers with higher outside option than supervisor, as proxied with whether the worker has a strictly higher education level than the supervisor or an average hourly earnings from any 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). The dependent variable in col. (5) is equal to (the average number of visits reported by the health worker per month) * (the incentive allocated to the supervisor - net transfer from the supervsior to the worker*z), where z = 3.61 (as estimated in the structural model). The variable is expressed in 1,000 SLL. 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: Job Satisfaction

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Health worker is satisfied with [...] = {0, 1}			Supervisor is satisfied with [...] = {0, 1}				
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}
Worker incentives	0.344***	0.074***	-0.039	0.086	0.012	-0.067	-0.046	0.006
	(0.032)	(0.024)	(0.036)	(0.056)	(0.049)	(0.061)	(0.044)	(0.026)
Supervisor incentives	-0.003	0.014	-0.040	0.305***	0.082*	-0.073	-0.023	-0.019
	(0.030)	(0.028)	(0.038)	(0.062)	(0.042)	(0.062)	(0.043)	(0.030)
Shared incentives	0.198***	0.048**	-0.054	0.329***	0.092**	-0.040	-0.024	-0.018
	(0.036)	(0.024)	(0.036)	(0.063)	(0.041)	(0.059)	(0.043)	(0.029)
Unit	Worker	Worker	Worker	Supervisor	Supervisor	Supervisor	Worker	Worker
Observations	2,709	2,825	2,876	360	364	359	2,923	2,923
Mean dep. var.	0.357	0.870	0.793	0.311	0.909	0.738	0.727	0.829
Mean dep. var. in Control	0.219	0.837	0.828	0.132	0.860	0.787	0.746	0.838
p-value Worker = Supervisor	<0.001	0.023	0.986	0.001	0.098	0.929	0.598	0.379
p-value Supervisor = Shared	<0.001	0.186	0.709	0.738	0.757	0.605	0.987	0.963
p-value Worker = Shared	<0.001	0.254	0.679	<0.001	0.048	0.664	0.609	0.394

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

Table A.10: Parameter Estimates (No Effort Complementarities)

	(1)
Marginal product of worker effort α	0.14
Marginal product of supervisor effort β	0.03
Worker cost of effort parameter c_1	0.00
Supervisor cost of effort parameter c_2	2.96
Worker baseline incentive b_1	20.47
Supervisor baseline incentive b_2	41.18
Calibrated friction z	3.61

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.11: Moment Fit (No Effort Complementarities)

Moments	Targeted Real	Simulated
Supervisor effort in worker incentives treatment	0.198	0.206
Supervisor effort in supervisor incentives treatment	0.225	0.204
Supervisor effort in shared incentives treatment	0.228	0.191
Supervisor effort in control group	0.164	0.206
Output in worker incentivs treatment	59.679	64.480
Output in supervisor incentives treatment	58.896	58.053
Output in shared incentives treatment	66.895	59.784
Output in control group	41.040	42.499
Value loss function	43.0	

Notes: This 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.12: Heterogeneity by Promotion Incentives

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

Notes: Data source is the household survey, aggregate at the 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.13: Sample of Health Workers without Promotion Incentives

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

Notes: Data source is the household survey, aggregate at the 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 and the Incentive Amount

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 5.

An alternative design than ours 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.

The Ministry of Health and Sanitation selected an incentive amount of 2,000 SLL. This amount was deemed convenient due to its easy divisibility by 2 in shared incentive scenarios, all the while maintaining its status as a round number. The Ministry regarded any amount above 2,000 as excessive, given that the variable component of a health worker's pay was projected to exceed one fourth of their fixed pay.

B.4 Intervention: Location of the Experiment and Randomization

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.

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.

B.5 Intervention: 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.12 shows that our main treatment effects on visits are orthogonal to the random variation in the promotion system and orthogonal to providing information about the supervisor's fixed wage. This is not surprising as the 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.13 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.6 Intervention: 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.[44] 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.

[44]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.

B.7 Intervention: Training

Before initiating the intervention, workers and supervisors from each PHU were invited to attend a 2-days training session together. These sessions usually occurred within or near the main health facility. Across all 372 PHUs, regardless of treatment, the training first provided instructions on the reporting process (i.e., how to send the SMS, at which cost, what to include in the SMS for the report to be "valid"). It further detailed the back-checks and potential consequences of dishonest reporting. For the worker incentives and shared incentives treatments, all attendees, both workers and supervisors, were informed about the reporting incentives to be given to the workers. In the supervisor incentives treatment and control group, attendees were informed that no incentive would be given to the worker. This approach makes supervisors' aware of the workers' incentives.

At the conclusion of the training, supervisors were isolated into a separate room (without the workers), where they were informed about their incentives per report (if any). This information was not provided to the workers. Supervisors were then explicitly made aware that they held the discretion to share these incentives, either wholly or partially, with their workers. While we did not suggest a specific amount to be transferred, we told the supervisors that these transfers could be viewed as potential incentives to stimulate worker effort.

B.8 Data: Variables Definitions

Accompanied visits: Each household was asked whether they received a visit from the health worker (on her own), the supervisor (on her own), the health worker and the supervisor jointly. 71% of household report having received a visit from the health worker, of which 16% are accompanied by the supervisor. No household reports having received a visit from the supervisor without the presence of the health worker. Among households who received at least one accompanied visit, 97% received one accompanied visit only, and 3% received two accompanied visits. 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.

Health knowledge of health workers: we administered a quiz on health knowledge to all health workers at baseline and endline. 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?

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).

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 experimental design and main outcomes of this study were pre-registered in the AEA RCT Registry with the number AEARCTR-0003345. This paper follows the pre-registration closely. The outcomes variables we use in the paper were all pre-registered.

The only deviation 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 $\times$ 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

In this Appendix, we will first analyze a simplified version of the model that does not have any heterogeneity. This will enable us to establish some key analytical results. We will then present formulas for the headline version of the model, which we use for the structural calibration. Throughout this Appendix, we assume that $z \geq 1$.

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:

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

<u>What level of p maximises output?</u>

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.

<u>$\gamma > 0,\ z = 1$:</u>

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:

<u>Period 2:</u>

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 $\eta = b_1z + b_2 + m(1 + p(z-1))$.

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

In collaboration with the Social Science Prediction Platform,[45] 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 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.

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