Barcelona School of Economics

WORKING PAPERS

The Allocation of Incentives in Multi-Layered Organizations

BSE Working Paper 1346

April 2022 (Revised: September 2024)

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

bse.eu/research

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†

September 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, Paulina Koenig, Margaux Jutant, Mustapha Adedolapo Kokumo, Virginia Manzo, Francesco Marchese, 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. This project was funded with UK International Development from the UK government, awarded through the J-PAL Governance Initiative and the Economic Development and Institutions initiative, and the Rockefeller Foundation (through the IGC). 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) and grants MCIN/AEI/10.13039/501100011033, 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 CIREP: 2024-352) 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@unibocconi.it. 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 is a first-order question in economics since the emergence of complex, hierarchical organizations play 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 — strengthens 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]

We designed a field experiment that creates random variation in the recipient of a new incentive scheme across 372 health units in Sierra Leone. This incentive pays 2,000 Sierra Leone Leones (SLL; approximately $0.25) for each health visit that is completed and reported by SMS by a health worker, in addition to a pre-existing fixed salary. The allocation of this incentive occurs in one of three ways: (i) exclusively to the health worker who performed the visit, (ii) solely to the worker's supervisor, or (iii) equally divided between the worker and their supervisor. To assess whether the incentive allocation impacts output, we collect measures of the quantity and quality of the completed visits, and health outcomes by interviewing a random sample of households in each village 18 months after the start of the intervention. We then assess how these outcomes vary across treatments. We do not rely on the number of visits reported by workers, as these figures may deviate from the actual visits conducted due to reporting costs, which we discuss in detail below.

In the first part of the paper, we present our main results. 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. 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; León-Ciliotta, Zejcirovic, and Fernandez, 2024).

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

[1] Such potent complementarities would not exist if the supervisor's role was merely confined to monitoring. Therefore, our study diverges from recent literature that focuses on the surveillance role of middle managers (Callen et al. 2020; Muralidharan et al. 2021; Bandiera et al. 2021; Dal Bó et al. 2021; Dodge et al. 2022).

facility (as opposed to at home) in the shared incentives treatment than in the one-sided 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.

Finally, we investigate workers' reporting behavior, which is key because our incentive scheme rewards both workers and supervisors based on the number of visits *reported* by the health worker via SMS. Due to a system of extensive back-checks, we find that over-reporting (reporting a visit that did not occur) is minimal. Instead, under-reporting (not reporting a visit that actually occurred) is more prevalent, likely due to high costs of reporting. As expected, under-reporting diminishes with the increase in the share of the incentive offered to the worker: health workers in the worker incentives treatment exhibit the highest reporting rate (90%), followed by those in the shared incentives treatment (51%), the supervisor incentives treatment (41%) and finally the control group (23%). This suggests that workers are willing to bear the cost and hassle of reporting only if the incentive is significant enough to warrant the effort. This is an important result as many of the incentive schemes worldwide reward workers based on reported rather than actual output, as detailed in Section 2.2.

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. The model has three time periods. First, the worker chooses a reporting rule, determining the threshold of reporting cost above which she will report a completed visit. Second, the supervisor chooses how much effort to invest in training and advising the worker, and whether to offer the worker a side payment conditional on the number of reported visits. Third, the worker chooses how much effort to exert to provide visits. The key intuition of the model is that two-sided incentive schemes, such as our shared incentives treatment, are optimal when: (i) the efforts of workers and supervisors are strategic complements, and (ii) the supervisor's ability to motivate the worker through side payments is limited due to contractual frictions.

We assess the relevance of effort complementarities and contractual frictions in explaining our headline results through a series of empirical tests derived from our model. To support the presence of effort complementarity, 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. Additionally, 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. Lastly, 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. Lastly, we provide evidence that the limited use of side transfers is due to contractual frictions. We show indeed that supervisor welfare is higher under shared than supervisor incentives, a result that would be difficult to reconcile with the absence of contractual frictions, as without these frictions, supervisors should be able to replicate any incentive allocation through side transfers if such an allocation maximized their welfare.

Overall, our findings suggest that shared incentives outperform one-sided incentives due to the presence of effort complementarities and contractual frictions. We rule out three alternative explanations: 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 146% higher due to the complementarity with supervisor effort. Furthermore, our calibrated contractual friction parameter implies that difficulties in contracting increase the cost of side transfers by more than threefold. From our counterfactual policy analysis, we derive several key lessons for optimal policy. First, we investigate the optimal allocation of the incentive to maximize household visits and show that the worker should be offered 56% of the overall incentive, which is very close to the equal share we offered in the shared incentives treatment. Second, we estimate the decrease in complementarity as workers gain more experience and show that shared incentives continue to be uniquely effective at boosting output, even at levels of experience greater than those observed in our sample. However, we also observe that the optimal allocation of incentives shifts in response to large changes in the complementarity parameter. This emphasizes the importance of recalibrating the policy in new contexts. Finally, we show that shared incentives remain the most effective policy for maximizing visits, even when under-reporting is eliminated. Thus, we conclude that the fundamental insight of this paper — that shared incentives surpass one-sided incentives — remains robust, irrespective of whether the organization raises reporting or not.

This paper contributes to three 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. 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. 2023b; 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;

[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.

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; Sen 2024).

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 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.[3] Contrary 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.

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. 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 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 and 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.

[3] See the Conclusion and Appendix E for more details on the results and the platform.

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.[4] 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

[4] 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.

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.[5] In the supervisor incentives treatment (T_{supv}), the 2,000 SLL 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), there is no monetary incentive tied to health visits. 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.[6]

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 (2024). We describe the change in the promotion system and show that our results are orthogonal to this variation in Appendix B.5.[7]

[5] 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.

[6] 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.

[7] 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.

Structure of the incentives The incentive scheme has two important features. First, the incentives were disbursed by a reputable external organization independent from the government. Subjects were paid monthly through mobile money and without any delay. This enabled us to establish the credibility of the new incentive scheme in the eyes of all experimental participants.[8]

Second, the payment of incentives was based on *self-reports* by the workers. To report a visit, a worker must send an SMS from their main phone number to a toll-free number. The SMS must include the service date and the patient's contact number and be sent from the worker's registered phone number. This requirement ensures that neither supervisors nor households can report services on behalf of the workers. All health workers participating in our study, including those in the control group, were asked to report their visits.

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 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. Our design, however, does not prevent under-reporting. Even though the SMS reporting tool is free, reporting inherently involves various costs. Firstly, it takes time and necessitates the gathering of information, such as the patient's name and a contact phone number. This process can be further complicated if the patient lacks a personal phone, requiring the worker to source a number from a neighbor or family member. Secondly, mobile phone coverage in rural areas of Sierra Leone, similar to many other low-income countries, is erratic and unreliable. Additionally, common issues such as low phone battery life — often exacerbated by limited access to electricity — can further hinder reporting efforts. These challenges frequently lead to multiple attempts to send a report or, in some cases, prevent reporting altogether.

When considering the policy relevance of our paper, it is important to note that incentive schemes based on self-reported outputs are widely used. These schemes are especially common in settings where direct monitoring of actual outputs is impractical. For example, in various decentralized contexts such as rural healthcare, community health workers and nurses are often compensated based on self-reports, as observed in Bangladesh, India, Rwanda, and other parts of Africa (Perry 2020). Similarly, incentive schemes based on self-reported outputs are prevalent in bureaucracies, where compensation often depends on achievements that are difficult to verify (Besley et al. 2021; Bandiera et al. 2019). Despite the wide use of incen-

[8] 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.

tive scheme based on self-reports, getting accurate self-reported performance data has proven difficult in many settings (Callen et al. 2023a; Bossuroy, Delavallade, and Pons 2024). For example, Karing (2021) found that health facilities in Sierra Leone frequently under-report vaccination records, even when incentivized financially. More broadly, issues with phone connectivity and functionality, and access to electricity, are prevalent in many low-income countries, and often lead to under-reporting.

Information provided to workers and supervisors All workers and their respective supervisors underwent a two-day training conducted at the PHU that provided instructions on the reporting process (i.e., how to send the SMS, what to include in the SMS for the report to be "valid"). It further detailed the back-checks and potential consequences of dishonest reporting. To maintain consistent focus on the training material across treatments, workers in T_{worker} and T_{shared} were informed of the incentives 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.

At the end of the training, supervisors were taken to a separate room (away from the workers) where they were informed about their incentives (if any) and the linkage to reporting. Supervisors' incentives were not disclosed to the workers. In this manner, our context mirrors 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). This setup also 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).

Throughout the duration of the experiment, supervisors were not informed about the number of SMS messages sent by individual workers or details regarding workers' incentive payments. This decision was made because, as we will show in Section 3.2, the reporting behavior of workers varies considerably across treatments. 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.[9] The fact that supervisors are unaware of workers' earnings also further minimizes the possibility that the supervisor and the worker would collude to report visits that have not actually been carried out.

[9]Supervisors in T_{supv} and T_{shared} received monthly information on the *aggregate* number of reported visits in the PHU through their monthly paycheck, but were never informed about the number of visits each specific worker reported.

Side payments and contractual frictions After supervisors were informed about the incentives allocated to them (if any), we 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 told 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.

When asked why they opted for such low transfers, 75% of the supervisors 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). Moreover, side agreements are informal in our context, 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). The challenge of making binding commitments might compel supervisors to compensate workers for the perceived risk of non-compliance (i.e., the risk of default). Consistent with this, 55% of the surveyed supervisors mentioned that "health workers would lack trust in their payment."[10]

Overall, this suggests that, due to contractual frictions, supervisors are unlikely to make a major use of transfers to alter the initial allocation of incentives. We will document and explore this in depth in Section 4.3. Such contractual frictions are not unique to our setting but are prevalent across and within organizations and firms, as noted by Coase (1937); Gibbons (2005); Lafontaine and Slade (2007); Lee, Whinston, and Yurukoglu (2021), and particularly in developing countries, as documented by Macchiavello (2021) and Adhvaryu et al. (2020). The specific challenge of making binding commitments has been identified in Bubb, Kaur, and Mullainathan (2018) and Casaburi and Macchiavello (2019), and in the literature on relational contracts that emphasizes the importance of trust in work relationships (Macchiavello and Morjaria 2015; McMillan and Woodruff 1999). More generally, Carranza et al. (2022); Breza, Kaur, and Krishnaswamy (2023); Fehr and Schurtenberger (2018) underscore the significant role of social norms in work environments.

2.3 Data and Balance Checks

Data sources We leverage three main sources of data.

Supervisor and health worker surveys. We surveyed all 372 supervisors and 2,970 health workers across the 372 PHUs regarding their demographic backgrounds, health knowledge,

[10] 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" or because of "reporting issues."

and jobs. They were surveyed at baseline in April-May 2018 and again at endline in June-September 2019, fifteen to sixteen months post-treatment implementation. In March 2024, we conducted a follow-up survey with the supervisors (henceforth, the "supervisor perception survey") that collects information on supervisors' perceptions, specifically their perceptions of the main contractual frictions and workers' reporting behavior.[11]

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 three 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. Third, we have access to village-level information from a leaflet given to each health worker by the PHU.

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).[12] 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 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

[11]The survey covers a sub-sample of 218 supervisors who were successfully contacted via phone, evenly distributed among different treatment groups.

[12]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.

E shows that health facilities record 48 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 150 pairwise comparisons, 14 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 incentive 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, and the reporting rate). $T_{worker,j}$, $T_{supv,j}$, 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).

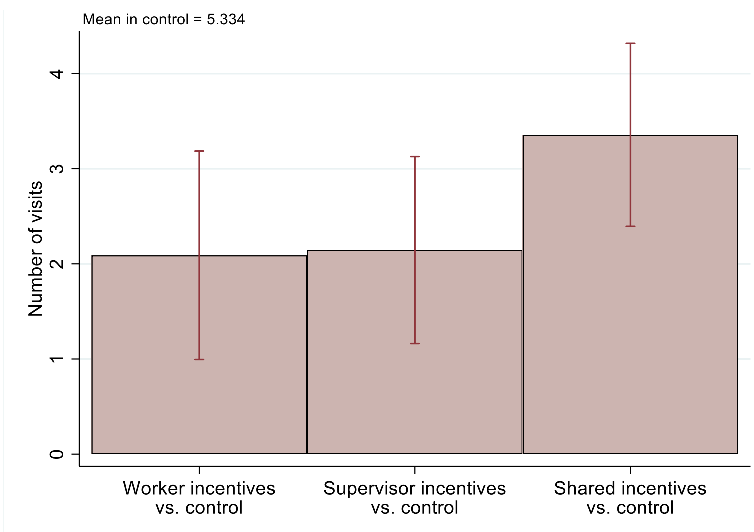
We report p-values corrected for multiple hypothesis testing using three alternative procedures in Table A.2.[13] For the sake of brevity, we do not discuss the corrected p-values in the main text as these are notably similar to the non-corrected ones we discuss below. Following Asiedu et al. (2021), we detail key aspects of research ethics in Appendix C.

[13]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.

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.[14] 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. These results are all statistically significant at the 1% level. Interestingly, offering the whole incentive to the health workers is equally effective as offering the whole incentive to the supervisor. Both interventions, however, are outperformed by the shared incentives treatment, which achieves 17% more visits overall. Relative to the control group, the boost

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

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.3).

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). 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). We show that this is not the case in Table A.4 where we explore targeting. 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.3), 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).[15] 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.[16] 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 3), presumably due to the relatively short timeframe of the experiment. The results are robust to multiple hypothesis testing corrections (Table A.2).

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

[15] 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.

[16] 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).

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 Incentive 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 SMS reports per month in T_{worker}, 19 in T_{shared}, 11 in T_{supv}, and 5 in the control group.

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.[17] 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). Over-reporting is however minimal. Only 3.6% of the 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.

Why is under-reporting so common? When queried about the reasons for not reporting visits, 50% of the workers pointed to mobile networks and phone functionality. These network and phone-related issues are not always predictable and might significantly prolong and complicate the reporting process. Our findings suggest that workers are willing to tolerate these inconveniences only if the incentive is substantial enough to outweigh the effort.[18]

[17]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.

[18]The most prevalent phone functionality issues include diminished battery life and scarce access to electricity for charging, technical failures associated with the age of the phone (including instances where the workers' phones stop functioning), and situations where households lack a phone, preventing workers from including a phone number in their SMS reports. Fewer than 3% of workers cited difficulties in reporting due to a lack of knowledge on how to submit the report, and, as we will show later, only a few workers

Contrary to other variables in this study, we are able to track the progression of reporting over time between baseline and endline. As demonstrated in Figure A.1, the differences in the number of reported visits across treatments stay fairly stable over time.

Incentive payments per actual visit and supervisor perceptions Systematic under-reporting decreases the average payment that agents receive for a completed visit. Accounting for under-reporting, the "effective" average incentive payment *per actual visit* equates to 510 SLL for the supervisor and the worker in T_{shared} (1,000 SLL at a 51% reporting rate), 820 SLL for the supervisor in T_{supv} (2,000 SLL at a 41% reporting rate), and 1,800 SLL for the worker in T_{worker} (2,000 SLL at a 90% reporting rate). This indicates that, accounting for the worker reporting behavior, workers earn three times more *per actual visit* in T_{worker} than in T_{shared}, while supervisors earn 61% more per actual visit in T_{supv} than in T_{shared}.

To understand supervisors' reactions to the treatments — an analysis we will conduct in Section 4.2 — it is important to consider their perceptions of reporting in each treatment group. As mentioned earlier, supervisors were not informed about workers' reporting rates during the experiment, which could lead to misperceptions about these rates. We assess supervisors' *perceptions* of reporting with a specific question in the endline survey: "Do you think health workers regularly report their visits?" The findings, presented in Table A.7 (column 1), indicate that supervisors correctly perceive reporting in T_{shared} and T_{supv} to be lower than in T_{worker}. This indicates that supervisors are capable of discerning differences in workers' reporting behavior across treatments. However, since this perception measure is mostly qualitative, it lacks direct comparability with the actual reporting statistics. To address this, we supplement the analysis with a more quantitative measure collected in the 'supervisor perception survey' where we asked supervisors, "For every 10 visits, how many were reported via SMS by the health workers you supervise?" Table A.7 (column 2) shows that supervisors accurately estimate the reporting rate in the worker incentives treatment, while they slightly overestimate it in the supervisor and shared incentives treatments. Despite these discrepancies and the possible recall bias in the supervisor perception survey, the relative ranking of the treatments remains consistent between perceived and actual reporting rates.[19]

Total monthly incentive payments The treatments generated clear differences in total monthly incentive payments (Table 4, columns 3-5). Workers receive the largest payments in

reported being assisted by the supervisor to report. This is not surprising, as all health workers received comprehensive training on reporting procedures at the beginning of the experiment. Less than half of the visits are reported the same day of the visit. This suggests some potential for batch reporting.

[19] The inflated beliefs of supervisors in T_{shared} and T_{supv} might be partly affected by the fact that these supervisors have information on the aggregate number of reported visits in the PHU through their monthly paycheck.

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). This is because the higher number of reports in T_{shared} offsets the larger incentives paid out in T_{supv}. 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}.

Actual visits resulting from every 2,000 SLL spent on the incentive scheme Our estimates suggest that, in comparison to the control group, the organization gains an additional 2.2 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 5.4 visits for each 2,000 SLL spent on T_{shared}.[20]

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. It is important to also acknowledge that this second driver of cost-effectiveness is driven by endogenous reporting and may thus have more limited external validity. In particular, it would not be at work in the absence of under-reporting or in settings where differences in reporting are limited.

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 incentive schemes. 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

[20]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 incentive payment is zero.

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.[21] The worker's task is to visit households, offer them health services during these visits, and report these visits to the principal. The supervisor's task is to facilitate the worker's delivery of health services through training and advice. At the beginning of the game, the principal sets the incentive scheme faced by the worker and the supervisor. The principal's goal is to maximize the number of actual household visits, but only observes the number of visits reported by the worker. Therefore, the principal is restricted to incentive schemes that reward agents for reported (rather than actual) visits.

After the principal sets an incentive scheme, the worker and supervisor make three sequential decisions. First, the worker chooses a reporting rule, determining the threshold of reporting costs below which she will report a visit. Second, the supervisor chooses a level of effort e_2, and a side payment $s \in [0, \infty)$ to offer the worker for every visit she reports. Third, the worker observes the effort choice of the supervisor and the side payment, and then chooses her level of effort e_1. At this point, output is realized and reported according to the reporting rule (see below for more details on how we model reporting). This sequential structure captures key aspects of our setting: the fact that at least some of the supervisor's training is given in advance of the worker's choice of effort, and the fact that the worker may pick her reporting rule strategically, to influence the choices of the supervisor.

Importantly, the supervisor may face a different cost of inducing worker effort compared to the principal. On the one hand, the supervisor may find it easier to induce effort due to a *motivation effect*: thanks to her close proximity to the worker, the supervisor may be more effective than the principal at observing worker effort, providing feedback and boosting worker morale. 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

[21] In our empirical setting supervisors are responsible for multiple workers. We abstract from this feature, as including it would not alter the fundamental insights but would diminish the model's clarity.

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 $\beta = 0$. This is based on our context where, if $e_1 = 0$, the supervisor cannot generate any visits, no matter how much effort she invests in 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.

The worker and the supervisor both receive a benefit of b_i for every visit that is *carried out*. This reflects the mix of personal motivation and employer- and community-reputation concerns that may drive their effort in the absence of performance-based incentives. In addition, they receive a financial incentive for every visit that is *reported*. The worker earns a monetary payment of pm per reported visit, 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}; while the supervisor is paid an incentive of $(1 - p)m$ for each visit reported by the worker. Moreover, the worker receives from the supervisor a transfer of s per visit reported, and the supervisor pays an amount zs to make this transfer.

Household visits are reported to the principal on the basis of the reporting rule chosen by the worker and a stochastic reporting cost. In particular, in the last stage of the game, each completed visit (independently) draws a reporting cost k from distribution F, which is common knowledge.[22] In the first stage of the game, the worker commits to a reporting rule: she will report all visits that cost less than some amount $t(p)$ to report. Given this rule, a fraction $\bar{q}$ of all visits is reported, at an average cost to the worker of $\bar{k}$ per visit reported, with $\bar{q} = Pr[k \leq t(p)]$ and $\bar{k} = E[k|k \leq t(p)]$.[23] Importantly, this structure allows the reporting rate to respond to the financial incentive offered to the worker — a point that we explore in detail below.

The worker and the supervisor both maximize a private expected payoff that is given

[22] These assumptions about reporting costs are natural and make the framework tractable. One possible limitation is that, since each cost draw is independent, the framework does not allow for the possibility of economies of scale in reporting. Given the substantial under-reporting, we believe that it is unlikely that significant economies of scale exist in our context. We thus view our formulation as a reasonable approximation of local conditions.

[23] We assume that the worker cannot report visits that have not taken place. This is in line with the fact that in our setting the threat of back-checks was sufficient to virtually eliminate all over-reporting. Further, in this section, we assume supervisors hold accurate beliefs on $\bar{q}$ and $\bar{k}$. However, as shown in Section 3.2, their beliefs are slightly inaccurate. When we take the model to the data in Section 5, we will allow supervisors' beliefs to be incorrect.

by the (expected) benefits received, minus the cost of effort, the cost of reporting (for the worker) and the cost of the side payment (for the supervisor). The payoff functions are as follows:

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

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

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 the effort functions for both the supervisor and the worker, as well as the 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.[24]

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 (panel A).

Second, when contractual frictions are large ($z > 1$) and there are no effort complementarities (or they are close to zero), 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.2 (panel B).

Third, when contractual frictions are large ($z > 1$) and effort complementarities are large enough, the optimal scheme is two-sided. In this final scenario, the presence of 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 shown in Figure A.2 (panel C).

[24] For this analysis, we assume that the reporting rule takes the following form $t(p) = kpm + a$, with $a > 0$ and $0 < k \leq 1$. We discuss the implications of endogenous reporting in the next subsection.

Reporting There are two central considerations on endogenous reporting that emerge from our framework. First, in the absence of strategic considerations (and without side payments), the optimal reporting rule would simply be to set $t(p) = pm$. However, the worker may opt to commit to a *higher* reporting threshold to motivate the supervisor to exert more effort. Note that strategic considerations would not prompt the worker to lower the reporting threshold, as doing so would lead to missing valuable reports and lower supervisor's effort.

Secondly, our framework sheds light on how the treatment effects would change if the organization could eliminate under-reporting, for example, by implementing penalties for missed reports or by introducing technologies that significantly reduce the costs of reporting. These policies, which increase the reporting rate, $\bar{q}$, would generate direct incentives for the supervisor to raise her effort, proportionally to the change in $\bar{q}$.[25] Therefore, if under-reporting was eliminated, treatments with the lowest initial reporting rates would likely see the most significant increase in supervisor effort. However, such policies would also affect the average cost of reported visits, $\bar{k}$, and would thus have a more nuanced impact on worker effort. Specifically, under a penalty, the worker might actually decrease her effort, especially if forced to deviate significantly from her preferred reporting rule. And, under an improvement in reporting technology, the worker would experience the largest drop in $\bar{k}$ in those treatments that have the highest initial reporting rate (and hence the largest initial $\bar{k}$). Therefore, changes in worker effort might actually benefit treatments that initially have the highest reporting rates, potentially moving in the opposite direction of changes in supervisor effort. Consequently, whether eliminating under-reporting favors treatments with the lowest or highest initial reporting rates is theoretically ambiguous. This is an empirical question that we investigate in the structural estimation section.

Reduced form test of complementarities and frictions Our framework motivates a number of reduced form tests for the presence of effort complementarities and contractual frictions. We describe these in turn here, and test them in the data in the next two subsections.

Effort complementarities — Test 1: As depicted in Figure A.2, under effort complementarities, an agent's effort does not necessarily increase monotonically with the size of the incentive offered. This is because a decrease in the direct incentive offered to the agent can be compensated by the indirect incentive arising from increased effort from the other agent. Showing that supervisor effort (which we can observe in our data) does not increase monotonically with the incentive size would provide evidence of effort complementarities.

Effort complementarities — Test 2: Intuitively, the difference in output between shared and one-sided incentives is expected to be most pronounced for worker-supervisor pairs that

[25] On top of this change in direct incentives, worker effort would also change, and this would indirectly affect the return to supervisor effort.

exhibit higher effort complementarity. Thus, showing heterogeneity in treatment effects by a plausible proxy for the strength of γ would also support the presence of effort complementarities.

Effort complementarities — Test 3: A final natural test to corroborate the presence of effort complementarities is to show that, holding everything else constant, higher supervisor effort stimulates higher worker effort and increases output.

Contractual frictions — Test 1: The model indicates that high contractual frictions reduce the level of side payments offered by the supervisor, as detailed in the formula provided in Appendix D. However, the model also indicates that side payments can be limited for reasons other than contractual frictions, particularly when $b_2 < b_1$ since the worker is inherently motivated to provide effort even without side payments.[26] Therefore, to corroborate the presence of contractual frictions, it is necessary to show that side payments are limited not only on average but also specifically in supervisor-worker pairs with $b_2 > b_1$, where limited side payments can only be explained by contractual frictions.

Contractual frictions — Test 2: Another implication of contractual frictions is that the supervisor's welfare may be maximized by an incentive allocation different from $p = 0$.[27] The rationale is that while the supervisor may highly value additional worker effort, contractual frictions may prevent her from eliciting this effort effectively, and she may be better off with an arrangement where some of the incentive is directly allocated to the worker. In the absence of frictions, this would never occur, as the supervisor would be able to use transfers to mimic any allocation of the incentive, and would never benefit more from not controlling the incentive allocation. Finding that supervisor welfare is higher under shared incentives than supervisor incentives would support the presence of contractual frictions constraining transfers.

4.2 Empirical Evidence of Effort Complementarities

This section presents the empirical tests for effort complementarities, and discusses the nature of these complementarities. The subsequent section presents the empirical tests for contractual frictions, and discusses the nature of these frictions.

Test 1: Supervisor effort does not monotonically increase with supervisor incentives. Our first test evaluates how supervisor effort varies with the level of supervisor incentives. Table 5 (column 1) shows 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

[26] 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.

[27] See Figure A.3 for an example of when this is the case.

supervisors. Specifically, we find 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}). As explained above, the fact that supervisors exert similar effort in T_{supv} and T_{shared}, despite T_{supv} offering a higher payment per actual visit, is consistent with the presence of effort complementarities.[28]

Table 5 (column 2) shows that neither treatment enhances the likelihood of supervisors inviting health workers to attend general training sessions, while worker health knowledge improves relative to the control group but only in the shared incentives treatment.[29] Table A.7 (columns 3-4) shows that the treatments also have no impact on the frequency of one-on-one meetings or the number of phone calls health workers received from their supervisors. Interestingly, Table A.7 (column 5) shows that only 9% of the health workers report that their supervisors ever helped them with SMS reporting, and this share is comparable across treatments. This suggests that the introduction of supervisor incentives did not shift supervisors' focus from essential tasks, such as training on health issues, to assisting with reporting tasks.

Test 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, having limited training and less recognition in the community, likely depend more on supervisor support. Consequently, we anticipate higher strategic complementarity of effort for these workers.

Table A.8 shows that shared incentives outperform worker incentives by a larger margin when workers have less experience, and hence, effort complementarities are likely to be higher. For workers with experience below the median (less than 4 years), shared incentives nearly double the impact on household visits compared to individual worker or supervisor incentives. However, for workers with over 4 years of experience, the effect of shared incentives on household visits is only slightly greater than that of one-sided incentives. Additionally, supervisors respond differently to the incentives based on worker experience. For less experienced workers, shared incentives enhance supervisor effort, whereas for more experienced workers,

[28]This similar effort cannot be explained by supervisors in T_{supv} anticipating a significantly lower reporting rate compared to T_{shared} (which could offset the difference in incentive levels). In fact, as shown in Table A.7 (columns 1 and 2), the perceived reporting rates are only slightly lower in T_{supv} than in T_{shared}, and certainly not half as low.

[29]The lack of effect on general training is expected as supervisors are required to organize these trainings monthly, with 99% compliance. We measure the improvement in health worker knowledge by administering to a health knowledge test to all health workers at baseline and endline; see Appendix B.8 for details.

they do not. Overall, these findings further substantiate effort complementarities as a pivotal mechanism in our experiment.[30] Adding controls for worker characteristics correlated with experience does not qualitatively affect these results (columns 2 and 4).

Test 3: Important mediating role of supervisor effort in the worker incentives treatment. As a third test, 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 of the worker incentives treatment on household visits. This gauges the impact of worker incentives on visits while keeping the level of a mediator (here, supervisor effort) fixed.[31]

In Figure A.4, 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).

Nature of the effort complementarities Complementarities may arise from the supervisor playing an "enabling" role. The supervisor's effort may help workers gain health knowledge, which in turn enables them to discuss a broader range of health topics during household visits, gain more trust within the community, increase demand for their services, and ultimately conduct more household visits. Complementarities may also arise from the supervisor playing a "coordination" role. Here, the primary contribution of the supervisor would be related to logistics: they would organize the worker's tasks in a way that makes the worker more effective. This could then lead to the worker providing a higher number of visits and, through this, may increase trust and topics discussed if these are by-products of more visits. However, in our context, supervisors are primarily tasked with training/advising workers, as opposed to performing coordinative or logistical roles. In line with this, 81% of health workers describe the supervisor's presence during household visits as "very useful" for improving the quality of their services. We thus believe that complementarities are mostly stemming from an enabling rather than a coordination role.

[30] Column (5) documents a difference in the probability of a side payment between experienced and inexperienced workers. The interpretation of this result is complicated by the fact that experienced workers conduct more visits, so it does not necessarily imply that side payments per visit (s in our model) are higher for more experienced workers. Column (7) shows that the more experienced workers do not report a higher fraction of the visits. Therefore, reporting is unlikely to explain these heterogeneous results.

[31] This quantity captures the treatment effect that would be observed if supervisor effort was fixed at an exogenous level, while worker 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.4.

A final question is whether our results could be explained by the supervisors playing a "monitoring" role. Our evidence indicates that this is not the case. If monitoring was the primary driver of our results, we would expect health workers in T_{supv} and T_{shared} to primarily visit households that had *not* been in direct contact with a supervisor (never received an accompanied visit), as these would be harder for supervisors to re-contact and verify visits. Table A.4 (column 4) contradicts this; it does not show a preference for targeting such households. Instead, the fact that T_{supv} and T_{shared} result in more visits to households that have never received an accompanied visit, and were thus never in direct contact with the supervisor, suggests that health workers in these treatments have received better training and are capable of increasing demand for their services, even when unaccompanied.

4.3 Empirical Evidence of Contractual Frictions

Test 1: Side payments are limited, even when workers have better outside options than their supervisor ($b_2 > b_1$). 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 had transferred a portion of their incentive to health workers since baseline. If they had, we then asked each health worker to assess this side payment's value (in-cash or in-kind).[32] 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 Table A.9, we examine side payments when the worker has a better outside option than her supervisor ($b_2 > b_1$), indicated by higher education levels or higher hourly wages from an outside job.[33,34] We expect the worker to exert less effort than the supervisor would

[32] 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.

[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_i and will, all else equal, exert less effort and produce lower output. In line with this, the correlation between a worker's outside option and output is negative in the data.

[34] While the variable "hourly wages from an outside job" provides a more direct measure of the outside option, it should be noted that this variable is not available for workers or supervisors who do not have an outside job. Consequently, the analysis is limited to a subset of the workforce, resulting in a smaller sample size.

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) show that, even in these (relatively rare) cases, side payments are modest across all treatments. This indicates that contractual frictions must play a role in explaining the limited use of side payments.

Test 2: Supervisor welfare is higher with shared incentives. We present two pieces of suggestive evidence showing that supervisor welfare is higher under shared than supervisor incentives, pointing to the presence of contractual frictions that prevent supervisors from fully capitalizing on the supervisor incentives treatment. First, we assess 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: $((1-p)m - zs)y)$, (ii) the cost of effort ($c(e_2)$), and (iii) the direct benefit generated by household visits ($b_2 y$). In Table A.9 (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) appear to be approximately equal in the two treatments, while component (iii) is significantly higher under shared incentives. While suggestive, this indicates that supervisor welfare is higher under shared incentives than under supervisor incentives.[35]

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 through an incentivized question, nearly two-thirds of supervisors opted for an equal division of incentives between themselves and the workers.[36] These findings reinforce the notion that supervisor welfare is likely higher under the shared incentives, and that contractual frictions are at play.[37]

[35]This analysis assumes that the cost function is not overly sensitive to minor changes in effort levels. If it was, the small differences in effort observed across treatments could result in significant cost variations, making welfare calculations considerably more complex.

[36]Supervisors were asked to decide the allocation of a total incentive of SLL 2,000 between themselves and their health workers. They were given five allocation options: SLL 2,000, 1,500, 1,000, 500, or 0 to themselves. They were informed that there was a small chance, 1 in 100, that their chosen distribution would be implemented at their PHU for six months after the endline.

[37]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.

Nature of the contractual frictions When asked about the frictions, most supervisors highlighted social norms against redistribution, and some also pointed to the challenge of making binding commitments because health workers would lack trust in their payments (see Section 2.2 for details). Interestingly, these frictions are present in our setting despite the long-standing relationships between supervisors and health workers. If these incentives had been introduced in settings with less established work relationships, the frictions might have been even stronger due to lower levels of trust and stronger adherence to established norms.

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 incentive 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 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 11% (resp., 14%) 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.[38] Finally, we show that one-sided incentives do not decrease the job satisfaction of unincentivized agents relative to the control group. Table A.10 shows that there is no evidence that unincentivized agents (workers in T_{supv} and supervisors in T_{worker}) 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.[39] All in all, these three

[38] The fact that only a few workers were aware of the incentives for supervisors also eliminates "altruism" as an alternative explanation — e.g., the notion that workers exert more effort/report more visits in T_{supv} than in the control group out of altruism towards their supervisor.

[39] We find similar results for the subsample of workers with above-median experience, who may perceive T_{supv} as particularly unfair.

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.10 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, or a religion) and to find the work environment competitive: columns (7) and (8) in Table A.10. 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 those provided in the shared incentives treatment 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, the finding that supervisor welfare is higher under shared incentives than under supervisor incentives, and the finding that the shared incentives treatment is particularly effective for less experienced workers.

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. 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.5 we show non-parametrically that the impacts of incentives on household visits are approximately constant over the distribution of worker and supervisor wealth.

Alternatively, there may be a discontinuity in the cost function. Here, the marginal cost of effort would need to rise sharply at the level of effort agents provide for the incentive provided in the shared incentives treatment. Such a scenario might occur if the distance households-worker or supervisor-worker is bimodally distributed, which is not the case in our setting. Moreover, Panel B of Figure A.5 shows non-parametrically that the treatment effects on visits and supervisor effort, relative to the distribution of household-worker distance (a proxy for the worker's cost of visiting a household) and worker-supervisor distance (a proxy for the supervisor's cost of training or monitoring a health worker), are approximately constant.

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, and for supervisor misperceptions about workers' reporting behavior (as in Appendix D.1). We first 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.[40] 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 $s = a - \frac{z+1}{2z}mp$.[41] This suggests that the slope of a regression line of side 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 $1 * m$. As frictions grow, the slope drops below

[40] 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.

[41] Where $a = \frac{b_2+\hat{\hat{q}}m-zb_1+\hat{\hat{q}}\hat{\hat{\bar{k}}}z}{2\hat{\hat{q}}z}$.

$1 * m$ and approaches $0.5 * m$ from above. This result is intuitive: the stronger the frictions, the less responsive the side payment is to changes in p.[42]

In the model we estimate, we allow workers to under-report visits, and supervisors to have an inaccurate perception of this under-reporting. We do so in three steps. *Step 1:* For each treatment, we calculate $\bar{k}$, the average reporting cost for the visits that are actually reported (this differs from the mean of the reporting cost distribution as visits whose cost exceeds a threshold are not reported). Specifically, we assume that the reporting costs follow a uniform distribution: $k \sim U(a, b)$. Additionally, we assume that workers in the worker incentives treatment use the non-strategic reporting threshold $t(1) = m$, while workers in the shared incentives treatment use a threshold that is higher than the non-strategic one. For our main results, we assume this threshold is 10% higher: $t(.5) = 1.1 * 0.5 * m$,[43] but we will show that the findings remain robust under different assumptions about the extent of this increase. Under these assumptions, we can then (i) back out the parameters of the distribution of reporting costs $U(a, b)$, (ii) calculate a reporting threshold for the supervisor incentives and control groups (i.e., the threshold that rationalizes the empirical reporting rate), and (iii) compute a value of $\bar{k}$ for each treatment. *Step 2:* We introduce supervisor misperceptions. We assume that supervisors hold reporting rate beliefs as in the data (Table A.7, column 2). Given that the cost distribution is assumed to be common knowledge, we then calculate the perceived reporting costs implied by supervisor reporting rate beliefs in the four experimental conditions.[44] *Step 3:* We feed the true and perceived reporting rates and reporting costs into the model.

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

[42]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 a, 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.

[43]When asked if "any worker under their supervision had ever threatened not to report their visits," only about 10% of supervisors answered "yes."

[44]While we take supervisor beliefs as given and we do not explicitly model belief formation, we note that beliefs may depart from true parameters because of imperfect communication (e.g., strategic persuasion) or information processing biases.

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.11 and A.12).[45] This model version wrongly predicts that worker incentives generate the largest increase in visits, and the value of its minimized loss function is more than five times larger than that of the model that features effort complementarities. 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 101% larger when the supervisor exerts the control level of effort, and 146% 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 a lower baseline incentive ($b_2 < b_1$).[46] Overall, this indicates that interventions

[45] See Appendix D.2 for the formal model without effort complementarities.

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

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 incentives, the supervisor will under-provide her key support to the worker.

Importantly, our core estimates and our simulated results are robust to different assumptions regarding strategic reporting. Our headline model assumes that in T_{worker} workers use the non-strategic reporting threshold, while in T_{shared} they use a threshold that is 10% above the non-reporting threshold. In Table A.13, we show that similar parameter estimates and simulated moments are obtained under a range of different assumptions. In particular, in all models, the estimated value of γ indicates that supervisor effort considerably raises the returns to worker effort. Further, in all models, shared incentives maximize the number of visits provided.

5.3 Counterfactual Policies

We conduct a series of counterfactual policy experiments to explore: (i) the effects of moderating or eliminating worker under-reporting and supervisor misperceptions; (ii) the optimal allocation of incentives between workers and supervisors; and how this optimal allocation changes as the complementarity of effort diminishes, and as the size of the incentive payout increases.

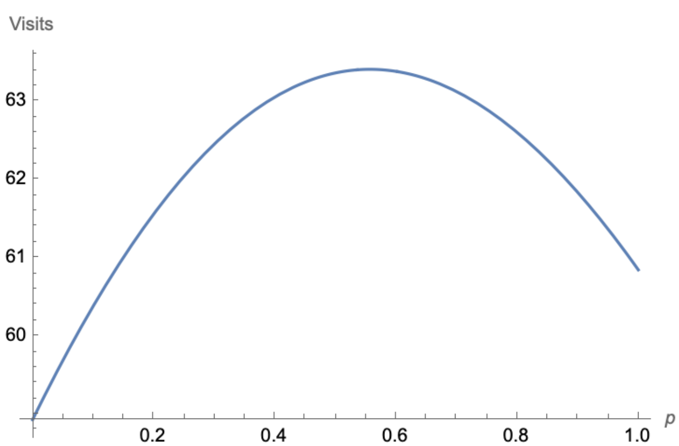
Removing under-reporting and supervisor misperceptions Our first sets of experiments show that shared incentives would remain the optimal policy even if the organization moderated or removed under-reporting, and eliminated supervisor misperceptions. Specifically, we examine two hypothetical policies that the organization could implement for this purpose, with results presented in Table A.14.

First, we consider a policy that ensures workers maintain a uniformly high reporting rate across all treatments, for example, a policy that introduces a penalty for under-reporting. In Experiment 1, the reporting threshold is set to a level that ensures, in all groups, the 90% reporting rate observed in T_{worker}. In Experiment 2, the reporting threshold is established to guarantee 100% reporting in all groups. In both instances, we assume these thresholds are common knowledge, preventing any misperceptions by supervisors. The results, presented in columns 2 and 3, show that the shared incentives treatment remains the most effective treatment. Also, supervisor effort increases as a result of higher reporting, while worker effort decreases. These changes in efforts roughly balance each other out, leaving shared incentives as the optimal treatment.

Second, we consider a policy that reduces reporting costs, e.g., by introducing a new reporting technology. Specifically, we assume that the organization shifts downward the upper bound of the cost distribution in a way that halves the mean of this distribution, and this change is common knowledge. In this case, the number of visits goes up substantially in all treatments (as workers exert a lot more effort), but the relative differences between the treatments are largely unaffected: see Experiment 3 (column 4). Again, this is because changes in worker and supervisor effort balance each other out (specifically, worker effort increases the most in those treatments where supervisor effort increases the least). In sum, our result appears robust to the removal of under-reporting and misperceptions, since the responses of supervisor and worker effort to interventions that raise reporting tend to offset each other.

Optimal allocation of the incentives Next, we investigate the optimal allocation of the incentive to maximize household visits.[47] As shown in Figure 2, to maximize household visits, the worker should be offered 56% of the overall incentive, which is very close to the equal share we offered in the shared incentives treatment. This suggests that, given the strong complementarity and large contractual frictions we have estimated, the optimal incentive scheme is one that rewards both the worker and the supervisor with a similar payment.

Figure 2: Optimal Incentive p^*

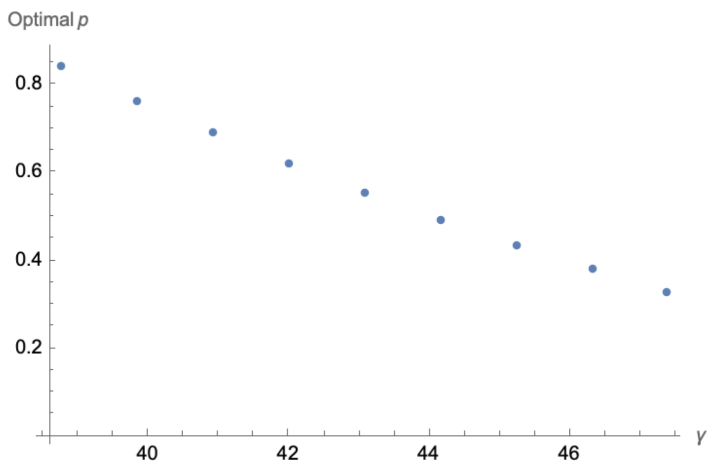


Would the near-optimality of shared incentives hold if we allowed the complementarity parameter to change? We examine this question with two additional counterfactual experiments. In the first counterfactual experiment, we examine the optimal share of the incentive offered to the worker (p^*) across varying levels of effort complementarity, as shown in Figure 3. 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. Quantitatively, if the complementarity parameter was 10% lower than what we estimate, the optimal incentive scheme would give about 80% of the piece rate to the worker. This suggests that in

[47] We have shown that differential reporting and supervisor misperceptions do not affect the optimal allocation of the incentives. For simplicity, we thus proceed by assuming uniform reporting rates across all treatments, consistent with those in T_{worker}, and that supervisors have no misperceptions about reporting.

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 output production — the optimal incentive scheme allocates the largest share of the piece rate to the supervisor.

Figure 3: Optimal Incentive p^* by Complementarity γ



In the second counterfactual experiment, we investigate how the effectiveness of shared incentives evolves over a worker's tenure, providing insights into the dynamics of our results. We present these findings in Figure A.6, where we analyze at what point in a worker's career the shared incentives treatment ceases to be the most effective strategy, as the complementarity between worker and supervisor effort declines due to the accumulated experience of the worker.[48] We find that shared incentives outperform worker incentives for up to 15 years of a worker's experience; beyond this point, worker incentives become more advantageous. Shared incentives continue to outperform supervisor incentives for an even longer duration, as indicated by the nearly flat slope of the orange dotted line in the figure. Given that many workers are likely to have exited the company by 15 years of experience — indeed, we find that only 2% of workers in our sample have that many years of experience — this suggests that although a supervisor's ability to help the worker perform does gradually deplete over time, shared incentives remain optimal for most of a worker's career.

Finally, we explore in Figure A.7 how the optimal share of the incentive offered to the worker (p^*) varies with the size of the payout (m). The main insight that emerges is that, as the size of the payment decreases, the optimal share of the incentive given to the worker

[48]To obtain the figure, we conduct a new estimation exercise where we match three moments capturing the difference in treatment effects between high and low experience workers, as in Table A.8. Our estimate suggests that this decline can be best reconciled with a value of γ that decreases by about 0.26 per additional year of worker experience.

increases. This is intuitive, since the supervisor has a much larger cost of effort. Hence, small payments may be insufficient to motivate her (and the sensitivity of p^* to changes in m is thus particularly pronounced for low levels of m).

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 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.[49] And they call into 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.

What we identify as the optimal incentive allocation in our context likely extends to many other organizations, particularly those where the role of supervisors is central to the production process — and effort complementarities are hence large — and those where contractual frictions hinder the redistribution of incentives across various levels. However, the optimal allocation might differ in organizations where the primary function of supervisors is oversight, such as those where supervisors focus on administrative duties and operations, or in environments with minimal contractual frictions, for example, family-owned businesses built

[49] 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. 26% for the shared incentives and 3% for the supervisor incentives. See Appendix E for more details about the prediction survey.

on longstanding trust, where redistributing incentives is presumably more common. It may also differ in organizations where supervisors are responsible for setting performance pay or making other personnel decisions (hiring, firing, promotions). Our structural results highlight the necessity of customizing incentives within each organization, taking into account the specific levels of effort complementarities and contractual frictions.

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. In other settings, (i) the cost of performing these checks may discourage organizations from introducing incentive schemes like ours, or (ii) organizations that use incentives may be overly reliant on unchecked worker reports and, as a result, may make the wrong inferences about the optimal allocation of incentives. As digital technologies improve, the costs of monitoring worker self-reports will likely decrease, enabling more organizations to set up incentive schemes like ours and enabling those that are already experimenting with such policies to form more accurate beliefs on the optimal allocation of incentives (Muralidharan et al. 2021; Dodge et al. 2022; Adhvaryu, Nyshadham, and Tamayo 2022).

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.

Besley, Timothy, Robin Burgess, Adnan Khan, and Guo Xu. 2021. "Bureaucracy and Development." *Annual Review of Economics* 14.

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.

Bossuroy, Thomas, Clara Delavallade, and Vincent Pons. 2024. "Biometric Monitoring, Service Delivery and Misreporting: Evidence from Healthcare in India." *Review of Economics and Statistics* :1–46.

Breza, Emily, Supreet Kaur, and Nandita Krishnaswamy. 2023. "Coordination without Organization: Collective Labor Supply in Decentralized Spot Markets." Working Paper.

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.

———. 2023a. "Personalities and Public Sector Performance: Evidence from a Health Experiment in Pakistan." Working Paper.

———. 2023b. "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.

Carranza, Eliana, Aletheia Donald, Florian Grosset, and Supreet Kaur. 2022. "The Social Tax: Redistributive Pressure and Labor Supply." Working Paper.

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. 2024. "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.

Fehr, Ernst and Ivo Schurtenberger. 2018. "Normative Foundations of Human Cooperation." *Nature Human Behaviour* 2 (7):458–468.

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.

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. 2024. "Policy-Making, Trust and the Demand for Public Services: Evidence from a Mass Sterilizations Campaign." *American Economic Journal: Economic Policy*. Forthcoming.

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.

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

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

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

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.

Sen, Ritwika. 2024. "Supervision at Work: Evidence from a Field Experiment." Working Paper.

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	0.317
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 Incentive 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)
	In-the-field training	*General training*	*Knowledge improvement*
Dep. Var.	% accompanied household visits	Supervisor organized general training in the past month = {0, 1}	Difference in health workers' knowledge between baseline and endline
Worker incentives	0.030	0.004	0.158
	(0.022)	(0.005)	(0.116)
Supervisor incentives	0.057**	0.006	0.063
	(0.023)	(0.005)	(0.113)
Shared incentives	0.062***	0.003	0.266**
	(0.021)	(0.005)	(0.121)
Unit	Worker	Worker	Worker
Observations	2,919	2,864	2,927
Mean dep. var.	0.204	0.994	-0.313
Mean dep. var. in Control	0.164	0.991	-0.433
p-value Worker = Supervisor	0.293	0.443	0.372
p-value Supervisor = Shared	0.846	0.463	0.074
p-value Worker = Shared	0.181	0.916	0.350

Notes: Data source is the household survey in col. (1), aggregated at the health worker level. Data source is the health worker survey in the other columns. The dependent variable in col. (1) is equal to the share of household visits accompanied by the supervisor. 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

Complementarity γ	43.1
	(4.5)
Worker baseline incentive b_1	76.4
	(19.7)
Supervisor baseline incentive b_2	49.9
	(15.3)
α	6.7
	(0.7)
Calibrated friction z	3.61
Δ in marginal product of worker effort (shared incentive)	146%
Δ in marginal product of worker effort (control)	101%
Total worker cost of effort (control)	1555.4
Total supervisor cost of effort (control)	1022.3

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 99^{th} 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 $\gamma = 43.1$ 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	0.203
Supervisor effort in supervisor incentives treatment	0.225	0.228
Supervisor effort in shared incentives treatment	0.228	0.226
Supervisor effort in control group	0.164	0.157
Output in worker incentives treatment	59.679	60.841
Output in supervisor incentives treatment	58.896	60.713
Output in shared incentives treatment	66.895	63.637
Output in control group	41.040	40.979
Value loss function	3.4	

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