Journal of Development Economics 163 (2023) 103096

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Regular article

Motivating volunteer health workers in an African capital city☆

Mattia Fracchia [a,b], Teresa Molina-Millán [c,b,d], Pedro C. Vicente [a,b,e,f,*]

[a] Nova School of Business and Economics, Portugal
[b] NOVAFRICA, Portugal
[c] University of Alicante, Spain
[d] IZA, Germany
[e] CEPR, United Kingdom
[f] BREAD, United States of America

ARTICLE INFO

Dataset link: http://dx.doi.org/10.17632/ppbfj4ttrt.1

JEL classification:
O12
D91
I15

Keywords:
Non-financial incentives
Motivation
Community Health Workers
Social status
Task significance
Guinea-Bissau
Africa

ABSTRACT

Community Health Workers (CHWs) are central to health systems. Still, they are typically unpaid volunteers in Sub-Saharan Africa. This paper follows all the CHWs in the capital city of Guinea-Bissau, who are non-salaried, and tests the impact of non-financial incentives on health indicators. We analyze two randomized interventions for CHWs: (i) an honorific award aimed at raising their social status; (ii) a video treatment aimed at increasing their perceived task significance. While employing administrative and survey data, we find that the social status intervention, differently from the task significance one, causes clear improvements in household health, particularly for young children.

1. Introduction

'Ensure healthy lives and promote well-being'. The United Nations currently prescribes this goal as one of the most fundamental ones for human kind. Be it a matter of fundamental human rights or just because healthy people are more productive, this is a shared objective for citizens and for public policy around the world. At the same time, substantial health challenges persist (Report of the UN Secretary-General on the Sustainable Development Goals, 2020): in the latest years available, 5.3 million children under 5 years of age died with almost half of these deaths occurring in the first month of life; close to 295 thousand women around the world died due to complications of pregnancy and childbirth. Crucially, almost all of these deaths are avoidable and the majority occurs in Sub-Saharan Africa.

This grim picture of health in Sub-Saharan African countries hides substantial improvements made on health indicators over the last few decades (Glassman and Temin, 2016). Important improvements in the supply of health services (infrastructure, personnel) were accompanied by increased education for health and demand for health services by the population. Community Health Workers (CHWs) have made a central contribution in linking their communities to the health system — see

☆ We thank the editor, two referees, as well as seminar and conference participants at DENeB (Berlin), IFPRI, NEUDC-Dartmouth, Alicante, Ca' Foscari, Paris School of Economics, Paris-Dauphine, Milan-Bicocca, and NOVAFRICA for useful comments. We are grateful to a large number of individuals who have made this project possible. First of all, we thank Helena Areal, Patrícia Carvalho, and Patrícia Maridalho, whom have headed the efforts by NGO VIDA in this project; Rita Aguiar, David Bastos, Ana Gaspar, Rita Pais, Luísa Rocha, and Ana Rita Simões also provided important contributions. We are also grateful to Sara Almeida, Fulvia Budillon, Giulia Caivano, Riccardo D'Ercole, Deoclécio Marinho, Frederica Mendonça, Melisa Rodrigues, Xaver Schenker, Maria Sousa, and António Vicente for outstanding field coordination and/or research assistance. We also convey a word of appreciation to all enumerators we worked with in this project. We acknowledge funding from the European Union, UNICEF, United States of America, Instituto Camões, Portugal, Fundação Calouste Gulbenkian, Portugal, USAID, United States of America, and Fundação para a Ciência e a Tecnologia, Portugal. Teresa Molina-Millán acknowledges financial support from Generalitat Valenciana CDEIGENT/2020/016 and from the Spanish Ministry of the Economy and Competitiveness, Grant ECO2017–87069-P. Ethics clearance was given by Universidade Nova de Lisboa as well as the National Institute of Public Health of Guinea-Bissau. A pre-analysis plan was registered for this project at the American Economic Association with the reference AEARCTR-0003399. All remaining errors are our responsibility.

* Correspondence to: Nova School of Business and Economics, Universidade Nova de Lisboa, Rua da Holanda 1, 2775-405 Carcavelos, Portugal.
E-mail addresses: mattia.fracchia@novasbe.pt (M. Fracchia), teresa.molina@ua.es (T. Molina-Millán), pedro.vicente@novasbe.pt (P.C. Vicente).

https://doi.org/10.1016/j.jdeveco.2023.103096
Received 22 April 2022; Received in revised form 22 March 2023; Accepted 24 March 2023
Available online 29 March 2023

Haines et al. (2007), Christopher et al. (2011), and Gilmore and McAuliffe (2013). Indeed, CHWs are typically community members who are tasked with providing the referred link while giving direct health support to their communities.[1] CHWs facilitated the spread of simple technologies allowing the prevention and treatment of prevalent health problems. As a result, many avoidable deaths have in fact been avoided (Björkman Nyqvist et al., 2019).

CHWs are however atypical health workers. They generally do not have medical or nursing formal education: just a few weeks of specialized training are needed in most CHW programs. And the majority is not salaried. Data for 34 Sub-Saharan African countries on 322,199 CHWs show that 69 percent of these workers do not receive a regular salary, and that 46 percent do not receive any type of monetary compensation.[2] Most CHWs are therefore voluntary, part-time workers. Many public health specialists have proposed the path of professionalization of CHWs: not surprisingly, many of the most successful CHW programs are those with full-time, salaried workers, fully integrated in the formal health system.[3] Still, these programs are not affordable in many countries without significant inflows of foreign aid (Taylor et al., 2017).

In this paper we ask how volunteer, part-time CHWs can be incentivized to do a better job without employing financial incentives. Specifically, we consider two main possibilities to incentivize CHWs. First, by improving the social status of CHWs in their communities. Second, by increasing the perceived task significance of CHWs. We follow the full contingent of 1015 CHWs in Bissau, the capital city of Guinea-Bissau during 2017–2019. Note that these CHWs cover the whole population of Bissau, with each of them assigned to a pre-defined grouping of households. Guinea-Bissau is one of the poorest countries in the world. It faces significant challenges in terms of health indicators, even for regional standards. CHWs constitute a central component of the country's strategy to address these challenges. However, like in most Sub-Saharan Africa, the CHWs in Bissau do not receive a regular salary and so lack clear incentives.

We employ a field experiment where CHWs faced randomized interventions. The first main intervention is defined by CHW access to honorific awards for good performance, which is assigned to CHWs at the neighborhood level. These awards are distributed in ceremonies with the presence of local authorities and representatives of international organizations; text messages announcing the award then follow to the households assigned to the awarded CHWs. The objective is to increase the social status of awarded CHWs in their communities. The second main intervention is the visualization, assigned at the CHW individual level, of a video recorded from the perspective of a CHW on a day of work, where the task significance of the CHW is emphasized through an interactive decision and the observation of its dramatic consequences.

Apart from aiming to assess the impact of these two main incentive interventions, we designed a few secondary sources of variation in our experimental design, which included variations of the video treatment on task significance, an information campaign, and the interaction between all treatments allowed by the cross-randomization we implemented. We can also identify the Local Average Treatment Effects (LATE) of the video treatments by employing data on compliance with those treatments.

We measure treatment effects on CHW performance and health-related outcomes of the corresponding households, covering the full chain of effects from interventions to factual household health. To assess CHW performance, we employ administrative data from our implementing partner on monitoring of CHWs, as well as household survey data on home visits and on household satisfaction, collected both face-to-face and by phone. We collected health-related indicators through administrative data on health-related appointments from the logbooks of all health centers and major hospitals in Bissau, as well as from household surveys. The latter include measures at the level of the household, the child under 5 years old, and the recently pregnant mother. We also collected data directly from CHWs through face-to-face surveys. Baseline survey data (namely in the case of households) allows improving the statistical power of the analysis.

Our main result is that the social status intervention improved the performance of CHWs and household health. Specifically, we identify significantly positive effects on CHW performance in terms of conducting home visits, health indicators at the level of the household, namely driven by knowledge of health practices, health of children under 5 years old, including the probability that children have all main five vaccinations (using data from hospitals' and health centers' logbooks), and peri-natal care (both from survey and logbooks data), including a higher probability of giving birth at a health facility and better post-natal care. The magnitudes of these effects vary between 0.07 and 0.2 standard deviation units. Overall, the pattern of results is consistent with a mechanism by which encouraging CHWs to elicit effort on learning leads to better performance in a wide range of tasks, including incentivized and non-incentivized activities, i.e., beyond those considered for the honorific awards.

We do not find systematic treatment effects for the task significance video. The clearest effects we encounter are positive impacts on vaccination of children under 5 years old. Note that the LATE of this treatment yields stronger effects, namely on measures of direct CHW performance — this pattern suggests that effects are concentrated on the compliers to the video treatment. On secondary hypotheses, it is difficult to distinguish the variations of the task significance video, there are no significant effects of the information campaign, and only limited evidence in favor of complementarity between incentive types.

Our paper relates to different strands of the literature. First of all, it relates to the theoretical literature on incentives and motivation.[4] Akerlof and Kranton (2000) model agents as motivated by the will to adopt an identity, which affects economic outcomes. In this world, identity is an important supplement to monetary compensation, which, as a sole motivator, can be both costly and ineffective (Akerlof and Kranton, 2005). The same idea is present in Bénabou and Tirole (2003) who model the mechanism underlying the phenomenon of financial incentives crowding-out motivation. Related, the concept of mission, as opposed to profit and as a fundamental driver of motivation and performance, is proposed by Besley and Ghatak (2005).[5] Bénabou and Tirole (2006) are closest to the conceptual structure of our paper in that they propose a model where agents' behaviors reflect a combination of three types of motivation: extrinsic or related to financial incentives

[1] The official definition offered by the International Labour Organization in their International Standard Classification of Occupations is: 'CHWs provide health education and referrals for a wide range of services, and provide support and assistance to communities, families and individuals with preventive health measures and gaining access to appropriate curative health and social services. They create a bridge between providers of health, social and community services and communities that may have difficulty in accessing these services'. There is however a range of definitions employed in practice — see Olaniran et al. (2017).

[2] Data Tool, One Million Community Health Workers Campaign, 2019.

[3] See for instance Zambruni et al. (2017) and 'WHO Guideline on Health Policy and System Support to Optimize Community Health Worker Programmes', WHO, April 2018.

[4] Many authors base their definition of motivation on Self-Determination Theory, developed in psychology by Ryan and Deci (2000b,a), which underlines the role of three innate psychological needs — competence, autonomy, and relatedness. For a comprehensive survey on the topic, we refer to Gagné (2014).

[5] Carpenter and Gong (2016) provide experimental evidence that matching the mission of the organization is a strong motivator for workers.

(which is well defined but outside the scope of this paper), reputational or related to social status, and intrinsic.[6]

Our study contributes to the vast and diverse literature on incentives in organizations (Kamenica, 2012; Ashraf and Bandiera, 2018). Our social status intervention relates directly to recent field experiments finding that social status and recognition have powerful effects on a range of behaviors including pro-social ones.[7] In this context, purely symbolic awards for voluntary work have shown significant positive effects on effort and performance (Kosfeld and Neckermann, 2011; Neckermann et al., 2014).[8] Furthermore, these awards can be effective even if they have no impact on future career opportunities (Gallus, 2017).[9] Close to our intervention, Cotofan (2021) finds that repeated public praise for the best teachers in a sample of Romanian schools positively impacts student performance. Our intrinsic motivation intervention, focused on task significance,[10] has received less attention in the economics literature.[11] However, from psychology, Grant (2008) provides evidence that manipulations of perceived task significance can have a positive effect on performance at work in different contexts. Following the work of DellaVigna and Pope (2017), who have tested the impact of a number of non-monetary inducements from psychology in a large-scale, real-effort experiment, taking task significance as a trigger of pro-social behavior to the real world is one of the contributions of our paper.

Our paper is related to recent empirical contributions to the study of incentives and motivation in the context of health workers and developing countries. The contributions by Ashraf et al. (2020) and Deserranno (2019) are devoted to analyzing the selection of health workers as well as the potential tradeoff between pro-sociality and talent. The first looks at recruitment of nurses in Zambia to test whether career benefits attract talent at the expense of pro-social motivation. The second estimates the effect of financial incentives on job candidates' perceptions about a new health-promoter position in Uganda. Both studies find that financial incentives reduce the ability to recruit the most socially motivated agents, although the first only reports this pattern for low-talented individuals.

A few studies test different incentive schemes for existing health workers. Björkman Nyqvist et al. (2019) test a novel approach to health care delivery in Uganda by incentivizing sales agents to conduct home visits, educate households on essential health behaviors, provide medical advice and referrals, as well as to sell preventive and curative health products. This mix of standard incentives with entrepreneurial incentives resulted in substantial health impacts. Ashraf et al. (2014a) compare the effect of financial and non-financial rewards on the performance of hairdressers and barbers (working on-the-job as health agents) in promoting HIV prevention and selling condoms in Zambia. Incentives similar to our social status treatment are found to be more effective than financial rewards at improving the direct performance of the referred agents. In a study with a one-year training program for health workers in Zambia, Ashraf et al. (2014b) unbundle public (non-financial) awards and find that employer recognition and positive social visibility are the main drivers of trainees' test scores. Our paper goes beyond these contributions in that it presents the full spectrum of measurement of treatment effects from multi-task CHW performance to household health.[12] Incentives on intrinsic motivation of health workers have also emerged as effective in improving the performance of those workers in a few contemporaneous studies to ours.[13]

Our paper adds to the literature by (i) studying the full contingent of CHWs in the capital city of Guinea-Bissau, by (ii) contrasting the impacts of non-financial incentives based on social status with those based on intrinsic motivation driven by task significance, as well as by (iii) adopting a set of administrative and survey measures of not only the direct performance of CHWs, but also and crucially, a comprehensive range of households' health outcomes.

The paper is organized as follows. In the next section we give the context of our experiment. Section 3 is dedicated to experimental design, including a description of treatments, randomization, sampling, and measurement. The following section explains our hypotheses and estimation strategy. Subsequently, we show our econometric results which are structured in balance, main (aggregated) treatment effects, disaggregated results, secondary treatment effects, LATE of the video treatments, as well as additional results and robustness. Section 6 concludes.

2. Context

Guinea-Bissau is one of the poorest countries in the world with more than two thirds of the population living below the poverty line. Its GDP per capita in current USD was 697 in 2019, ranking 199 in 213 countries.[14] The population of Guinea-Bissau is estimated at 1.92 million of which 56 percent live in rural areas and 30 percent live in the capital city of Bissau. The health situation in the country is characterized by the persistence of high morbidity and mortality in maternal, newborn, as well as child and youth health.[15] The country's life expectancy is 58 years, which is lower than the average in Sub-Saharan Africa.[16] The main causes of death are lower respiratory infections (accounting for 12 percent of deaths), maternal and neo-natal complications (12 percent), HIV/AIDS (11 percent), malaria (8 percent), and diarrheal diseases (6 percent).[17] The country's health system faces persistent challenges related to inadequate supply of health workers, low public spending, and poor infrastructure.[18]

In order to address the significant difficulties faced by the health-care system of Guinea-Bissau, international organizations have strongly

[6] We define an individual to be intrinsically motivated if willing to perform a task even in the absence of any reward or monitoring, similarly to Gagné and Deci (2005).

[7] See the recent review by Bursztyn and Jensen (2017). Indeed, social recognition has been found to play an important role in very diverse settings, ranging from academic research in economics (Chan et al., 2014) to fighter pilots squadrons during World War II (Ager et al., 2022).

[8] Ariely et al. (2009) show that social recognition is important for pro-social behaviors. Bradler et al. (2016) find that unannounced recognition in the workplace improves performance by non-recipients.

[9] Like Dana et al. (2007) show in lab games, reputational motivation could be related to self-image. Their experimental evidence shows that subjects behave fairly because they intrinsically dislike appearing unfair, either to themselves or others.

[10] Task significance signals agents that their efforts have an impact on the well-being of other people (Grant, 2007).

[11] A related theoretical literature in economics focused on the motivation of public servants (Francois, 2000; Prendergast, 2007).

[12] Gauri et al. (2021) underline the possibility that effectiveness of public awards is context-specific.

[13] In a rural health worker program in India, Lee (2018) finds that a novel mobile app that makes effort more intrinsically rewarding leads to a substantial increase in workers' performance (home visits). Khan (2020) finds in Pakistan that making salient the organization's public health mission can improve health workers' performance across incentivized (home visits) and non-incentivized tasks. In contrast, piece-rate financial incentives improved performance only on incentivized tasks. Finally, Banuri et al. (2018) find that task-based motivation beats mission-based motivation in eliciting effort among medical and nursing students in Burkina Faso.

[14] World Development Indicators, World Bank, 2020.

[15] At the time of the launching of this project (2016), Guinea-Bissau scored 7th globally for Neonatal Mortality Rate with 38 deaths per 1000 live births, 8th for Maternal Mortality Rate with 679 deaths per 100,000 live births, and 17th for Under 5 Mortality Rate with 87 deaths per 1000 live births (UNICEF Data Warehouse, 2021).

[16] Latest available years, World Development Indicators, World Bank, 2020.

[17] Latest available years, World Health Organization, 2019.

[18] See 'Guinea-Bissau: Service Delivery Indicators Report–Health', World Bank, June 2019.

supported the introduction of CHWs in the country. This is in line with World Health Organization policy[19] and recent efforts across Sub-Saharan Africa — see for instance the One Million Community Health Workers Campaign.

In this context, CHWs were introduced for the first time in the capital city of Bissau (Autonomous Sector of Bissau) in 2017. This effort was formally conducted by the Ministry of Public Health of Guinea-Bissau in collaboration with the European Union and UNICEF, which were the main funders. International NGO VIDA, which has been present in the health sector of the country since the 1990s, managed this contingent of CHWs. In close coordination with international guidelines, this CHW program focuses on improving maternal, newborn, and child health. It consists in training community members on a series of simple health practices, who then provide regular household visits within their communities. Every month, CHWs are expected to visit each one of the households within a pre-defined group of households they are allocated to (which is typically composed of around 50 households per CHW). CHWs are trained to give health education, refer households to the health centers, and offer simple medical treatments during their visits to the households.

CHWs were recruited through a local selection process organized in collaboration with community representatives. The CHW position was advertised as a volunteering one with no mention of any monetary compensation or career opportunities. Candidates had to be aged at least 18 years and to have at least nine years of education. As part of selection procedures, VIDA conducted a face-to-face interview and a test evaluating writing skills of the eligible candidates. After recruitment, agents received training on basic health practices for 21 days in January 2017, and, after the program started in March 2017, they also attended refresher training sessions every month.[20]

Like in most other health worker programs in Sub-Saharan Africa, CHWs in the Bissau program are treated as volunteers.[21] The setting of this study is thus representative of many CHW programs in the region, where community volunteers are in charge of following a group of pre-assigned households on a regular basis and where designing effective incentive schemes remains a major challenge.

3. Experimental design

3.1. Treatments

The interventions we followed in this project relate to non-financial incentives of the CHWs in the city of Bissau. Our field experiment included two incentive treatments and an information intervention. First, we analyze incentives targeting increased social status of the CHW in his/her community. Second, we devote attention to incentives aimed at increasing the significance attributed by CHWs to their role/task. Third, we also follow an information campaign about the role of CHWs in their communities. There were three rounds of treatment for each one of the interventions, i.e., we repeat each of the interventions three times over a period of 9 months. Figure B1 in the Online Appendix presents a timeline of the interventions. We now turn to the details of these interventions.

The first intervention aimed to improve CHWs' performance through increased social recognition of the CHW in the community. We label it *Social status (award, ceremony, and SMS)*. This intervention was assigned to CHWs at the neighborhood level, i.e., all CHWs in a given neighborhood either received this treatment or not. All agents assigned to this treatment who (individually) performed above an absolute performance threshold were awarded with an honorific prize during a ceremony with the presence of health authorities and community-relevant figures.[22] On top of the award and the ceremony, information on the awarded CHWs was also passed to the corresponding households. All treated CHWs were announced the possibility of awards in initial meetings at the neighborhood level. Figure C1 in the Online Appendix presents the distribution of the neighborhoods in the city of Bissau by treatment status.

In collaboration with the research team, NGO VIDA built for each CHW a score of performance using individual (administrative) information collected by VIDA on a monthly basis. The score was based on three sources of information: (i) the number of monthly reports submitted by each CHW[23]; (ii) test scores from short exams submitted to CHWs every month during the monthly CHW general meeting and refresher training sessions[24]; and (iii) supervisors' evaluations of CHWs' performance.[25] The specific threshold was never made public to CHWs during the three rounds of awards. However, all along, CHWs were informed that the three referred sources of data would be used to decide the awards.[26]

To avoid spillovers to non-treated agents, attendance at the award ceremonies was conditional on receiving an invitation. During the ceremony, the awarded CHWs were called individually and received an honorific award with residual monetary value. The awards were traditional objects, slightly different between rounds, associated with community honor.[27] In addition and in order to increase awareness in the community about the awards, all households assigned to an awarded CHW received a text message to inform them that their CHW

[19] Refer to: 'Global Strategy on Human Resources for Health: Workforce 2030', WHO, 2016; 'WHO Guideline on Health Policy and System Support to Optimize Community Health Worker Programmes', WHO, April 2018.

[20] Table A1 in the Online Appendix shows the 16 essential family practices promoted by the program.

[21] Note however that CHWs receive a small monthly allowance, which is meant to cover transportation and communication costs. It is a function of the number of household visits CHWs perform: this is approximately USD 0.2 per household visited each month. Since CHWs are expected to visit (and communicate to) around 50 households, the typical costs allowance is around USD 10 per month. Every semester, there is also a small compensation for achieving pre-established health goals at the health area level: this is at most USD 21 per semester, i.e., the equivalent of USD 3.5 per month. We employ the average exchange rate for 2017 and 2018: 1 USD = 577.831 XOF (West African CFA francs).

[22] Deserranno et al. (2021) show the importance of ensuring that the evaluation of health workers is perceived as meritocratic. In a field experiment in partnership with the Ministry of Health in Sierra Leone, these authors find that promotions perceived as meritocratic lead to higher productivity. However, promotions that are perceived as non-meritocratic reduce productivity by triggering a negative morale effect.

[23] Each CHW is expected to submit a monthly report with aggregate information on the number of households visited, and a headcount of children and pregnant women tracked and treated. However, all the information is self-reported and is typically not validated externally, which may induce over-reporting of activities by CHWs. The score did not consider the specific information provided in the report.

[24] These meetings were typically held at the level of the health area.

[25] The CHWs were organized in teams, which corresponded to sections of the health areas. Supervisors were assigned to each team, both by VIDA and by the National Health System. The supervisors from VIDA were selected, trained, and employed full time for this activity, with direct responsibility over their teams' performance: they oversaw day-to-day activities, collected data, and filled reports, thus carrying most of the administrative tasks. There were 26 supervisors from VIDA in total. The supervisors from the National Health System were selected among doctors, nurses, and administrative personnel from the health system, with limited time to devote to the supervision of the CHWs.

[26] The specific joint criteria used for assigning the awards were as follows: (i) CHWs had to submit all monthly reports in the period under evaluation; (ii) CHWs had to be given an average score of 15 out of 20 or more in the quizzes submitted during the monthly meetings; (iii) CHWs had to be given an average score of 4.5 out of 5 or higher in the supervisors' reports.

[27] See Section C.1 in the Online Appendix for photos of these objects and of the ceremonies.

had been given a performance award.[28] Note that in each of the three rounds of awards all CHWs in treated neighborhoods had the possibility of winning the award (18 percent of all CHWs treated won at least one round of awards).[29]

The second intervention aimed to improve CHWs' performance through increased intrinsic motivation towards performing their role as health workers. The intervention manipulated CHWs' perceived task significance using an interactive video, a novel format for communicating with CHWs in Guinea-Bissau. The video aimed to make salient the social impact of the CHW task, i.e., the extent to which CHWs' actions improve the welfare of the members of their communities (Grant, 2008). We label this intervention *Task significance (video)*. This intervention was assigned to CHWs at the individual level. The video was recorded from the point of view of a CHW performing daily activities.

There are three versions of the full video, which allowed showing a different version on each round of treatment. Each version covers a different health problem arising on a given day of the CHW activity, directly related to the essential family practices promoted by the program. The three health problems covered are related to: (i) assistance to a pregnant woman, (ii) treatment of diarrhea, and (iii) treatment of severe malaria. The videos were watched individually in tablets using headphones.[30]

The full video has three components, which we describe as follows.[31]

1. **Presentation**: The video begins by showing a CHW visiting a household where he/she encounters a household member facing an health problem.
2. **Interactive decision and ending**: The agent needs to make a single central decision about how to solve the problem raised in the first part of the video. After presenting the health issue, the video stops and offers the agent two different paths: one in which he/she needs to exert/elicit some effort, and another in which he/she leaves the household. Depending on the decision taken by the CHW, the video continues with a positive or a negative ending for the health condition of the referred household member. The negative ending follows a low-effort decision by the CHW and involves the death of that person. After the decision is taken by the CHW watching the video, and the corresponding ending is visualized, the CHW is instructed to play again the video and visualize the other possible ending. The objective of this interactive video is that the CHW clearly sees the potential (dramatic) consequences of his/her actions during household visits.
3. **Endorsement of traditional healers**: A group of eminent traditional healers from outside Bissau appears sequentially on the video, one at a time, making a speech about the importance of CHWs for the welfare of the communities, in practice endorsing their activity. Traditional healers are labeled as such in the video. These figures are very influential in the sphere of tradition and spirituality in Guinea-Bissau.[32]

Importantly, we divide the submission of the video intervention into three cumulative versions, each one constituting a different treatment condition in our experimental design. The first is composed of Part 1 — Presentation only. We label this treatment as *Information/placebo (video)*. This is because this part of the full video just reminds CHWs of specific health problems they can encounter, thus providing some information. At the same time, this part of the video constitutes a strong placebo for the remaining parts. The second version includes both Part 1 and Part 2 — Interactive decision and ending. We label this treatment as *Task significance alone (video)* provided it embeds the simple message of task significance of CHWs. The third version includes all three parts (in addition to the previous two, Part 3 — Endorsement of traditional healers). We label this treatment as *Task significance plus endorsement (video)*. This treatment is intended to be a strong version of task significance, with cultural adherence.

We also designed an additional intervention aiming to improve CHWs' performance through enabling higher levels of cooperation from households. In the context of the recent introduction of CHWs in the city of Bissau, low levels of information in the urban neighborhoods about the role of CHWs could constitute an impediment to their performance. Hence, this intervention disseminated information to households via text messages on the role of the CHWs. We label this intervention *Information campaign (SMS)*. This intervention was assigned to CHWs at the individual level, meaning all households for a given treated CHW were assigned information SMSs.[33]

3.2. Randomization and sampling

Our study includes the full number of CHWs active in the city of Bissau by September 2017, i.e., 1015 individuals. This means our study encompasses the whole of the Autonomous Sector of Bissau.

The randomization procedure for the allocation of treatments to CHWs was implemented following a three-step stratified clustered design. First, within health areas,[34] after the 76 neighborhoods were paired based on population size (number of households), half of them (38) were randomly allocated to the social status intervention.[35] Second, within neighborhood, after pairs of CHWs were formed based on observable characteristics (age, gender, civil status, education, and employment), half of the CHWs were randomly allocated to the information campaign intervention. As mentioned, all the households assigned to those health workers received text messages during the intervention. Finally, within neighborhood and within information campaign treatment status, after quadruplets of CHWs were formed based on observable characteristics (age, gender, civil status, education, and

[28] Section C.1 in the Online Appendix reproduces the specific contents of the text messages that were sent to the households.

[29] Although the Social status intervention followed an absolute threshold of performance minimizing competition between CHWs, awarded only a small minority of the treated CHWs, and conveyed a clear positive distinction sense in the ceremonies and the SMSs, we are unable to rule out a more negative mechanism producing fear of shame if not given the awards.

[30] After watching the video, treated CHWs participated in focus groups to discuss the content of the video and the main messages.

[31] In the Online Appendix to this paper, in Section C.2, we show video snapshots and online links to the videos. Note that the contents of the videos were piloted extensively at design stage in collaboration with NGO Vida, namely with CHWs from other regions in Guinea-Bissau.

[32] Note that the fact that the endorsing traditional healers were from outside Bissau (also labeled as such in the videos), and so unlikely to be socially related to the targeted CHWs, is crucial for minimizing the possibility of direct social status effects of these endorsements. Still, we cannot rule out impacts of the task significance video interventions on perceptions of CHWs about their social status.

[33] Prior to the beginning of the program, NGO VIDA completed a census of the Autonomous Sector of Bissau and collected phone numbers for every household. Each household assigned to the information campaign received three rounds of 2–3 text messages about CHW activities and their role in the community. The information campaign started by presenting the program in the first round. In the second round it provided detailed information about the practices that the CHWs were trained to implement. It also encouraged households to learn more about the program. In the third and last round, the messages included information on the success of some of the activities implemented by the CHW. Section C.3 of the Online Appendix reproduces all the text messages sent to the households.

[34] The city of Bissau is divided in 14 health areas. Each health area is typically defined as the territory covered by a health center.

[35] Neighborhoods are geographically delimited areas within the health areas which are socially accepted. VIDA NGO's supervisors and the official health areas' representatives identified the neighborhoods' borders for the research team.

employment), CHWs were randomly assigned to one of four groups in relation to the video interventions: (i) the information/placebo group, only exposed to the first component of the full video, i.e., the presentation; (ii) the task significance alone group, exposed to the presentation and the interactive component of the video; (iii) the task significance plus endorsement group, exposed to the presentation, the interactive video, and the endorsements by the traditional healers; (iv) a control group not exposed to any video intervention.

This crossed randomization procedure produced 15 treatment groups and one pure control group. These comparison groups are shown in the CONSORT diagram of Figure D1 in the Online Appendix. As expected, the numbers of CHWs are similar across these 16 groups.

As part of the measurement in this project we sampled households for surveying face-to-face and by phone. This was done by randomly selecting a fixed number of households from the list of households of each CHW. In the face-to-face survey, two households were sampled from each CHW for the baseline and endline surveys. In the phone survey, four households per CHW were sampled from half the CHWs, whom were randomly selected. Note that for the phone survey, the sampling process was conditional on the existence of phone numbers for the corresponding households and happened after the interventions finished.

3.3. Measurement

Our measurement in this project includes a broad range of data sources. These encompass: (i) administrative data from NGO VIDA, the local counterpart managing the CHWs; (ii) baseline and endline CHW surveys; (iii) baseline and endline household face-to-face surveys; (iv) a household phone survey administered after the end of the interventions; and (v) daily health-provision activities from hospitals' and health centers' logbooks from October 2017 to October 2018. We now turn to providing some details about the design of these data.

The administrative data from our implementing partner include CHW retention rates until three months after the end of the interventions, i.e., until February 2019. Apart from these data, we also had access to self-reported monthly reports of CHWs' home visits, test scores before and after the training sessions, and evaluation records of CHWs by their supervisors. We employ as auxiliary data the administrative records on supervisors' basic demographic characteristics.

The survey data we designed and conducted includes face-to-face surveys to all CHWs and to a random sample of households before the start of the intervention (July–September 2017) and 14 months later (October–November 2018). The survey questionnaire targeting CHWs includes questions on their demographic and socioeconomic characteristics. It also includes a module on psychometric questions related to motivation, on social connections to other agents in the program, and on participation in community activities. The face-to-face survey questionnaire targeting households includes questions on demographic and socioeconomic characteristics for all household members. Importantly for our analysis, we asked questions to the household head on health and sanitation practices in the household, as well as on health outcomes for all children living in the household who were 5 years of age or younger. We gathered information on fertility for all women between the ages of 12 and 49, and we asked questions on peri-natal care to all women with children born alive in the previous two years. In the endline questionnaire we include questions on knowledge of the 16 essential family practices and on the household's experience with the CHW program.

The household phone survey we designed and conducted was administered after the end of the treatments in November 2018. It included simple demographic questions as well as a small number of questions about the household's experience with the CHW program. Submitting the phone survey lasted on average 10 min.

The research team visited all 10 health centers and the three hospitals in Bissau and digitized logbooks with registries on vaccination, post-natal care, and family planning from October 2017 to November 2018. Logbooks are homogeneous across facilities. Since patient identifiers in the logbooks were imperfectly registered, we opted for merging these data at the level of the place of residence/neighborhood. As we collected data from all the health centers and the main hospitals in Bissau, we manage to match information on patients' visits to neighborhoods even in those cases in which households visited a health center/hospital to which they were not geographically assigned. We use these data to evaluate the impact of the Social status treatment.

All outcome questions employed in our study, structured by data source, are fully described in Section E of the Online Appendix to this paper.

4. Hypotheses and estimation strategy

Our experiment is designed to study the impact of two distinct types of non-financial incentives, one on social status, the other on intrinsic motivation via task significance. We are mainly interested in assessing impacts on CHWs' performance and on households' health outcomes. Hence, our main hypotheses are the following.

Hypothesis 1. The incentive treatment on Social status (award, ceremony, and SMS) improves the performance of CHWs as well as the health outcomes at the level of the households. It is likely that an increase in effort by the CHWs translates into better practices among the visited households.

Hypothesis 2. The incentive treatments on the Task significance video (blending the groups with and without endorsement by traditional healers) improves the performance of CHWs as well as the health outcomes at the level of the households. It is likely that an increase in effort by the CHWs translates into better practices among the visited households.

This is the backbone of the Pre-analysis Plan (PAP) that we registered for this project at the American Economic Association (AEARCTR-0003399) under the title 'Non-Financial Incentives of Community Health Workers in Guinea-Bissau'. The PAP makes clear that the main objective of the project is to test the impact of non-financial incentives on the performance of CHWs as well as the health outcomes of the corresponding households, and that there are only two dimensions of incentives studied in the project. Still, we acknowledge the large set of hypotheses in our design given the different video treatments relating to task significance, the information treatment, and the tests of complementarity between treatments enabled by the cross-randomization we implemented.

For these reasons, we take the latter as secondary hypotheses in this paper. The corresponding results are referred in the paper but are only shown in the Online Appendix. These hypotheses are the following: first, that Task significance improves on Information/placebo, i.e., that the visualization of CHW impact in community health (through the video) produces in itself improvements in CHWs' performance and households' health; second, that there is a positive difference in the outcomes when comparing Task significance plus endorsement with Task significance alone, implying that endorsements by traditional figures are impactful; third, that the Information campaign improves the outcomes; fourth, that the two main incentive treatments, and that each of them and the information treatment are complementary regarding improvements in the performance of CHWs as well as in the health outcomes of the households.

To evaluate the main hypotheses we estimate a set of specifications, where the treatments are labeled SS for Social Status, TS for Task Significance both with and without endorsements by traditional healers, IP for Information/Placebo video, and IC for Information Campaign.

The first specification we consider is the following.

$$y_i = \alpha + \beta_1 \mathrm{SS}_i + \beta_2 \mathrm{TS}_i + \beta_3 \mathrm{IP}_i + \beta_4 \mathrm{IC}_i + X_i'\gamma + \epsilon_i \quad (1)$$

where y_i is the outcome of interest at the endline, i.e., related to CHW performance or household health (assumed to be measured in such a way that higher values signify better outcomes). Note that individual i can be a CHW, a household head, a child under 5 years old belonging to a household, a woman with a child born alive in the past two years belonging to a household, a woman in fertile age (12–49 years old) belonging to a household, or a phone-survey respondent belonging to a household. Treatment indicators are binary variables taking value 1 for CHWs or households whose CHWs were assigned the corresponding treatment. X_i is a set of controls including strata fixed effects, defined from the blocks formed before randomization.[36] ϵ_i is an idiosyncratic error term. To account for possible correlation in outcomes within neighborhoods, the error term is clustered at the neighborhood level.

When baseline values of the outcome variable are available, we can employ an ANCOVA specification which can be described as follows:

$$y_i = \alpha + \beta_1 \mathrm{SS}_i + \beta_2 \mathrm{TS}_i + \beta_3 \mathrm{IP}_i + \beta_4 \mathrm{IC}_i + X_i'\gamma + \delta y_{i0} + \epsilon_i \quad (2)$$

where y_{i0} is the baseline value of the dependent variable.

Specifications (1) and (2) allow testing Hypotheses 1 ($\beta_1 > 0$) and 2 ($\beta_2 > 0$).

Finally, we modify Eq. (1) to analyze the treatment effects of Social status on health-related outcomes from hospitals' and health centers' logbooks. We employ the following specification:

$$y_{ig} = \alpha + \beta_1 \mathrm{SS}_g + X_{gi}'\gamma + \epsilon_g \quad (3)$$

where y_{gi} is the outcome of interest for patient i in neighborhood g. X_{gi} is a set of controls including strata fixed effects and a vector of neighborhood-specific characteristics.[37] This specification allows revisiting Hypothesis 1 ($\beta_1 > 0$).

For all specifications, we estimate linear regressions regardless of whether the outcomes are continuous or discrete.[38] In the results section we check whether the main results of the paper are robust to using the Post-Double Selection LASSO procedure to select control variables.

The outcome variables studied in this paper are included in the PAP. Despite this, the fact that we analyze in this paper a large number of outcome variables raises concerns about multiple-hypothesis testing: as the number of single hypotheses under consideration increases, the probability that at least one of them is falsely rejected given that all of them are true, i.e., the family-wise error rate, increases as well. In order to reduce this concern, we apply two strategies. First, we follow Kling et al. (2007) and aggregate similar individual outcomes into indices. This is done by calculating within-sample z-scores for each outcome variable, using the mean and the standard deviation of the pure control group, and applying non-weighted averages of z-scores between outcomes. Second, while employing the algorithm described in Romano and Wolf (2016), we also compute, for each null hypothesis under study, a corresponding p-value adjusted for the stepwise multiple hypothesis testing method proposed in Romano and Wolf (2005a,b). This method is stepdown like other improvements over Bonferroni (Holm, 1979), and resampling-based, which allows accounting for dependence between hypotheses. Hence, the underlying procedure allows increasing the power of the testing over other previous methods. The main test, reported in each table, considers each treatment effect separately: we test whether each one of the treatment effect parameters is significantly different from zero across all outcomes reported in the table grouped by the source of data (this is the same grouping we implement in the averages of z-scores). Following a more conservative strategy, we test Hypotheses 1 and 2 simultaneously on top of the previous approach, i.e., across outcomes grouped by the source of the data presented in each table. Results are reported in Section J of the Online Appendix.

5. Results

5.1. Balance and descriptive statistics

We show balance tests in Section F of the Online Appendix. Our randomization procedure was able to identify comparable groups, namely in terms of demographic characteristics of both CHWs and households. Appendix Table F1 presents balance tests on baseline characteristics for the full set of CHWs across the three main treatment arm dimensions, i.e., social status, video treatments, and information campaign, when compared to the corresponding control groups. We also employ a joint F-test to test for all main differences together and report p-values for this test. We run 90 tests and find statistical significance in only four cases.

We can also employ Table F1 to provide an overall description of CHWs' characteristics by looking at the mean of the pure control group. The average age of these CHWs is 26 years, 48 percent are female, and 49 percent are Catholic. Seventy-six percent have completed 12 years of schooling, and 51 percent were studying at the time the CHW program started.[39] Fifty-one percent worked in the 12 months previous to the beginning of the CHW program and 11 percent had a business when the program started. Sixty-two percent of the CHWs had done volunteer work at a health center before the beginning of the program and 81 percent had had a position in the community.

Appendix Table F2 reports balance tests for the sample of households interviewed using the face-to-face survey. We run 96 tests and find statistically significant ones for eight cases.

Table F2 provides an overall characterization of the demographic characteristics of the sample of households interviewed face-to-face. The average age of the household head is 44.5 years, 33 percent are female, and 41 percent are Catholic. Twenty-four percent have completed 12 years of schooling, and 69 percent worked in the 12 months previous to the beginning of the CHW program. Twenty percent are Balanta, 14 percent Papel, and 19 percent Fula. Households are composed on average of 7 members, among whom 2 are women in fertile age, and 1 is a child under 5 years old. The average number of assets is 4.29 (over a maximum of 15) and the average number of

[36] Control variables include CHW characteristics (gender, age, and education) and households characteristics when interviewed face-to-face (age and gender of the household head as well as household size). When analyzing health of children under 5 years old, women in fertile age, or women with a child born alive in the past two years, controls include the age of the corresponding individual.

[37] Control variables include averages of CHWs characteristics at the neighborhood level (gender, age, and education), number of households targeted by CHWs in the neighborhood, and the average of households' size at the neighborhood level. In addition, when analyzing vaccine records, controls include children's age fixed effects and number of children under 2 years old living in the neighborhood. When analyzing post-natal care records, controls include quarter fixed effects for date of visit and the number of births in the neighborhood. When analyzing family planning records, controls include quarter fixed effects for date of visit, and the number of women in fertile age living in the neighborhood.

[38] The specifications we describe were registered in the PAP. There are however two caveats. First, we are aggregating together the task significance video treatments with and without endorsements by traditional healers in the main analysis of the paper. In the Online Appendix, we distinguish all the video treatments. Second, we also included a difference-in-differences specification in the PAP, which we decided to drop in favor of the ANCOVA specification. If the autocorrelation of the outcome variable is low, which is the case for most survey outcomes, the latter specification maximizes statistical power in field experiments (McKenzie, 2012).

[39] This is above regional standards of education of CHWs: data from the One Million Community Health Workers Campaign, 2019, shows that 40.7 percent of the CHWs registered in Sub-Saharan Africa have completed secondary education. It is important to note that this figure includes rural areas, which is likely to explain the average difference to the CHWs that we study in the capital city of Bissau.

mosquito nets per household member is 0.48. Eighty-one percent have access to piped water and only 3 percent use latrines.[40]

An analysis of attrition is given in Section G in the Online Appendix. Table G1 shows data availability rates for the pure control group as well as differences across comparison groups. We look at both CHW and household-level data. We have complete administrative records for 90 percent of the CHWs, and 86 percent of the CHWs were interviewed in the endline survey (rates for the pure control group). The division of CHW attrition between comparison groups is in the CONSORT diagram in Figure D1 of the Online Appendix. Turning to households, we have at least one household interviewed by phone at the endline for 48 percent of the CHWs in the pure control group. Attrition in the face-to-face household survey was 12 percent (for the pure control). We test for differences across treatment arms and for all differences together. Attrition rates for CHWs and households are not significantly different across treatment arms. The exceptions, which yield marginal significance, are that CHWs assigned to task significance plus endorsement (video) are 4 percentage points less likely to have complete administrative records and 7 percentage points more likely to have at least one household interviewed by phone (compared to the video control group).

Tables G2–G3 in the Online Appendix verify that the CHWs and households surveyed face-to-face at endline, i.e., after attrition, are similar in treatment and control groups.

A final note goes to the logbooks of health centers and hospitals. The final matching rates on the basis of place of residence are 34 percent for the logbook on vaccines, 37 percent for the post-natal care logbook, and 24 percent for the family planning logbook. A potential concern is whether the number of records matched differs by comparison groups. To analyze whether this is the case, we use information on CHWs, the number of households living in each neighborhood, and the composition of these households. With the caveat of more limited statistical power, Table G4 in the Online Appendix shows that the differences between the Social status treatment group and its control are not statistically significant for any of the population variables analyzed.

5.2. Treatment effects of incentives — aggregated outcomes

We begin by showing treatment effects on aggregated measures of our outcomes of interest. This is in order to address the risks posed by the analysis of multiple outcomes. We bundle outcomes in indices that are built using the procedure detailed in Kling et al. (2007). We calculate within-sample z-scores for each individual outcome, employing the mean and the standard deviation of the pure control group. Subsequently, we obtain the unweighted average z-score for each set. We define indices as a function of level of analysis and source. We employ specification (1) in the case of outcomes that can be matched with CHWs, but focus attention on the effects of the incentive treatments, i.e., of Social status (award, ceremony, and SMS) and of Task significance (video), which blends the task significance treatments with and without endorsements by traditional healers. We also use specification (3) in the case of outcomes measured through the logbooks of health centers and hospitals. This analysis enables reporting about the validity of Hypotheses 1 and 2.

Although we follow in this section a specific aggregation procedure, it puts together similar or related outcome variables. This is not a substitute to showing the full detailed results per outcome variable, which we discuss in the next section of the paper. Specifically, we consider indices aggregating the dependent variables corresponding to sections of Tables 1–5, which depict our disaggregated results.

Our index on CHW performance employs administrative data at the level of the CHW, built from outcomes in columns (1)–(4) of Table 1, which include CHW dropout in February 2019, three months after the end of the interventions, the share of monthly reports submitted by CHWs during the time they were active, the test score of CHWs in examinations taken during monthly meetings, and the evaluation score of CHWs attributed by their supervisors.

We consider two indices of home visits, one using the phone survey at the level of the household, built from outcomes in columns (5)–(6) of Table 1, and the other utilizing the face-to-face survey at the level of the household, built from outcomes in columns (7)–(8) of Table 1. The specific outcome variables we observe are the total number of CHW visits reported by the households, as well as household satisfaction with the CHWs.

Our index of health indicators at the level of the household, employs the face-to-face survey data, built from the outcomes of Table 2, which include knowledge of health practices by the survey respondent, whether the household treats water with bleach or chlorine, the number of mosquito nets impregnated with insecticide in the household, and whether the household uses latrines.

We compose two indices relating to the health of children under 5 years old. The first index uses the face-to-face survey data at the level of the child in the household, built from the outcomes in Table 3. These encompass the extent to which children were vaccinated by employing an index of taking the five most important vaccines (BCG, polio, diphtheria–tetanus–pertussis, measles–mumps–rubella, and yellow fever), both measured from self-reports and the observation of vaccination bulletins for individual children. We also add outcomes on whether children are reported to have been sick in the last 15 days before the survey, as well as on whether they took a malaria test at all and conditionally on having malaria symptoms. The second index is dedicated to vaccination of children, but this time using health records from the health centers' and hospitals' logbooks on take-up of the five most important vaccines, as defined above and employed in Table 5.

Finally, we devote our attention to family planning and peri-natal care. First, we employ the face-to-face survey data at the level of the woman, built from the outcomes in Table 4. These include whether women used family planning methods in the 12 months before the survey, the number of pre-natal visits to a health facility during pregnancy, an index of quality of pre-natal care averaging indicator variables for taking pre-natal care exams (blood pressure, blood, and urine tests) and taking a vaccine (tetanus), whether women attended a post-natal visit to a health center after giving birth, whether women nursed their children after birth, and on whether women administered vitamin A to their children in the 45 days after birth. Second, we employ a measure of family planning, using health records from the logbooks, also employed in Table 5, which concern first visits to health centers/hospitals for family planning appointments. Third, we include a measure of whether the births happened at home, using health records from the logbooks, also used in Table 5. Fourth, we look at post-natal care, employing records from the logbooks, built from outcomes in columns (4)–(5) of Table 5, which include measures on the time elapsed between delivery and post-natal check-up, and on whether a post-natal check-up happened in the 10 days after giving birth.

Fig. 1 shows treatment effects of Social status on the aggregate indices we describe above.[41] Confidence intervals are built using statistical significance at the 5 and 10 percent levels. Given the standardization of outcome variables embedded in the procedure we adopted, all treatment effects are in standard deviation units.

We find significant effects for the Social status treatment in the case of home visits (phone survey), health indicators at the level of the household (face-to-face survey), health of children under 5 years

[40] When employing the census data for Bissau (2009), we find a similar overall picture. The average age of the household head is 43 years, 31 percent are female, and 44 percent are Catholic. 20 percent have completed 12 years of schooling. 20 percent are Balanta, 16 percent Papel and 19 percent Fula. The average number of household members is 7.

[41] See Table H4 in the Online Appendix for the full table corresponding to this figure.

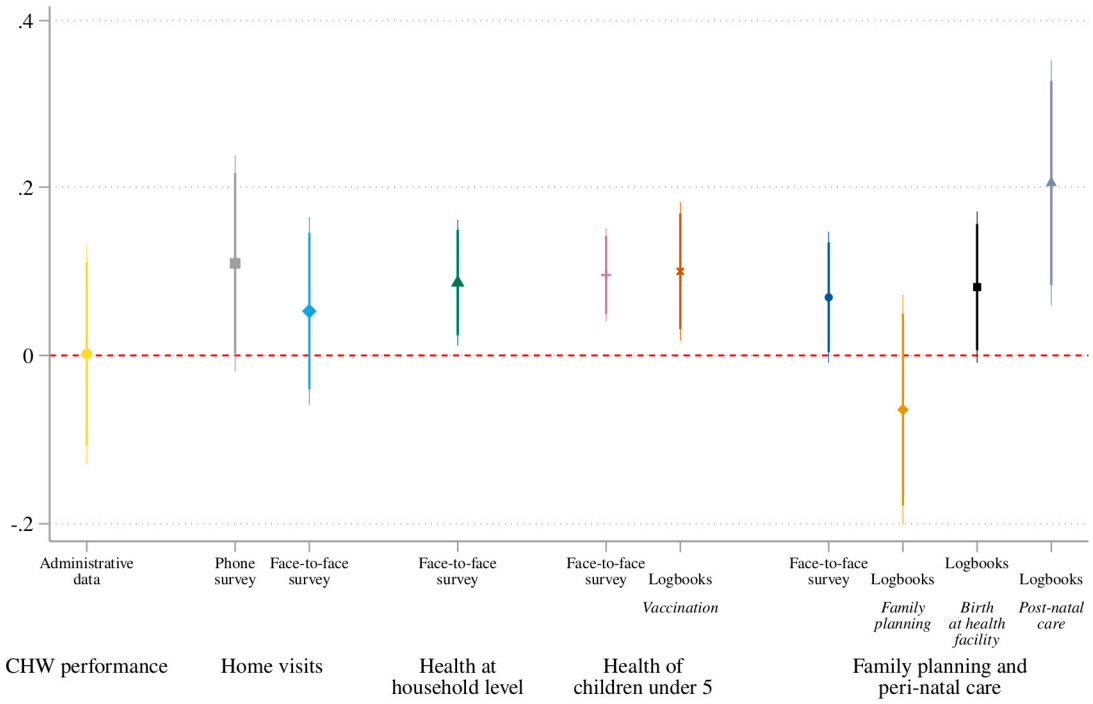

Fig. 1. Main treatment effects – aggregated outcomes employing z-scores – Social status.
Note: Estimates based on OLS regressions. The first to fifth, and seventh bars present estimates using Eq. (1). The sixth, eighth, ninth and tenth bars present estimates using Eq. (3). Outcomes are grouped in indices that are built using the procedure in Kling et al. (2007). We calculate within-sample z-scores for each individual outcome, employing the mean and the standard deviation of the pure control group. We then obtain the unweighted average z-score for each category. The indices are defined by the following outcomes: (1) CHW performance (administrative data): formed from outcomes in columns (1)–(4) of Table 1; (2) Home visits (phone survey): formed from outcomes in columns (5)–(6) of Table 1; (3) Home visits (face-to-face survey): formed from outcomes in columns (7)–(8) of Table 1; (4) Health at household level (face-to-face survey): formed from outcomes in columns (1), (2), (4), and (6) of Table 2; (5) Health of children under 5 (face-to-face survey): formed from outcomes in columns (1), (3), and (5)–(7) of Table 3; (6) Vaccination of children (logbooks): index measuring the take-up of the five most important vaccines as defined in Table 5; (7) Family planning and peri-natal care (face-to-face): formed from outcomes in columns (1), and (3)–(7) of Table 4; (8) Family planning (logbooks): indicator variable that takes value of 1 for women visiting for the first time a health center or hospital for an appointment on family planning, as defined in Table 5; (9) Birth at a health facility (logbooks): indicator variable that takes value of 1 for women who gave birth at a health facility; (10) Post-natal care (logbooks): formed from outcomes in columns (4)–(5) of Table 5. Specifications in the first to the fifth bar and specification in the seventh bar include an indicator variable for assignment to the Task significance treatment, an indicator variable for assignment to the Information/Placebo treatment, an indicator variable for assignment to the Information campaign treatment, strata fixed effects, and CHW level controls. Specifications in the third to the fifth bar and specification in the seventh bar include household level controls. The specification in the fifth bar includes age fixed effects for the children under 5 years old. The specification in the seventh bar includes respondent's age. Specifications in the sixth, eighth, ninth, and tenth bars include neighborhood characteristics. The full list of controls is presented in Section 4. Confidence intervals are built using statistical significance at the 10 and 5 percent level. Standard errors are clustered at neighborhood level.

old (face-to-face survey), vaccination of children (logbooks), peri-natal care (face-to-face survey), birth at a health facility (logbooks), and post-natal care (logbooks). The magnitudes of these effects vary between 0.07 and 0.2 standard deviation units. These findings are clearly in line with Hypothesis 1, following the whole chain of expected effects: from performance of CHWs, including their effort, to better practices among the visited households and their health outcomes.

Fig. 2 shows treatment effects of Task significance on the indices built using administrative data, as well as phone and face-to-face survey data.[42] Treatment effects of Task significance are much less clear than effects of Social status. However, driven by the outcome variables on vaccinations of children, we find a treatment effect of 0.09 standard deviations, (significant at the 5 percent levels) for Task significance when considering the index on health of children under 5 years old.[43]

We conclude, when employing indices of outcome variables, that the Social status intervention was generally effective, according to our corresponding hypothesis. We do not find the evidence on the impact of the Task significance intervention to be as systematic.

5.3. Treatment effects of incentives — disaggregated outcomes

5.3.1. CHW performance

We now turn to the analysis per individual outcome variable, and start by devoting attention to our measures of direct CHW performance. We employ specification (1) and focus attention on the effects of the incentive treatments, i.e., of Social status (award, ceremony, and SMS) and of Task significance (video). This analysis enables reporting about the validity of Hypotheses 1 and 2. Our treatment effects are shown in Table 1. Columns (1) to (4) are dedicated to administrative data for CHWs, i.e., at the level of the CHW. Columns (5) to (8) are devoted to household survey data from the endline phone survey, and from the endline face-to-face survey, i.e., at the level of the household.

We find positive treatment effects of Social status on several outcome variables related to CHW direct performance, consistently with Hypothesis 1. This is the case for performance scores of CHWs from administrative data, and household satisfaction with the CHWs. We observe that test scores improve by 0.09 standard deviation units and that supervisory scores improve by 0.07 standard deviation units.[44] These

[42] See Table H5 in the Online Appendix for the full table corresponding to this figure.

[43] Note that the effect sizes we report from Figs. 1 and 2 are relatively low compared to the minimum detectable effects we would have at the design stage (reported in Table D1 in Appendix). This is mainly due to the unavailability of data on control variables at that stage, allowing for the full consideration of the statistical power we have in our experimental design.

[44] We show in Figures H1 and H2 in the Online Appendix that these effects are slightly increasing during 2018 in the case of text scores but peak in the

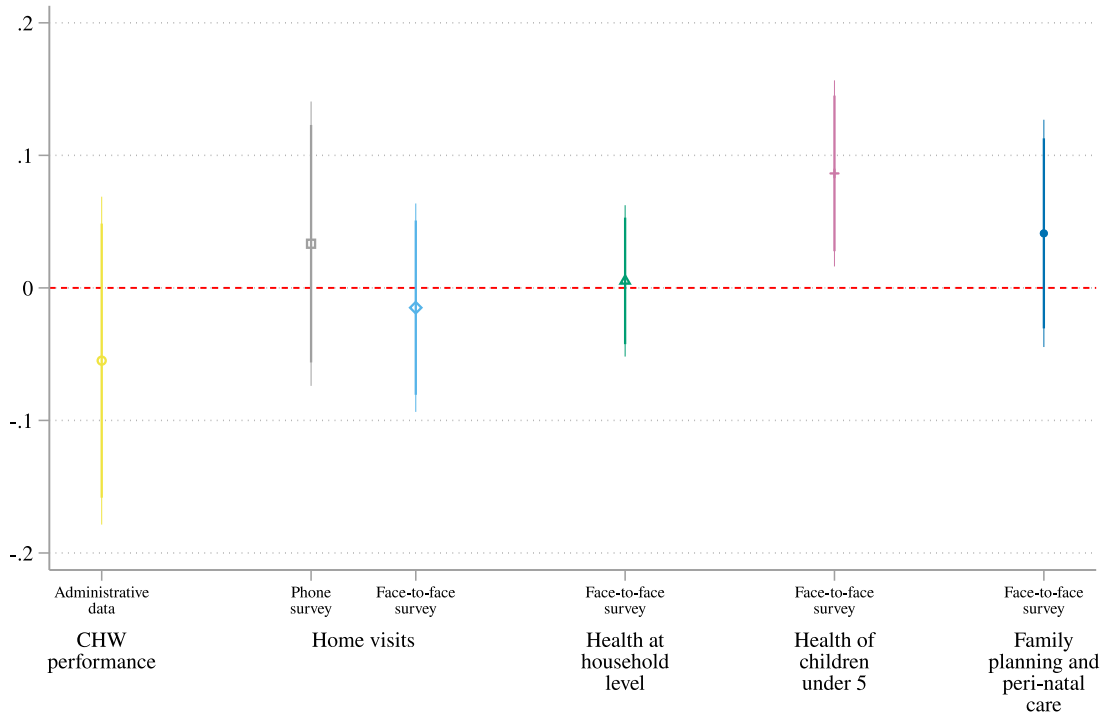

Fig. 2. Main treatment effects – aggregated outcomes employing z-scores – Task significance.
Note: Estimates based on OLS regressions using Eq. (1). Outcomes are grouped in indices that are built using the procedure in Kling et al. (2007). We calculate within-sample z-scores for each individual outcome, employing the mean and the standard deviation of the pure control group. We then obtain the unweighted average z-score for each category. The indices are defined by the following outcomes: (1) CHW performance (administrative data): formed from outcomes in columns (1)–(4) of Table 1; (2) Home visits (phone survey): formed from outcomes in columns (5)–(6) of Table 1; (3) Home visits (face-to-face survey): formed from outcomes in columns (7)–(8) of Table 1; (4) Health at household level (face-to-face survey): formed from outcomes in columns (1), (2), (4), and (6) of Table 2; (5) Health of children under 5 (face-to-face survey): formed from outcomes in columns (1), (3), and (5)–(7) of Table 3; (6) Family planning and peri-natal care (face-to-face): formed from outcomes in columns (1), and (3)–(7) of Table 4. Specifications employed include an indicator variable for assignment to the Social status treatment, an indicator variable for assignment to the Information/Placebo treatment, an indicator variable for assignment to the Information campaign treatment, strata fixed effects, and CHW level controls. Specifications in the third to the sixth bar include household level controls. The specification in the fifth bar includes age fixed effects for the children under 5 years old. The specification in the sixth bar includes respondent's age. The full list of controls is presented in Section 4. Confidence intervals are built using statistical significance at the 10 and 5 percent level. Standard errors are clustered at neighborhood level.

effects are statistically significant at the 10 percent level, although not robust to the Romano–Wolf multiple hypothesis correction. Household satisfaction increases by 0.20 standard deviation units (phone survey) and by 0.24 standard deviation units (face-to-face survey). These effects are statistically significant at the 10 percent level, and robust to the Romano–Wolf procedure. Other treatment effects of Social status are generally positive, although not significant at standard levels. We also find one significant and positive effect of the Task significance video, namely for household satisfaction with the CHWs in the phone survey data, consistently with Hypothesis 2. While this effect is robust to the Romano–Wolf procedure, it is not robust when considering the face-to-face survey data. The effects of Social status are significantly different from those of Task significance when considering the test scores and household satisfaction with CHWs (face-to-face survey).

A potential concern with performance-based incentives is the presence of multitasking problems, i.e., effort allocated toward targeted indicators may come at the expense of other, non-incentivized indicators (Holmstrom and Milgrom, 1991). On treatment effects related to the direct performance of CHWs, we conclude that Social status incentives are effective in improving incentivized outcomes, i.e., test scores and evaluation scores, as well as a non-incentivized outcome, i.e., households' satisfaction with the CHWs.[45] Among the three targeted indicators in assessing CHW performance for the honorific awards, the Social status intervention had the largest impact on test scores in the context of refresher training, suggesting effects on learning and improved skills.[46] Results relating to the Task significance video are not as clear when comparing to those of the Social status treatment. This means that only Hypothesis 1 seems systematically verified.

middle of the year in the case of supervisory scores. This is not suggestive of a treatment effect of Social status that arises solely in the short run.

[45] One could argue that along with a motivation related to the social status of the CHW in strict sense, this intervention could trigger effects related to being monitored, or to a different perception of the health authority. Indeed, household satisfaction with CHWs improves as shown in Table 1. We estimate treatment effects on self-reported satisfaction by the CHWs about different aspects of the CHW program and on their perception of the satisfaction of households about the CHW program. These are shown in Table H1 of the Online Appendix. We do not find significant effects of the Social status treatment on CHW satisfaction about the program or their perceptions about the satisfaction of households. Possibly related to our design of absolute thresholds of performance for the honorific awards, we do not find effects on the CHWs' evaluation of their peers.

[46] We show in the Online Appendix treatment effects on specific components of both the test scores in the context of monthly meetings and the evaluation of supervisors. Table H2 shows treatment effects on test scores by topic. We find a positive effect of Social status on knowledge of the protocol for home visits, which is in line with the findings on the satisfaction of households with CHW performance during home visits. We also report an increase in scores measuring knowledge about prevention of illnesses and their identification. These results open the possibility that the Social status intervention impacts health outcomes at the household level. Table H3 shows treatment effects on supervisors' evaluation score of CHWs by component (theoretical knowledge, relationship with families, protocol for home visits, transfer of know-how, as well as management and monitoring protocol). We find that the positive Social status treatment effect observed on supervisors' evaluation score of CHWs is driven by an increase of 0.9 standard deviations on theoretical knowledge.

Table 1
Direct CHW performance.

	Administrative data				Phone survey		Face-to-face survey	
	CHW dropout	CHW reports submitted — share	CHW test score — training	CHW evaluation score by supervisor	Total home visits	Household satisfaction with the CHWs	Total home visits	Household satisfaction with the CHWs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Social status	0.01	−0.03	0.09*	0.07*	0.37	0.20*	0.17	0.24*
	(0.03)	(0.03)	(0.05)	(0.04)	(0.24)	(0.11)	(0.27)	(0.12)
	[0.73]	[0.34]	[0.18]	[0.18]	[0.05]	[0.01]	[0.33]	[0.01]
Task significance	0.04	−0.02	−0.08	0.06	0.01	0.17**	0.12	−0.10
	(0.03)	(0.03)	(0.09)	(0.07)	(0.28)	(0.08)	(0.21)	(0.10)
	[0.24]	[0.49]	[0.49]	[0.49]	[0.95]	[0.01]	[0.46]	[0.31]
Observations	1015	1015	939	936	1797	1241	1645	685
R^2	0.10	0.15	0.09	0.10	0.11	0.11	0.09	0.18
Mean (control group)	0.19	0.62	0.02	−0.10	3.30	0.00	1.72	0.00
P-values:								
Joint test all treatments	0.515	0.630	0.338	0.344	0.494	0.022	0.464	0.377
Social status = Task significance	0.450	0.891	0.084	0.838	0.360	0.811	0.875	0.071
ANCOVA specification	No	No	No	No	No	No	No	No

Note: Estimates based on OLS regressions using Eq. (1). Estimation sample in columns (1)–(4) consists of CHWs for whom we have administrative records; estimation sample in columns (5)–(6) consists of households interviewed in the endline phone survey; estimation sample in columns (7)–(8) consists of households interviewed in the endline face-to-face survey. Depending on the column the dependent variables are defined by the following. (1) CHW dropout: indicator variable equal to 1 if the CHW dropped out the program by February 2019. (2) CHW reports submitted — share: number of monthly reports submitted by the CHW divided by the number of months that the CHW was active from October 2017 to November 2018. (3) CHW test score — training: average score in the monthly meetings' examinations from May to October 2018; score ranges from 0 to 20 and is normalized (z-score) within supervisor. (4) CHW evaluation score by supervisor: average score from supervisors' monthly report on CHWs performance, from January to November 2018; score ranges from 1 to 5 and is normalized (z-score) within supervisor. (5) and (7) Home visits — total: total number of CHW home visits received since the start of the program. (6) and (8) Household satisfaction with the CHWs: respondent's level of satisfaction with the activity of the CHW conditional on being visited at least once by a CHW; this variable is normalized (z-score) relative to the pure control group. All specifications include an indicator variable for assignment to the Information/placebo treatment, an indicator variable for assignment to the Information campaign treatment, strata fixed effects, and CHW level controls. Specifications (7)–(8) include household level controls. The full list of controls is presented in Section 4. Standard errors are reported in parentheses. Standard errors are clustered at neighborhood level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. *P*-values adjusted for multiple hypothesis testing grouped by rows and source of data, following the structure of the aggregate indices in Figure 1, are presented in squared brackets.

Table 2
Health indicators at the level of the household.

	Face-to-face survey						
	Knowledge of health practices	Household treats water		Number of mosquito nets		Use of latrines	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Social status	0.28**	−0.03*	−0.04**	0.23	0.22	0.01	0.00
	(0.11)	(0.02)	(0.02)	(0.15)	(0.15)	(0.04)	(0.04)
	[0.01]	[0.08]	[0.08]	[0.08]	[0.08]	[0.80]	[0.80]
Task significance	0.01	0.00	−0.00	−0.05	−0.04	0.02	0.02
	(0.06)	(0.02)	(0.02)	(0.11)	(0.10)	(0.03)	(0.03)
	[0.99]	[0.99]	[0.99]	[0.93]	[0.93]	[0.93]	[0.93]
Observations	1744	1752	1749	1752	1747	1752	1749
R^2	0.24	0.07	0.10	0.41	0.43	0.14	0.14
Mean (control group)	0.00	0.29	0.29	2.98	2.98	0.57	0.57
P-values:							
Joint test all treatments	0.019	0.434	0.193	0.473	0.443	0.615	0.604
Social status = Task significance	0.050	0.296	0.218	0.193	0.206	0.831	0.775
ANCOVA specification	No	No	Yes	No	Yes	No	Yes

Note: Estimates based on OLS regressions. Columns (1), (2), (4), and (6) present estimates using Eq. (1), columns (3), (5), and (7) present estimates using Eq. (2), which includes the lagged dependent variable (ANCOVA). Estimation sample consists of households interviewed in the endline face-to-face survey. Depending on the column the dependent variables are defined by the following. (1) Knowledge of health practices: number of correct answers to 28 questions measuring household knowledge of the 16 basic essential family practices and of family planning; this variable ranges from 0 to 28 and is normalized (z-score). (2)–(3) Household treats water: indicator variable that takes value of 1 for households who report treating water with bleach or chlorine. (4)–(5) Number of mosquito nets: total number of mosquito nets impregnated with insecticide available in the house. (6)–(7) Use of latrines: indicator variable that takes value of 1 for households who report using latrines. All specifications include an indicator variable for assignment to the Information/placebo treatment, an indicator variable for assignment to the Information campaign treatment, strata fixed effects, as well as CHW and household level controls. The full list of controls is presented in Section 4. Standard errors are reported in parentheses. Standard errors are clustered at neighborhood level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. *P*-values adjusted for multiple hypothesis testing grouped by rows (at the row-level) are presented in squared brackets.

5.3.2. *Household health — survey*

We now turn to measures of household health from the endline face-to-face household survey we conducted. Here, we seek to identify the treatment effects of both types of incentives in our design to inform on Hypotheses 1 and 2 (like in the previous section). Table 2 depicts results on outcome variables related to the whole surveyed household, i.e., at

Table 3
Health of children under 5 years old.

	Face-to-face survey						
	Vaccination index (5 vaccines)				Sick in the last 15 days	Took a malaria test	
	Self-reported		Observed bulletin			All	If sick
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Social status	0.01	0.02	0.02	0.04*	−0.08**	−0.03*	0.07
	(0.01)	(0.01)	(0.02)	(0.02)	(0.04)	(0.02)	(0.05)
	[0.27]	[0.27]	[0.27]	[0.27]	[0.10]	[0.19]	[0.27]
Task significance	0.05***	0.05***	0.09***	0.07***	0.00	−0.01	−0.00
	(0.02)	(0.01)	(0.03)	(0.03)	(0.03)	(0.03)	(0.07)
	[0.00]	[0.00]	[0.00]	[0.00]	[1.00]	[0.99]	[1.00]
Observations	1295	1057	1018	656	1295	1295	386
R^2	0.38	0.25	0.32	0.30	0.13	0.06	0.18
Mean (control group)	0.87	0.87	0.75	0.75	0.34	0.14	0.41
P-values:							
Joint test all treatments	0.022	0.002	0.003	0.012	0.198	0.491	0.463
Social status = Task significance	0.056	0.103	0.091	0.377	0.101	0.413	0.367
ANCOVA specification	No	Yes	No	Yes	No	No	No

Note: Estimates based on OLS regressions. Columns (1), (3), (5), (6) and (7) present estimates using Eq. (1), columns (2) and (4) present estimates using Eq. (2), which includes the lagged dependent variable (ANCOVA). Estimation sample in columns (1), (2), (5), (6) and (7) consists of children under 5 years old living in households interviewed in the endline face-to-face survey; estimation sample in columns (3) and (4) is restricted to children under 5 years old with a vaccination bulletin. Depending on the column the dependent variables are defined by the following. (1)–(4) Vaccination index (5 vaccines): index variable averaging five indicator variables for taking each of the following vaccines: BCG, polio, DTcoq (diphtheria–tetanus–pertussis), MMR (measles–mumps–rubella), and yellow fever; columns (1) and (2) employ self-reported data, columns (3) and (4) employ information directly observed in the vaccination bulletins. (5) Sick in the last 15 days: indicator variable that takes value of 1 for children who were reported to be sick (had any fever or diarrhea) in the 15 days previous to the interview. (6) Took a malaria test: indicator variable that takes value of 1 for children who took a malaria test, and zero for children who did not take a malaria test. (7) Took a malaria test if sick: indicator variable that takes value 1 for children who had malaria symptoms and took a malaria test, and 0 for children who had malaria symptoms but who did not take a malaria test. All specifications include an indicator variable for assignment to the Information/placebo treatment, an indicator variable for assignment to the Information campaign treatment, strata fixed effects, CHW and household level controls, as well as children's age fixed effects. The full list of controls is presented in Section 4. Standard errors are reported in parentheses. Standard errors are clustered at neighborhood level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. *P*-values adjusted for multiple hypothesis testing grouped by rows (at the row-level) are presented in squared brackets.

Table 4
Family planning and peri-natal care.

	Face-to-face survey						
	Use of family planning		Number of pre-natal visits	Pre-natal care index	Post-natal visit	Nursing	Administered vitamin A in the 45 days after giving birth
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Social status	0.03	0.01	0.07	0.06	−0.02	0.07	0.06**
	(0.02)	(0.02)	(0.26)	(0.05)	(0.06)	(0.05)	(0.03)
	[0.51]	[0.94]	[0.94]	[0.56]	[0.94]	[0.55]	[0.10]
Task significance	0.02	0.01	−0.15	−0.04	0.10	0.07	0.02
	(0.02)	(0.02)	(0.34)	(0.06)	(0.08)	(0.08)	(0.04)
	[0.77]	[0.94]	[0.94]	[0.86]	[0.62]	[0.80]	[0.94]
Observations	3166	2576	208	209	209	209	209
R^2	0.10	0.16	0.34	0.28	0.42	0.33	0.37
Mean (control group)	0.36	0.36	5.00	1.00	0.44	0.89	0.00
P-values:							
Joint test all treatments	0.418	0.870	0.597	0.618	0.554	0.315	0.201
Social status = Task significance	0.843	0.990	0.606	0.183	0.278	0.996	0.366
ANCOVA specification	No	Yes	No	No	No	No	No

Note: Estimates based on OLS regressions. Columns (1) and (3)–(7) present estimates using Eq. (1), column (2) presents estimates using Eq. (2), which includes the lagged dependent variable (ANCOVA). Estimation sample in columns (1) and (2) consists of women between the ages of 12 and 49, sexually active, and living in households interviewed in the endline face-to-face survey; estimation sample in columns (3)–(7) consists of women living in households interviewed in the endline face-to-face survey with children born alive in the two years previous to the interview. Depending on the column the dependent variables are defined by the following. (1)–(2) Use of family planning: indicator variable that takes value of 1 for women who report having used a family planning method in the last 12 months. (3) Number of pre-natal visits: number of pre-natal visits to a health facility during pregnancy, (4) Pre-natal care index: index variable averaging four indicator variables for the following pre-natal care exams and vaccine: blood pressure, blood test, urine test, and tetanus vaccine. (5) Post-natal visit: indicator variable that takes value of 1 for women who attended a post-natal visit to a health center after giving birth. (6) Nursing: indicator variable that takes value 1 for women who breast fed after giving birth. (7) Administered vitamin A in the 45 days after giving birth: indicator variable that takes value of 1 for women whose newborn was given vitamin A until 45 days after giving birth. All specifications include an indicator variable for assignment to the Information/placebo treatment, an indicator variable for assignment to the Information campaign treatment, strata fixed effects, CHW and household level controls, as well as women age. The full list of controls is presented in Section 4. Standard errors are reported in parentheses. Standard errors are clustered at neighborhood level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. *P*-values adjusted for multiple hypothesis testing grouped by rows (at the row-level) are presented in squared brackets.

the household level. Note that for most of these outcomes, beyond specification (1), we can employ baseline data, which allows us to use the ANCOVA specification (2).

We observe a clearly positive effect of Social status when considering our knowledge dependent variable. The magnitude is 0.28 standard deviation units, statistically significant at the 5 percent level. This

Table 5
Hospitals' and health centers' logbooks.

	Vaccination of children	Family planning	Birth at a health facility	Post-natal care	
	Vaccination index (5 vaccines)	First visit to the health center/hospital		Number of days between delivery and a post-natal check-up	Post-natal check-up in the 10 days after delivering
	(1)	(2)	(3)	(4)	(5)
Social status	0.02**	−0.03	0.03*	−12.21***	0.09**
	(0.01)	(0.03)	(0.02)	(4.30)	(0.04)
				[0.40]	[0.40]
Observations	3999	2331	1225	1066	1066
R^2	0.13	0.11	0.10	0.11	0.17
Mean (control group)	0.60	0.62	−0.14	38.92	0.29

Note: Estimates based on OLS regressions using Eq. (3). Estimation sample in column (1) consists of children who were under 2 years in 2018 with vaccination records from a health center or hospital in Bissau. Estimation sample in column (2) consists of women in fertile age in 2018 with family planning records from a health center or hospital in Bissau. Estimation sample in columns (3) to (5) consists of women in 2018 with post-natal records from a health center or hospital in Bissau. Depending on the column the dependent variables are defined by the following. (1) Vaccination index (5 vaccines): index variable averaging five indicator variables for taking each of the following vaccines: BCG, polio, DTcoq (diphtheria–tetanus–pertussis), MMR (measles–mumps–rubella), and yellow fever. (2) First visit to the health center/hospital: indicator variable that takes value of 1 for women visiting for the first time a health center or hospital for an appointment on family planning. (3) Birth at a health facility: indicator variable that takes value of 1 for women who gave birth at a health facility. (4) Number of days between giving birth and a post-natal visit: number of days between the registered date of birth and the visit to the health center/hospital for a post-natal check-up. (5) Post-natal visit in the 10 days after giving birth: indicator variable that takes value of 1 for women visiting the health center/hospital for a post-natal check-up in the 10 days after giving birth. All specifications include strata fixed effects, and neighborhood characteristics. The full list of controls is presented in Section 4. Standard errors are reported in parentheses. Standard errors are clustered at neighborhood level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. *P*-values adjusted for multiple hypothesis testing grouped by rows, following the structure of the aggregate indices in Figure 1, are presented in squared brackets.

is robust to employing the Romano–Wolf procedure for estimating *p*-values taking into account multiple hypotheses. Note that this effect is consistent with the direct effects on CHW performance, namely those on CHW learning (assessed through the test scores of CHWs, which included knowledge of the essential family practices) and house visiting efforts: it is likely that Social status incentives triggered effects on education for health in the households. Note that this effect is significantly different from that of the Task significance video intervention. We do not find other clear effects of Social status, except for a negative effect on the treating of water by the households (which did not have a significant influence over the treatment effect on the corresponding aggregated index as shown in Fig. 1). This is possibly evidence that there was more emphasis on other aspects of education for health, which were less known to households, with negative effects on this specific dimension. We do not find significant impacts of the Task significance treatment in Table 2.

In Table 3 we show treatment effects on outcomes related to the health of children under 5 years old. Our estimations are at the level of the child under 5 living in households interviewed in the endline face-to-face survey. We are able to employ ANCOVA specifications for the vaccination index.

We report positive effects of both incentive treatments on the probability of getting all five vaccines. In the case of Task significance, these effects are robust across specifications, i.e., with or without baseline dependent variables as controls, and robust across data sources, i.e., considering self-reports or observed bulletins. Note that levels of vaccination are relatively high: 75–87 percent of control children get all five vaccines. Magnitudes are 4 percentage points for the Social status treatment (observed bulletins, ANCOVA specification), significant at the 10 percent level, and 5 to 9 percentage points for the Task significance treatment, significant at the 1 percent level. Only the effects of task significance are robust to the Romano–Wolf correction. Differences across the two incentive treatments are statistically significant when baseline dependent variables are not employed as controls. Turning to the other outcome variables in the table, we find statistically significant impacts of the Social status treatment. Specifically, the probability of being sick in the last 15 days before the survey decreases by 8 percentage points — this effect is significant at the 5 percent level of statistical confidence (at the 10 percent level if employing the Romano–Wolf procedure). The probability of taking a malaria test decreases by 3 percentage points, but in case symptoms of malaria appear, it increases by 7 percentage points, although not significantly by any standards. This result suggests that the negative effect of Social status on taking a malaria test is driven by a lower probability of getting sick. We do not identify clear effects on these outcomes for the video intervention. In fact, the treatment effect of Social status is marginally different from the one of Task significance (*p*-value of 0.1) for the probability that the child was sick just before the endline survey.

Table 4 displays measures related to family planning and peri-natal care. The level of analysis is that of women living in households interviewed in the endline face-to-face survey. In the case of columns (1) and (2) these women had to be between the ages of 12 and 49 (sexually active). We restrict the analysis to the household head or the spouse of the household head. In the case of the remaining columns, women in our sample had children born alive in the two years before the survey interview. The variable on using family planning was available at the baseline and so we are able to employ the ANCOVA specification for this outcome variable.

We do not find clear effects of our incentive treatments on family planning and peri-natal care. The exception is the probability that women administered vitamin A to their newborns in the 45 days following their birth: the Social status induces an increase in this probability of 6 percentage points, which is significant at the 5 percent level (with a *p*-value of 0.10 when employing Romano–Wolf). Most other treatment effects of Social status are positive but do not reach significance at standard levels. We do not find any statistically significant differences between Social status and Task significance.

We conclude that Social status incentives had several positive effects on household health as measured in the context of our survey implementation. Knowing that the prospective of receiving employer and community recognition led to higher levels of CHW direct performance as measured in both incentivized and non-incentivized outcomes, including important aspects of CHW learning about how to effectively interact with households (shown in the previous section), it is not surprising that we find positive treatment effects of Social status on non-incentivized indicators of household health. This is the case specifically for knowledge about health practices in the household, vaccination of children, likelihood that children are not sick, and administration of vitamin A to newborns, although not always robustly to the Romano–Wolf procedure. Note that apart from positive effects on vaccination

of children, we do not find clear impacts of the Task significance intervention on household health. This means Hypothesis 1 for Social status is verified when considering household health. We find less evidence in favor of Hypothesis 2, which is limited to vaccination of children.

5.3.3. Household health — logbooks from health centers and hospitals

We now turn to analyzing health records from the health centers' and the hospitals' logbooks in the city of Bissau. The analysis is restricted to the Social status intervention (see Section 3.3 for details). Table 5 depicts the corresponding results. In the case of column (1), our estimation is at the level of children who were two years or younger in October 2018 while having received a first vaccine at a health center or hospital in Bissau after September 2017.[47] Column (2) takes as sample those women visiting a health center or hospital for a family planning appointment in 2018. In columns (3)–(5), our sample is formed by women visiting a health center or hospital for a post-natal check-up in 2018.[48]

We find positive and significant effects of Social status on the probability of getting all five vaccines. This result supports our previous findings employing the survey-based measurements. Indeed, the magnitude of the effect is very similar, around 2 percentage points, significant at the 5 percent level. We do not find effects of the Social status treatment on the probability of visiting for the first time a health center or hospital for an appointment on family planning. Turning to the outcome variables taken from the post-natal logbook, we also find several statistically significant impacts of the Social status treatment. Specifically, the probability of giving birth at a health facility increases by 3 percentage points, which is significant at the 10 percent level. We also report a drop of 12 days on the number of days elapsed between delivery and a post-natal check-up, and an increase of 9 percentage points on the probability of attending a post-natal check-up in the first 10 days after birth. These estimates are statistically significant at the 1 or 5 percent levels, respectively, although not robustly to the Romano–Wolf multiple hypothesis correction for post-natal outcomes.

We conclude that the Social status intervention had important treatment effects on actual health indicators related to immunization of children, as well as on peri-natal care, namely on assisted births. These findings reinforce the validity of Hypothesis 1 on the positive impact of incentives related to Social status.

5.4. Other treatment effects

We now devote our attention to the secondary hypotheses in our experimental design. First, we distinguish between all the video treatments, including the comparison between Task significance alone and Information/Placebo, as well as between Task significance alone and Task significance plus endorsement by the traditional healers. Second, we evaluate the impact of the remaining treatment in our experimental design, i.e., the SMS information campaign. Third, we test for the complementarity between our interventions. We undertake the analysis by employing the aggregated indices.

First, as shown in Table H7, Panel A, in the Online Appendix, we find some statistically significant differences across the video treatment effects. For some of our outcomes of interest, we find smaller effect sizes for the Placebo/information video, when compared to Task significance alone. This is the case for health indicators at the level of the household and for health of children under 5 years old. This is indicative of an effect of task significance over that of information, as we hypothesized. More surprisingly, we find smaller effect sizes for the full video on Task significance plus endorsements, relative to Task significance alone, when considering health of children under 5 years old, as well as family planning and peri-natal care. This is suggestive that the endorsements of traditional healers could actually be detrimental.

Second, we do not find any statistically significant treatment effects of the Information campaign delivered through text messages to households. This is included in Table H7, Panel B, in the Online Appendix. Possible interpretations for this null result are that: (i) SMSs did not reach households to the extent we expected; (ii) households were sufficiently aware of CHWs activity; and (iii) CHWs did not use increased trust by the households to improve their productivity.

Third, we turn to testing the complementarity between incentive treatments, i.e., Social status (award, ceremony, and SMS) and Task significance (video), and between each one of these incentive treatments and the Information campaign (SMS). Our cross-randomization design enabled all the interactions between the different treatment groups. These are displayed in Tables H8 and H9 in the Online Appendix. We do not find statistically significant interaction coefficients. We conclude that there is no systematic evidence in favor of complementarities between interventions. A possibility is that limited statistical power, together with limited impacts of the Task significance intervention and of the Information campaign, prevent us from achieving statistical significance on some of these interaction effects.

5.5. LATE of the video treatments

We now explore the availability of data on compliance with the video treatments to identify the treatment effects of having visualized the different video interventions. Table H10 in the Online Appendix shows the number of treatment rounds actually attended by CHWs in the different video comparison groups. There we see that only 7 to 11 percent of the CHWs in the video treatment groups were not exposed to any round of treatment — the number of non-complying CHWs per video treatment is included in the CONSORT diagram in Figure D1 of the Online Appendix. The average number of rounds of video treatments ranged from 2.2 to 2.3, with most CHWs in each video treatment group having watched the full three rounds of treatment. At the same time no video control CHWs watched any round of video treatment.

Table 6 shows the effects of having visualized each type of video, where we instrument visualization of a given video by the random assignment to that treatment condition. In other words, the endogenous variables of interest are defined as having visualized at least one round of the corresponding video treatments. We are thus estimating the LATE of the video treatments. While the relevance of the three instruments is difficult to dispute (notwithstanding, we show tests of weak instruments), the exclusion restriction is also likely to be valid in face of the implausibility of direct impacts of invitations to watch the video treatments. We conduct our analysis by employing as outcome measures the aggregate z-scores we introduced before.

We find clearly positive effects of watching the video dedicated to task significance alone. These are effects on CHW performance as well as on health of children under 5 years old. Magnitudes are 0.39 and 0.10 standard deviation units (respectively), statistically significant at

[47] Each child getting a vaccine in a health center or hospital is registered at the date of the first visit, with all the follow-up visits recorded in the same logbook page. Thus, older children's vaccination records are included in previous years' logbooks, to which we did not have access.

[48] We note that our data relates to public health provision. Although private provision of health services is limited in our setting, and mainly related to private medical appointments by individual doctors, we check treatment effects on endline survey data reporting public vs. private provision in a variety of health services. This is shown in Table H6 in the Online Appendix. We do not find any treatment effects in this dimension, which reassures us against interpretations of treatment effects on the public logbooks driven by more/less use of the public vs. private providers.

Table 6
LATE of the video treatments — aggregated outcomes employing z-scores.

	Administrative data	Phone survey	Face-to-face survey			
	CHW performance	Home visits	Home visits	Health at household level	Health of children under 5 years old	Family planning and peri-natal care
	(1)	(2)	(3)	(4)	(5)	(6)
Watched the video on task significance alone	0.39***	0.06	0.08	0.01	0.10**	0.14
	(0.10)	(0.06)	(0.06)	(0.07)	(0.04)	(0.09)
Watched the video on task significance plus endorsement	0.39***	0.04	0.06	−0.13	0.03	0.00
	(0.11)	(0.06)	(0.06)	(0.08)	(0.05)	(0.09)
Watched the placebo/information video	0.41***	−0.01	0.03	−0.10	0.03	−0.06
	(0.11)	(0.07)	(0.05)	(0.07)	(0.04)	(0.09)
Observations	1015	2018	1748	1765	1295	1448
Mean (control group)	−0.12	−0.10	−0.18	−0.18	0.08	0.06
Test of weak instruments — F-statistic						
Task significance alone	64179.4	33413.8	49721.2	53126.3	19043.9	51489.1
Task significance plus endorsement	25117.5	53120.0	27351.8	29820.7	6080.1	12923.1
Information/placebo	44892.5	25486.6	35604.8	35931.0	11824.5	27551.9
Task significance alone = Task significance plus endorsement	0.958	0.766	0.775	0.039	0.058	0.043
Task significance alone = Information/placebo	0.794	0.263	0.296	0.074	0.037	0.009
Task significance plus endorsement = Information/placebo	0.806	0.370	0.576	0.740	0.964	0.447

Note: Estimates based on Two-stage Least Squares (2SLS) estimation, where random assignment to Task significance alone, Task significance plus endorsement, and Information/placebo are used as instrumental variables for having visualized (respectively) the interactive video on task significance alone, the interactive video on task significance plus the endorsements by the traditional healers, and the information/placebo video. The specification of the second stage is given by Eq. (1) including treatment variables for the two comparison groups of the Task significance video. The F-statistics of the tests of weak instruments are displayed at the bottom of the table. Outcomes are grouped in indices that are built using the procedure in Kling et al. (2007). We calculate within-sample z-scores for each individual outcome, employing the mean and the standard deviation of the pure control group. We then obtain the unweighted average z-score for each category. Depending on the column, the indices are defined by the following outcomes: (1) CHW performance (administrative data): formed from outcomes in columns (1)–(4) of Table 1; (2) Home visits (phone survey): formed from outcomes in columns (5)–(6) of Table 1; (3) Home visits (face-to-face survey): formed from outcomes in columns (7)–(8) of Table 1; (4) Health at household level (face-to-face survey): formed from outcomes in columns (1), (2), (4), and (6) of Table 2; (5) Health of children under 5 (face-to-face survey): formed from outcomes in columns (1), (3), and (5)–(7) of Table 3; (6) Family planning and peri-natal care (face-to-face): formed from outcomes in columns (1), and (3)–(7) of Table 4. All specifications employed include an indicator variable for assignment to the Information campaign treatment, strata fixed effects, and CHW level controls. Specifications in columns (3)–(6) include household level controls. The specification in column (5) includes age fixed effects for the children under 5 years old. The specification in column (6) includes respondent's age. The full list of controls is presented in Section 4. Standard errors are reported in parentheses. Standard errors are clustered at neighborhood level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

the 1 and the 5 percent levels of confidence.[49] We conclude that watching the task significance alone video may have led to improvements in CHW performance and the health of children under 5 years of age. The clearer LATE when compared to the weaker intent-to-treat effects suggests that treatment effects are more centered around compliers to the video treatments.

5.6. Additional results and robustness

We now turn to a few auxiliary results shown in the Online Appendix. First, we analyze treatment effects on self-reported CHW motivation. This is assessed through standard survey questions on the role of monetary awards and social recognition (Amabile et al., 1994), as well as of social impact (Grant and Campbell, 2007) in motivating CHWs. These questions were submitted in both baseline and endline face-to-face surveys of CHWs. Hence, we are able to employ ANCOVA specifications when using these data as dependent variables. Table H11 is devoted to the estimation of the main treatment effects, i.e., on the impact of Social status and Task significance. We find positive and significant effects of Social status on CHWs motivation through social impact. In other words, CHWs rationalize additional motivation in face of the Social status intervention, not through added social recognition but through a perception that their impact in their communities has increased. Magnitudes are 1 percent of the scale employed, with confidence levels at the 1 percent. These effects are however indistinguishable from those of task significance.[50]

Second, we report in Table H13 the main treatment effects on different dimensions of household knowledge about the health practices conveyed by CHWs to households as part of their CHW mandate. This is disaggregating the outcome variable of regression (1) in Table 2. The outcomes are constructed from questions in the endline face-to-face household survey.[51] We distinguish between knowledge about newborn care, nutrition, hygiene/washing hands, use of latrines, water treatment, preventive measures about tuberculosis and HIV, pre-natal care, alert signals of illnesses, and family planning. We observe positive treatment effects of Social status on knowledge about newborn care, nutrition, washing hands, and use of latrines. These effects range from 1 to 8 percentage points.

Third, we provide evidence that the effects of the incentives schemes are not contaminated by spillovers, namely by agents in other treatments reacting to not having had the possibility of getting an award or visualizing the task significance video. We exploit CHWs' networks within the program at the baseline to test whether individuals in the control group, who know CHWs assigned to the treatments and therefore are more likely to be affected by spillovers, show different levels of the outcome variables. Figure H3 shows that the number of CHWs known in each treatment group by the individuals in the control group does not systematically affect the aggregated outcomes.

[49] Like in the reduced form results, there are positive and statistically significant differences between Task significant alone and both Placebo/information and Task significance plus endorsement. Here that happens when employing the index on health at the household level, the index on health of children under 5 years old, and the index on family planning and peri-natal care.

[50] In fact, looking at Table H12, which reports on the other treatment effects in our design, we find significant effects of Task significance alone on social impact. We also observe positive and significant differences between the simple Task significance video treatment and the full video treatment (with endorsements) on social impact.

[51] The full list of questions is included in Section E in the Online Appendix.

Fourth, we study whether our performance-based, Social status incentive, creates heterogeneous effects among CHWs, namely demotivating those with a low chance of winning the awards. To assess this possibility, we allow the effects of Social status to differ between those who won the award and those who did not. Figure H4 shows larger treatment effects for those who were awarded, of 0.15–0.18 standard deviations on home visits (significant at the 5 percent level), and of 0.09 standard deviations on health of children under 5 (significant at the 5 percent level). Our results suggest that both groups (awarded and non-awarded) performed at a similar level when considering health indicators at the household level, and peri-natal care.[52] We can then conclude for some limited evidence of heterogeneous effects of Social status.

A final note goes to robustness exercises we conduct on the choice of control variables for CHWs and households, as well as on multiple hypothesis testing. In Section I of the Online Appendix we show the replication of the main results of the paper while employing the Post-double Selection Lasso procedure for selecting the referred control variables. In Section J of the Online Appendix we report *p*-values of the procedure described in Romano and Wolf (2016), which we employ to account for multiple hypothesis at the row level grouped by the source of the data (like in our indices) and accounting for the two main treatment effects at the same time, for each of the main tables of the paper. As expected, *p*-values are generally higher, although still significant for some of the individual results in the paper.

6. Concluding remarks

In this paper we report on the results of a field experiment testing the impact of non-financial incentives for CHWs in Guinea-Bissau. Specifically, we follow the activation of social status, through the attribution of honorific awards for good performance, and of intrinsic motivation, through a video treatment that establishes the task significance of CHWs in saving lives. We vary the components of the video to isolate the impact of a basic video without the task significance component, and the marginal impact of endorsements by traditional healers.

The main finding is that raising the social status of CHWs is effective at improving their direct performance in terms of home visits, as well as household health in terms of household knowledge of good practices, the health of children under 5 years old, and peri-natal care. We find positive effects of the task significance video, particularly for compliers, which are difficult to distinguish from the basic video treatment. Endorsements of traditional healers are not improving our outcomes of interest.

Volunteer health workers constitute an essential part of the health system, not only in settings like the one we study in Guinea-Bissau, but also in most of Sub-Saharan Africa. It is a shared belief by all stakeholders in these health systems that no easy path to professionalizing these volunteers is available due to limited resources. Although the role of financial incentives/professionalization is likely important, this paper devotes attention to short-run and inexpensive strategies to keep CHWs motivated. We show that increasing the social status of these health workers, while incentivizing worker's learning about health practices, is an effective strategy to improve their performance and impact on relevant dimensions of household health. We should note that it is possible that increasing the social status of CHWs could have a future financial impact on them, e.g., through better jobs and financial opportunities. In this sense, clear treatment effects of social status could be interpreted as a confirmation that financial incentives are effective. In any case, the implied direct policy recommendation on increasing the social status of CHWs for settings like the one we study in Sub-Saharan Africa is clear in face of the negligible costs of that intervention. Beyond our attempt to employ task significance through videos to incentivize CHWs, more work is needed to understand the triggers of intrinsic motivation for pro-social behavior by CHWs.

[52] Consistently, we find that treatment effects spread over the whole distribution of the main performance indicator we employed (Figure H5).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Replication files are available at: http://dx.doi.org/10.17632/ppbfj4ttrt.1

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jdeveco.2023.103096.

References

Ager, Philipp, Bursztyn, Leonardo, Leucht, Lukas, Voth, Hans-Joachim, 2022. Killer incentives: Rivalry, performance and risk-taking among German fighter pilots, 1939–45. Rev. Econom. Stud. 89 (5), 2257–2292.

Akerlof, George A., Kranton, Rachel E., 2000. Economics and identity. Q. J. Econ. 115 (3), 715–753.

Akerlof, George A., Kranton, Rachel E., 2005. Identity and the economics of organizations. J. Econ. Perspect. 19 (1), 9–32.

Amabile, Teresa M., Hill, Karl G., Hennessey, Beth A., Tighe, Elizabeth M, 1994. The work preference inventory: Assessing intrinsic and extrinsic motivational orientations. J. Personal. Soc. Psychol. 66 (5), 950–967.

Ariely, Dan, Bracha, Anat, Meier, Stephan, 2009. Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. Amer. Econ. Rev. 99 (1), 544–555.

Ashraf, Nava, Bandiera, Oriana, 2018. Social incentives in organizations. Annu. Rev. Econ. 10, 439–463.

Ashraf, Nava, Bandiera, Oriana, Davenport, Edward, Lee, Scott S, 2020. Losing prosociality in the quest for talent? Sorting, selection, and productivity in the delivery of public services. Amer. Econ. Rev. 110 (5), 1355–1394.

Ashraf, Nava, Bandiera, Oriana, Jack, B. Kelsey, 2014a. No margin, no mission? A field experiment on incentives for public service delivery. J. Publ. Econ. 120, 1–17.

Ashraf, Nava, Bandiera, Oriana, Lee, Scott S., 2014b. Awards unbundled: Evidence from a natural field experiment. J. Econ. Behav. Organ. 100, 44–63.

Banuri, Sheheryar, Keefer, Philip, De Walque, Damien, 2018. Love the Job…or the Patient? Task vs. Mission-Based Motivations in Health Care. Policy Research Working Paper 8338, World Bank.

Bénabou, Roland, Tirole, Jean, 2003. Intrinsic and extrinsic motivation. Rev. Econom. Stud. 70 (3), 489–520.

Bénabou, Roland, Tirole, Jean, 2006. Incentives and prosocial behavior. Amer. Econ. Rev. 96 (5), 1652–1678.

Besley, Timothy, Ghatak, Maitreesh, 2005. Competition and incentives with motivated agents. Amer. Econ. Rev. 95 (3), 616–636.

Björkman Nyqvist, Martina, Guariso, Andrea, Svensson, Jakob, Yanagizawa-Drott, David, 2019. Reducing child mortality in the last mile: Experimental evidence on community health promoters in uganda. Am. Econ. J.: Appl. Econ. 11 (3), 155–192.

Bradler, Christiane, Dur, Robert, Neckermann, Susanne, Non, Arjan, 2016. Employee recognition and performance: A field experiment. Manage. Sci. 62 (11), 3085–3099.

Bursztyn, Leonardo, Jensen, Robert, 2017. Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure. Annu. Rev. Econ. 9, 131–153.

Carpenter, Jeffrey, Gong, Erick, 2016. Motivating agents: How much does the mission matter? J. Labor Econ. 34 (1), 211–236.

Chan, Ho Fai, Frey, Bruno S., Gallus, Jana, Torgler, Benno, 2014. Academic honors and performance. Labour Econ. 31, 188–204.

Christopher, Jason B., Le May, Alex, Lewin, Simon, Ross, David A., 2011. Thirty years after alma-ata: a systematic review of the impact of community health workers delivering curative interventions against Malaria, Pneumonia and Diarrhoea on child mortality and morbidity in Sub-Saharan Africa. Hum. Resour. Health 9 (1), 1–11.

Cotofan, Maria, 2021. Learning from praise: Evidence from a Field Experiment with Teachers. J. Publ. Econ. 204, 104540.

Dana, Jason, Weber, Roberto A., Kuang, Jason Xi, 2007. Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. Econom. Theory 33 (1), 67–80.
DellaVigna, Stefano, Pope, Devin, 2017. What motivates effort? Evidence and expert forecasts. Rev. Econom. Stud. 85 (2), 1029–1069.
Deserranno, Erika, 2019. Financial incentives as signals: Experimental evidence from the recruitment of village promoters in uganda. Am. Econ. J.: Appl. Econ. 11 (1), 277–317.
Deserranno, Erika, Kastrau, Philipp, León Ciliotta, Gianmarco, 2021. Promotions and Productivity: The Role of Meritocracy and Pay Progression in the Public Sector. Discussion Paper 15837, CEPR.
Francois, Patrick, 2000. 'Public service motivation' as an argument for government provision. J. Publ. Econ. 78 (3), 275–299.
Gagné, Marylène, 2014. The Oxford Handbook of Work Engagement, Motivation, and Self-Determination Theory. Oxford University Press, USA.
Gagné, Marylène, Deci, Edward L., 2005. Self-determination theory and work motivation. J. Organ. Behav. 26 (4), 331–362.
Gallus, Jana, 2017. Fostering public good contributions with symbolic awards: A large-scale natural field experiment at wikipedia. Manage. Sci. 63 (12), 3999–4015.
Gauri, Varun, Jamison, Julian C., Mazar, Nina, Ozier, Owen, 2021. Motivating bureaucrats through social recognition: External validity—A tale of two states. Organ. Behav. Hum. Decis. Process. 163, 117–131.
Gilmore, Brynne, McAuliffe, Eilish, 2013. Effectiveness of community health workers delivering preventive interventions for maternal and child health in low-and middle-income countries: a systematic review. BMC Publ. Health 13 (1), 847.
Glassman, Amanda, Temin, Miriam, 2016. Millions Saved: New Cases of Proven Success in Global Health. Brookings Institution Press.
Grant, Adam M., 2007. Relational job design and the motivation to make a prosocial difference. Acad. Manag. Rev. 32 (2), 393–417.
Grant, Adam M., 2008. The significance of task significance: Job performance effects, relational mechanisms, and boundary conditions. J. Appl. Psychol. 93 (1), 108.
Grant, Adam M., Campbell, Elizabeth M., 2007. Doing good, doing harm, being well and burning out: The interactions of perceived prosocial and antisocial impact in service work. J. Occup. Organ. Psychol. 80 (4), 665–691.
Haines, Andy, Sanders, David, Lehmann, Uta, Rowe, Alexander K., Lawn, Joy E., Jan, Steve, Walker, Damian G., Bhutta, Zulfiqar, 2007. Achieving child survival goals: Potential contribution of community health workers. Lancet 369 (9579), 2121–2131.
Holm, Sture, 1979. A simple sequentially rejective multiple test procedure. Scand. J. Stat. 6 (2), 65–70.
Holmstrom, Bengt, Milgrom, Paul, 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. J. Law Econ. Organ. 7 (1), 24–52.
Kamenica, Emir, 2012. Behavioral economics and psychology of incentives. Annu. Rev. Econ. 4 (1), 427–452.
Khan, Muhammad Yasir, 2020. Mission motivation and public sector performance: Experimental evidence from Pakistan. Mimeo.
Kling, Jeffrey R., Liebman, Jeffrey B., Katz, Lawrence F., 2007. Experimental analysis of neighborhood effects. Econometrica 75 (1), 83–119.
Kosfeld, Michael, Neckermann, Susanne, 2011. Getting more work for nothing? Symbolic awards and worker performance. Am. Econ. J. Microecon. 3 (3), 86–99.
Lee, Scott, 2018. Intrinsic Incentives: A Field Experiment on Leveraging Intrinsic Motivation in Public Service Delivery. Working Paper 3537336, SSRN.
McKenzie, David, 2012. Beyond baseline and follow-up: The case for more T in experiments. J. Dev. Econ. 99 (2), 210–221.
Neckermann, Susanne, Cueni, Reto, Frey, Bruno S., 2014. Awards at work. Labour Econ. 31, 205–217.
Olaniran, Abimbola, Smith, Helen, Unkels, Regine, Bar-Zeev, Sarah, van den Broek, Nynke, 2017. Who is a community health worker? A systematic review of definitions. Global Health Act. 10 (1), 127–223.
Prendergast, Canice, 2007. The motivation and bias of bureaucrats. Amer. Econ. Rev. 97 (1), 180–196.
Romano, Joseph P., Wolf, Michael, 2005a. Exact and approximate stepdown methods for multiple hypothesis testing. J. Amer. Statist. Assoc. 100 (469), 94–108.
Romano, Joseph P., Wolf, Michael, 2005b. Stepwise multiple testing as formalized data snooping. Econometrica 73 (4), 1237–1282.
Romano, Joseph P., Wolf, Michael, 2016. Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. Statist. Probab. Lett. 113, 38–40.
Ryan, Richard M., Deci, Edward L., 2000a. Intrinsic and extrinsic motivations: Classic definitions and new directions. Contemp. Educ. Psychol. 25 (1), 54–67.
Ryan, Richard M., Deci, Edward L., 2000b. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am. Psychol. 55 (1), 68.
Taylor, Celia, Griffiths, Frances, Lilford, Richard, 2017. Affordability of comprehensive community health worker programmes in rural Sub-Saharan Africa. BMJ Glob Health 2 (3), e000391.
Zambruni, Jerome Pfaffmann, Rasanathan, Kumanan, Hipgrave, David, Miller, Nathan P., Momanyi, Maureen, Pearson, Luwei, Rio, Dolores, Romedenne, Magali, Singh, Sagri, Young, Mark, et al., 2017. Community health systems: Allowing community health workers to emerge from the shadows. Lancet Global Health 5 (9), e866–e867.