American Economic Review 2020, 110(5): 1355–1394
https://doi.org/10.1257/aer.20180326

Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services[†]

By Nava Ashraf, Oriana Bandiera, Edward Davenport, and Scott S. Lee*

We embed a field experiment in a nationwide recruitment drive for a new health care position in Zambia to test whether career benefits attract talent at the expense of prosocial motivation. In line with common wisdom, offering career opportunities attracts less prosocial applicants. However, the trade-off exists only at low levels of talent; the marginal applicants in treatment are more talented and equally prosocial. These are hired, and perform better at every step of the causal chain: they provide more inputs, increase facility utilization, and improve health outcomes including a 25 percent decrease in child malnutrition. (*JEL* H83, I11, I13, J24, M51, O15, Z13)

Economic development entails the professionalization of public service delivery, whereby career professionals replace informal local providers (Northcote and Trevelyan 1853, Weber 1922, North 1991). This has raised concerns of a possible trade-off between qualifications and skills on the one hand, and intrinsic motivation and local rapport on the other. In other words, does a career in the civil service attract talent at the expense of prosociality?[1]

*Ashraf: Department of Economics, LSE (email: n.ashraf1@lse.ac.uk); Bandiera: Department of Economics, LSE, and STICERD (email: o.bandiera@lse.ac.uk); Davenport: Department of Economics, MIT (email: edav@mit.edu); Lee: Department of Medicine and Institute for Global Health, Vanderbilt University (email: ssl@mail.harvard.edu). Roland Bénabou was the coeditor for this article. We thank the Ministry of Health of Zambia and especially Mutinta Musonda for partnership on this project. We thank the IGC, J-PAL Governance Initiative, USAID, and HBS DFRD for financial support. We also thank Robert Akerlof, Charles Angelucci, Tim Besley, Robin Burgess, Paul Gertler, Edward Glaeser, Kelsey Jack, Giacomo Ponzetto, Imran Rasul, Jonah Rockoff, and seminar participants at several institutions for useful comments. Adam Grant, Amy Wrzesniewski, and Patricia Satterstrom kindly provided guidance on psychological measurement. We thank Kristin Johnson, Conceptor Chilopa, Mardieh Dennis, Madeleen Husselman, Alister Kandyata, Allan Lalisan, Mashekwa Maboshe, Elena Moroz, Shotaro Nakamura, Sara Lowes, and Sandy Tsai for the excellent research assistance and the Clinton Health Access Initiative in Zambia for their collaboration. IRB approval was obtained from Harvard University, Committee on the Use of Human Subjects in Research, and University of Zambia, Biomedical Research Ethics Committee. This study is registered in the AEA RCT Registry and the unique identifying number is AEARCTR-0000064. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

[†] Go to https://doi.org/10.1257/aer.20180326 to visit the article page for additional materials and author disclosure statements.

[1] Ambition toward a career in public service, with both its ability to attract the most able but also the most self-interested, has a long intellectual history; ambition was used by Romans, as *ambitio*, exclusively to refer to those in public life. In *De officiis*, Cicero referred to *ambitio* as a "malady" that can cause individuals to "lose sight of their claims to justice," but one that seems to draw "the greatest souls" and "most brilliant geniuses" (King 2013).

We design a field experiment as part of a nationwide recruitment drive for a new health care position in Zambia to test whether career benefits attract talent at the expense of prosocial motivation. We collaborate with the government of Zambia as they formalize primary health care in remote rural areas by creating a new health worker position in the civil service. This cadre is meant to replace informal service provision by religious and other charitable organizations, thereby following a typical professionalization process. The stakes are high because, due to the shortage of medical staff, hiring effective agents can make a great difference for the quality of health services and, ultimately, health outcomes in these communities.

Our experiment varies the salience of a career in the civil service at the recruitment stage, exploiting the fact that this position is new to potential applicants. In control districts, the recruitment ads reflect the status quo, in which local health services are provided by individuals hired by nongovernmental and charitable organizations. Helping the community is listed as the main benefit, and local agents are listed as peers. In treatment districts, the ads reflect the fact that this is a formal employment opportunity within the government: career advancement is listed as the main benefit, and doctors and nurses are listed as peers. Treatment and control posters differ only in the salience of career opportunities, while all other factors such as application requirements are kept equal.

To isolate the effect of selection on performance, we must sever the link between treatment and incentives on the job. To this end, all hired agents are given the same information on career opportunities as soon as they join, and then are trained together for one year before deployment. A survey administered before and after the training program validates our design: before training, treatment and control agents differ in the perceived relevance of career benefits, but after training these perceptions converge.

The new health worker position effectively adds career opportunities to a job with social impact. Theoretically, the effect of extrinsic rewards is ambiguous because performance in public service delivery depends on the agents' skills (both cognitive and non) and their social preferences, that is the extent to which they internalize the utility of the recipients of the services. A tension then arises if extrinsic rewards attract talented agents, whose effort is more productive, at the expense of prosocial agents who, other things equal, exert more effort. This is the extensive margin equivalent of motivation crowding-out, whereby extrinsic rewards can reduce performance by reducing the agent's intrinsic motivation (Bénabou and Tirole 2003, 2006). This tension also underpins a frequent argument made by policymakers that extrinsic rewards should be kept low so as to draw in agents who care sufficiently about delivering services per se. A simple conceptual framework makes precise that, in line with prevailing policy concerns, this attracts applicants who are less prosocial conditional on a given level of talent. However, since the outside option is increasing in talent, adding career benefits will draw in more talented individuals, and the marginal, most talented applicant in both groups will have the highest prosociality. Intuitively, since a candidate with high ability will also have a high outside option, if they are applying for the health worker position it must be because they are highly prosocial. The treatment effect on recruited candidates will therefore depend on how candidates are chosen from the pool. If applicants are drawn randomly, there might be a trade-off between talent and prosociality. However, if only the most talented are hired, there will be no trade-off.

To evaluate the impact of treatment on the applicant pool and on hired candidates we collect information on the skills and prosociality of every applicant. This exercise reveals that, in line with common intuition, the average applicant in treatment is more talented and less prosocial. In line with the theoretical intuition, however, the most talented applicants have the same, high level of prosociality. We show that the selection panels in both treatment and control put a high weight on talent, leading them to recruit among the most talented in their pool; as a result, treatment recruits are more talented and equally prosocial.

To evaluate the impact of treatment on service delivery we combine three data sources: real-time data on service delivery in remote areas collected through a mobile platform, administrative data on health facility utilization, and our own survey of household health practices and outcomes, including immunization records and anthropometrics. This allows us to link the services delivered by the newly recruited health workers to the outcomes of the households who receive those services and, ultimately, their health impact.

We find that agents drawn by career opportunities are more effective at each step of the causal chain, from the inputs they provide to the outcomes of the recipients. They provide more inputs (29 percent more household visits, twice as many community meetings) at the same cost. They increase facility utilization rates: the number of women giving birth at the health center is 30 percent higher, and the number of children undergoing health checks is 24 percent higher, being weighed 22 percent higher, and receiving immunization against polio 20 percent higher. They improve a number of health practices among the households they serve: breastfeeding and proper stool disposal increase by 5 and 12 percentage points (pp), respectively, deworming treatments by 16 percent, and the share of children on track with their immunization schedule by 5 pp (relative to a control mean of 6 percent). These changes are matched by changes in objective health outcomes: the share of children under age 5 who are underweight falls by 25 percent.

Taken together, these results indicate that offering a civil service position with career opportunities attracts agents who deliver services with remarkable health impact. The fact that we observe consistently positive impacts from three distinct and entirely independent data sources further strengthens our confidence in the findings.

The study of how individuals sort into jobs according to their preferences, skills, and the jobs' own attributes has a long tradition in economics (Roy 1951). More recently, this has been enriched by the study of job mission as a selection and motivation mechanism (Besley and Ghatak 2005) and identity or self-image as components of preferences (Akerlof and Kranton 2005, Bénabou and Tirole 2011). Our findings provide empirical support to these contributions as they suggest that the identity associated with the job can affect those drawn to it and that this selection affects performance.[2]

The fact that career opportunities affect performance *through selection* complements the recent findings of Bertrand et al. (forthcoming) that, on the intensive

[2] In light of the evidence of poor bureaucratic performance in low-income countries (Collier 2009, Muralidharan et al. 2011) our findings suggest that this is not due to the fact that civil service careers attract poor performers when these jobs are first created.

margin, better promotion prospects improve the effectiveness of Indian civil servants. Our findings also complement a large literature on the impact of financial incentives. On the selection margin, Dal Bó, Finan, and Rossi (2013) and Deserranno (2019) study the effect of earnings levels on the traits of applicants for government and NGO jobs[3] while several papers evaluate the effect of performance pay on the performance of agents after these have been hired either for the delivery of health services (Ashraf, Bandiera, and Jack 2014; Miller et al. 2012; Miller and Babiarz 2014; Celhay et al. 2019) or education (Muralidharan and Sundararaman 2011; Duflo, Hanna, and Ryan 2012; Glewwe, Ilias, and Kremer 2010; Fryer 2013; Rockoff et al. 2012; Staiger and Rockoff 2010). Our contribution is to provide the first experimental evidence that selection affects performance in public service delivery. In particular, we show that job design, of which incentives are a component, affects who sorts into these jobs in the first place, and that the effect of this selection on performance is of the same order of magnitude as the largest incentive effects estimates.[4,5]

The rest of the paper is organized as follows. Section I describes the context and research design. Section II develops a conceptual framework to make precise the trade-off between talent and prosociality, and Section III tests for it in the applicant pool and among recruited candidates. Section IV evaluates the treatment effect on performance in delivering health services. Section V evaluates the treatment effect on facility utilization, health behaviors, and health outcomes. Section VI concludes with a discussion of external validity, welfare implications, and general equilibrium effects relevant for program scale-up.

I. Context and Research Design

A. *Context: Health Services in Rural Communities*

Delivering health services to remote rural areas is challenging at every level of development because trained medical staff are reluctant to be posted there and turnover rates are high (World Health Organization 2006).[6] In Zambia, as of 2010, the average health post (the first-level government health facility) had 1.5 staff from the Ministry of Health, including those not permanently based there. The government Community Health Assistant (CHA) position was created as a solution to this challenge. The position is meant to put formally trained government health staff in place of informal community health workers employed, often as volunteers, by religious

[3] Dal Bó, Finan, and Rossi (2013) finds that higher salaries for civil service jobs attract better qualified candidates with the same level of prosocial preferences. Deserranno (2019) finds that expectations of higher earnings discourage prosocial candidates from applying for an NGO job that encompasses both commercial and health promotion activities. While consistent with these selection effects, our experiment focuses on measuring the effect of selection on agents' performance and beneficiaries' outcomes, which encompasses the effect of all the attributes that determine effectiveness.

[4] There is a corresponding literature that studies the same issues in the private sector. This literature stresses the theoretical importance of the effect of incentives on selection, but empirical studies focus on incentives on the job (Lazear and Oyer 2012, Oyer and Schaefer 2011).

[5] Rothstein (2015) uses a model-based approach that simulates the selection effect of alternative teachers' contracts. He finds that bonus policies have small effects on selection while reductions in tenure rates accompanied by substantial salary increases and high firing rates can have larger effects.

[6] The US Health Resources and Services Administration estimates that 1 in 2 rural Americans lives in a medically underserved area with a shortage of primary care providers (PCPs), defined as a population-to-PCP ratio of greater than 3,500:1 (US Human Resources and Service Administration 2019).

and other nonprofit organizations.[7] In fact, the title of the newly created government position was changed to "Community Health Assistant" shortly after its inception precisely to distinguish it from the informal, untrained community health workers.

In 2010, the program's first year, the government sought to recruit, train, and deploy 2 health workers to each of 167 communities in 48 districts. The main task of CHAs is to visit households and refer them to health facilities as needed. The job requires both medical and social skills (World Health Organization 2006). Medical skills include taking vital signs, diagnosing and triaging common illnesses, filling out patient registries, and performing first aid. Social skills include counseling, supporting, advising, and educating patients and other laypeople.

Government-funded community health worker programs vary in the extent to which they integrate the health workers into the civil service. At one extreme there are programs that mimic the informal model with financing provided by the government and all other decisions including hiring, monitoring, and firing left to the nongovernmental sector. At the other extreme is the model adopted in this program in Zambia where health workers become a cadre of civil servants and can advance to higher-ranked and better-paid cadres. The pay gradient is steep as the starting monthly wage is US$290 for CHAs and US$530 for entry-level nurses.[8] Promotion into higher-ranked cadres within the Ministry requires additional training (for example, nursing or medical school). Being part of the civil service, the health workers are eligible for in-service training, meaning that they attend school while continuing to receive salary and in some cases sponsorship of tuition as well. The official policy of the Ministry is to periodically ask the district medical officers to nominate a number of candidates on merit, but there is no mechanical link between quantitative measures of performance (say, the number of visits that a health worker makes) and nominations. Promotions to higher cadres are therefore not automatic, but the expected payoff is high even with low success rates, especially because job opportunities that allow for a career in central government are rare in the remote communities from which the health workers are recruited.[9]

The government chose the latter model in the hope of attracting agents with strong technical skills to do community work. Nevertheless, they were fully aware that the

[7] The history of community health work goes back at least to the early seventeenth century, when a shortage of doctors in Russia led to training community volunteers in providing basic medical care to military personnel. This role later became formalized with China's "barefoot doctors," laypeople who sometimes could not afford shoes but were trained to meet primary health needs in rural areas, and then became widespread in Latin America, in underserved areas in the United States, and, more recently, across Africa (La Familia Sana Program 1992, Pérez and Martinez 2008). The original programs emphasized community self-reliance and participation. Like much of informal public service delivery, for example in the United Kingdom in the eighteenth and nineteenth centuries, these are provided by religious institutions, grassroots movements, and, more recently, nongovernmental organizations. For this reason, however, they are often uncoordinated, lower-skilled efforts.

[8] At the time of the launch of the recruitment process in September 2010, the government had not yet determined how much the health workers would be formally remunerated. Accordingly, the posters did not display any information about compensation. Although the job wage was unknown to applicants at the time of application, applicants would likely have been able to infer an approximate wage, or at least an ordinal wage ranking, based on the "community health" job description and the relatively minimal educational qualifications required: both of which would intuitively place the job below facility-based positions in compensation. In Section IB, we present evidence against the hypothesis that wage perceptions may have differed by treatment.

[9] As of 2019, Ministry of Health records showed that 5 of the original 307 CHAs had applied for higher-level professional training (2 in registered nursing, 2 in public health nursing, and 1 in registered midwifery). Of these 5, 3 were accepted by their training schools and granted study leave by the Ministry, with the understanding that, upon completing their programs, they would cease to work as CHAs and instead assume their new professional roles.

focus on career advancement could backfire by crowding out applicants motivated to help the community.[10] The possibility of this trade-off led to the experiment we describe below.

B. *Experimental Design*

Our experiment aims to assess whether a career in civil service attracts talent at the expense of prosociality, and whether this affects who is hired and their performance. This is relevant to evaluate the role of selection in public service delivery beyond health services in low-income countries, as the concern that material rewards attract the wrong types of people is pervasive. The key challenge is to separate the effect of selection from the effect of incentives on the job. We tackle this in two steps: the first opens the selection channel, and the second shuts down the incentive channel.

Step I: Opening the Selection Channel.—To open the selection channel we use the recruitment posters and the information materials distributed to health officers. In each community, paper advertisements for the job were posted in local public spaces, such as schools, churches, and the health post itself. District health officials were responsible for ensuring that the recruitment posters were posted.

To ensure that the recruitment process was carried out in a uniform manner across all the communities, the government included detailed written instructions in the packets containing the recruitment materials (posters, applications, etc.) that were distributed to district health officials (see online Appendix Section F).

The treatment poster stresses the civil service identity of the new position. It lists as the main benefit of the job the opportunity to ascend the civil-service career ladder to higher and better-paid positions such as environmental health technician, nurse, clinical officer, and doctor. This incentive is summarized in a bold caption stating, "Become a community health worker to gain skills and boost your career!" The poster also explicitly leverages a sense of belonging to the civil service by stating, "Become a highly trained member of Zambia's health care system." Finally it sets "experts in medical fields" as the peer group.

The control poster uses the standard approach of recruiting community health workers, stressing the social identity of the position by making salient community impact such as "[gaining] the skills you need to prevent illness and promote health for your family and neighbors." The message is summarized in a caption stating, "Want to serve your community? Become a community health worker!" Finally, it lists local health post staff as the peer group with whom candidates can expect to interact. Both posters are presented in Figure 1.

Three points are of note. First, the social identity poster functions as control because the status quo community health worker jobs do not offer career opportunities. Second, treatment and control posters have exactly the same structure except the wording of the benefits. We chose this over a "neutral" control poster with no benefits whatsoever because in that case, the treatment effect would conflate the

[10] The Director of Human Resources at the Ministry of Health expressed this trade-off clearly when he asked us: "What is going to happen now that they [potential health workers] will see themselves as civil servants? Will they be connected to the community?"

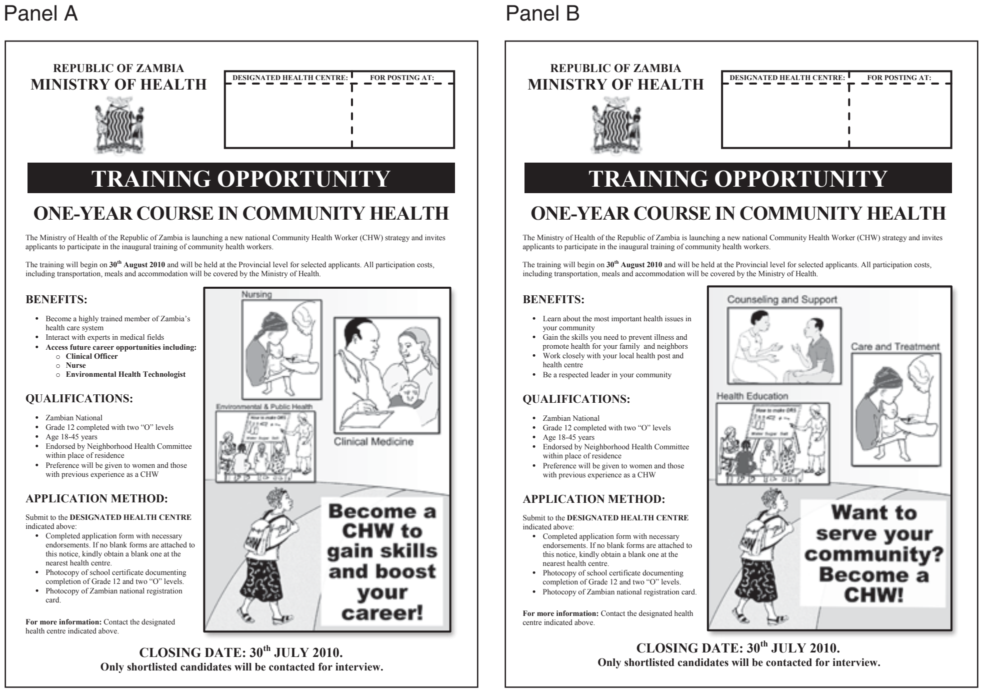

Figure 1. Treatment and Control Posters

Notes: The recruitment posters were created by the authors after consulting with the Ministry of Health of Zambia regarding basic design and content requirements. Content shared in common between the two posters was stipulated by the MOH as required regardless of treatment.

effect of interest with the effect of advertising benefits per se. While this might be of intrinsic interest, it would not allow us to answer the more general question of how agents who are attracted by a career in the civil service differ from those attracted by social impact and how this selection affects performance. Third, it is important to note that in these communities government jobs are scarce; therefore, a poster advertising a government job is likely to be highly visible.

Since recruitment was organized by district officials, we randomized treatment at the district level in order to maximize compliance with the experimental assignment, evenly splitting the 48 districts into two groups. This implies that each district official is only exposed to one treatment. As district officials are the main source of information for aspiring health workers, randomization at the district level minimizes the risk of contamination. Randomization at the district level also mitigates the risk of informational spillovers between communities, as the distance between health posts in different districts is large. Random assignment of the 48 districts is stratified by province and average district-level educational attainment.[11] To ensure compliance with the randomization protocol, we worked

[11] We stratify by the proportion of adults in the district who have a high school diploma, as reported in the most recent Living Conditions Monitoring Survey, conducted by the Central Statistical Office four years prior in 2006. We sort districts by province and, within each province, by high school graduation rate. Within each sorted, province-specific list of districts, we take each successive pair of districts and randomly assign one district in the pair to the career opportunities treatment and the other to the control group. For provinces with an odd number of

closely with the government to standardize the information given to the district officials to organize the recruitment process.[12]

Appendix Table A1 reports balance tests on three sets of variables that can affect the supply of health workers, the demand for their services, and their working conditions. Overall, Table A1 shows that the new health workers are recruited from similar areas and will work in similar areas. Besides showing balance between treatment and control, this exercise is useful to understand labor markets in rural Zambia. Two findings are of note. First, only 4.4 percent of the population have the necessary credentials (grade 12 education) to apply. Second, and more strikingly, just over half of the eligible were either out of work or in unpaid employment over the past 12 months.[13] Among the 45 percent engaged in income-generating activities (either as employees or self-employed), fewer than one-third are employed in high-skill occupations (such as teachers, which account for 9 percent of the eligible population), and about half are employed in low-skill occupations, mostly in agriculture, which accounts for 18 percent of the eligible population. Given the scarcity of skilled jobs, the program might have the added benefit of creating job opportunities in these communities.

Step II: Closing the Incentive Channel.—To close down the incentive channel, all successful applicants are eligible for career opportunities once hired. After being recruited, all agents train together for one year, during which they receive the same information about the career opportunities they are entitled to as civil servants. As treatment and control health workers face the same incentives once hired, performance differences, if any, are attributable to selection.

The experiment aims to create differences in career opportunities at the application stage and then to eliminate these differences after candidates have been hired. To check whether it succeeded, we ask all agents about perceived benefits of the job when they first arrive at the training school in June 2011 and then again 22 months later in April–May 2013, which is after they have completed the one-year training. To elicit this information, we give each health worker a bag of 50 beans and ask them to allocate the beans to different cards describing potential benefits of the job. This method has two desirable features: (i) it forces respondents to take into account the trade-off between different benefits, namely that giving more weight to one benefit necessarily implies that other benefits will be given less weight, and (ii) it allows

districts, we pool the final unpaired districts across provinces, sort by educational attainment, and randomize these districts in the same pair-wise manner.

[12] District officials were given a packet containing 10 recruitment posters and 40 application forms for each health post and were asked to distribute each packet to the respective health center and, from there, ensure that recruitment posters were posted, application forms were made available, and so forth. We conducted a series of follow-up calls over several weeks to the district point-persons to verify that the recruitment process was conducted as planned. To reinforce the treatment, we also included a basic written script that the district officials were invited to use to inform health centers and neighborhood health committees of the health worker program and recruitment process. In the career opportunities treatment, the script described the new program as follows: "This is an opportunity for qualified Zambians to obtain employment and to advance their health careers. Opportunities for training to advance to positions such as Nurse and Clinical Officer may be available in the future." In contrast, in the control group, the script stated, "This is an opportunity for local community members to become trained and serve the health needs of their community" (see online Appendix Section F).

[13] The 29 percent who were out of work are either unemployed (13 percent), housewives (7 percent), or full-time students (9 percent). Most (66 percent) of the unpaid jobs are in agriculture. These are balanced across treatments.

us to test whether the treatment affected other benefits besides career advancement and community service.

There are two sources of potential desirability bias, which might affect the magnitude of the treatment effects but not their sign. First, the fact that respondents say what they think the enumerators want to hear based on the information given on the posters does not invalidate this exercise; the aim of the exercise is precisely to test whether the information they have matches that given on the posters. Second, the fact that this is a community-based position, named "Community Health Worker," might lead the health workers to overstate community benefits. This will bias the share put on community benefits upwards and the difference between treatments downwards, making it less likely for us to be able to detect a difference between treatment and control. This should be kept in mind when interpreting the magnitudes reported below.

The answers tabulated in Appendix Table A2 show that differences in the perceived benefits reported by the health workers when they first arrive at the training school match those advertised in treatment and control posters and then disappear after the health workers are exposed to the training program. Panel A of Table A2 shows that service to the community is listed as the main benefit in both groups. This might truly reflect preferences or be inflated by desirability bias as discussed above. Despite the fact that this biases treatment effects toward 0, we find that the treatment group places 38 percent more weight on career opportunities ($p = 0.002$) and lower weight on both "allows me to serve the community" and "earn respect and status in the community" ($p = 0.046$ and $p = 0.031$, respectively). All other benefits are balanced across groups, suggesting that the poster did not convey different expectations about pay on the job or the nature of the job.

Panel B of Table A2 shows that the answers converge after exposure to training and that there are no significant differences between the two groups. In line with the fact that control health workers receive information about career opportunities during training, the weight they give to career opportunities rises by 27 percent, while the weight they give to service to the community falls by 15 percent. In contrast, treatment health workers, who receive no new information during training, do not change their answers.

The experimental design allows us to identify the effect of career opportunities on performance through selection if the salience of career opportunities at the recruitment stage does not affect the agents' behavior directly once the real career opportunities are known by both treatment and control health workers. This assumption fails if control agents react to the difference between advertised and actual benefits, rather than to the benefits themselves. If control agents value career benefits this will bias the treatment estimates downwards as they might respond to the positive surprise by working harder. Symmetrically, estimates will be biased upwards if control agents dislike career benefits or dislike finding out that the actual value of career opportunities is larger than the value advertised. Note that in this case, agents for whom the participation constraint is met ex ante but not ex post would drop out once hired. For instance, Deserranno (2019) finds that NGO health promoters who receive a negative surprise on earnings are 14 pp more likely to drop out than those who do not over a two-year period. In contrast, the drop-out rate of control CHAs was only 4 percent during the one-year training (the relevant period in which they could update their beliefs about the job), and, as shown in Section IV, once all CHAs are deployed, retention on the job does not vary by treatment.

II. Framework

The experiment can be modeled as a three-stage game. In Stage 1, potential applicants choose whether to apply based on the information conveyed by the posters. In Stage 2, the panels select the CHAs from the applicant pool. In Stage 3, the selected applicants choose how much effort to exert on the job. Our treatment occurs at Stage 0, where we assign different posters to different districts such that the information applicants have in Stage 1 depends on whether their district is in treatment or control.

Applicants differ in ability, $a \geq 0$, and social preferences toward the community, $s \in [0,1]$, or prosociality for short. Below we make precise why both are desirable for successful applicants. General ability a comprises all cognitive (IQ) and noncognitive (ambition, tenacity, work ethic) skills that make individuals productive in all occupations. Prosociality determines the utility individuals get from helping others, in this case by improving their health status. We assume that ability and prosociality are independently distributed in the population.[14] If selected as a CHA, an individual exerts effort, $e \geq 0$, to produce community health according to $H(a,e)$ which is increasing in a and e, generating utility $sH(a,e)$ net of disutility of effort $d(e)$. We assume that output is concave in effort, so $H_{ee} \leq 0$, and also that ability and effort may be complementary, i.e., $H_{ae} \geq 0$, reflecting the idea that a given unit of effort may generate more health output for a highly talented than for a less talented CHA. To ensure an interior solution, we set $d' > 0$ and $d'' > 0$. All selected CHAs also receive material benefits equal to M, which can be thought of as reflecting the discounted sum of future wages; these accrue to all agents regardless of performance. We proceed via backward induction.

A. *Stage 3: Selected Applicants' Choice of Effort*

If hired, the agent chooses e to maximize $U(a,s,e,M) = sH(a,e) - d(e) + M$. By the convexity of d, this yields an interior $e^*(a,s) \geq 0$ that satisfies $de^*/ds > 0$ and $de^*/da \geq 0$. This implies a realized health output function $H^*(a,s) \equiv H(a,e^*(a,s))$ with $H^*_a > 0$ and $H^*_s > 0$. This means that both ability and prosociality contribute to generating health output through different channels: prosociality increases effort, while ability increases output directly for any given level of effort and additionally may increase effort if $H_{ae} > 0$. This is the foundation of the trade-off: if career benefits lead to hiring CHAs with higher a but lower s, the net effect on health output is ambiguous. Finally, substituting realized health output into the utility function gives indirect utility $U^*(a,s,M) \equiv U(a,s,e^*(a,s),M) = sH^*(a,s) - d(e^*(a,s)) + M$, with $U^*_a > 0$, $U^*_s > 0$, and $U^*_M > 0$.

[14] This assumption is made, first, for analytical tractability and, second, to highlight most clearly that self-selection will by itself generate a positive correlation between ability and prosociality at the top of the ability distribution, even if there is none in the general population.

B. *Stage 2: Panel's Choice of Applicants*

In Stage 2, the panel observes applicants with different levels of ability and prosociality. By conducting interviews and using background information (such as test scores), the panel is able to observe these with some noise, such that for individual i the panel observes $\tilde{a}_i = a_i + \epsilon_{ai}$ and $\tilde{s}_i = s_i + \epsilon_{si}$, where ϵ_i is individual-specific noise. For simplicity, we abstract from the exact optimization problem and, as both effort and health output are increasing in a and s, simply assume that the panel follows some rule that picks candidates with high levels of $\tilde{a}$, $\tilde{s}$, or both. Even aside from the direct effect of s on effort and thus health output, we shall see below that ability and prosociality are positively correlated among the applicants with high ability, and thus the panel may also take s into account as a secondary signal to improve prediction of a.

Given the panel's selection rule, whether a particular applicant is hired depends on how her ability and prosociality compare to those of other applicants. As the panel observes these with noise, an individual applicant i can have a nonzero probability of being selected even if they do not have the highest levels of a_i and s_i, and thus we can observe applicants with a range of a_i and s_i in equilibrium. In addition, the probability of selection for any applicant i will also depend on the total number of applicants (as the number of CHA positions is fixed), which is an equilibrium quantity that will depend on M. We thus use a reduced-form probability function that captures the effects of a_i, s_i, and M, which we write as $p(a_i, s_i, M)$.[15] The panel's decision rule means that p is increasing and concave in both a and s, and we shall see below that p is decreasing in M.[16]

C. *Stage 1: Applicant's Decision to Apply*

In Stage 1, applicants decide whether to apply. If hired, an individual with traits (a_i, s_i) will receive utility in Stage 3 of $U^*(a_i, s_i, M)$. To apply, an individual needs to pay cost c. Taking into account the probability of selection in Stage 2 and the utility in Stage 3 if selected, individual i will apply if the expected utility of applying to be a CHA, net of application costs, exceeds the utility in their next best alternative occupation, which, in this setting, is mostly self-employment in agriculture or small trade where the agent is the residual claimant. We denote this by $V(a)$ and assume, as is standard, that the marginal return to ability is higher in the private sector $V_a > U_a^* > 0$ for every a. This is the empirically relevant case because, as is common in the public sector, CHAs' earnings are not linked to performance, while self-employed agents in the private sector are the residual claimants on the value they create.[17] Thus,

[15] Given our specification of the panel's decision rule, we use a reduced-form function rather than a genuine "equilibrium" function, which would arise from computing the fixed point from the problem of applicants applying, which in turn affects the probability, which in turn affects applications again, and so on. The reduced-form approach is considerably simpler analytically, and we also observe empirically that the panel's selection process does not differ between treatment and control, which seems more in line with panels using a sensible rule-of-thumb approach than strategic behavior that differs between treatment groups.

[16] We assume continuity of this and related quantities (i.e., that there are sufficiently many applicants) for analytical simplicity, although this is clearly not literally true.

[17] As an indication of the returns to ability, we note that Young (2012) runs Mincerian regressions for 24 countries (including Zambia), of which 14 are in sub-Saharan Africa, using a range of methods that yields estimates

individual i applies if and only if $E(a_i, s_i, M) = p(a_i, s_i, M)\, U^*(a_i, s_i, M) - c > V(a_i)$. To capture the fact that in practice there are minimum qualification requirements, we assume that application costs are high enough that $E(0, s, M) < V(0)$ for any s and M, so that low-ability individuals who have little chance of being hired do not apply. Formally, there is a threshold of ability $\underline{a}_i(s_i, M)$ such that all i with $a_i < \underline{a}_i(s_i, M)$ do not apply. We shall suppress the subscript on the threshold for notational simplicity, but it is important to note that it differs across individuals as it depends on their level of s_i. Online Appendix Section B shows that the structure of the solution depends on whether $E(a_i, s_i, M) > V(a_i)$ for all $a_i > \underline{a}$. If so, everybody with $a_i > \underline{a}$ will apply. If not, there is a further threshold defined by $E(\overline{a}_i, s_i, M) = V(\overline{a}_i)$ such that only i with $\underline{a} < a_i < \overline{a}$ apply.

D. *Stage 0: Treatment*

The intervention takes place at Stage 0 when treatment districts receive the career posters and control districts do not. We model treatment as $M_T > M_C$, representing the fact that the career emphasis suggests a higher expected net present value of lifetime material benefits despite the starting wage being the same. Our goal is to make precise the conditions under which treatment creates a trade-off by attracting applicants with lower prosociality and higher ability.

Treatment increases U^* through several channels. Of these, the increase in the net present value of exogenous material benefits M is the one that makes the job relatively more attractive to individuals with low prosociality.[18] This may occur through automatic salary progression (individuals are entitled to all increases negotiated collectively for government employees) or through promotion to higher-paying positions.

To create a trade-off between ability and prosociality, the increase in M must attract higher-ability applicants. This happens through the upper threshold of ability, $\overline{a}(s_i, M)$. In particular, as Result 1 shows, the threshold $\overline{a}(s_i, M)$ is increasing in M. This also creates an additional effect in equilibrium, as the fact that more (high-ability) candidates apply reduces the probability of being selected for all applicants. This means that $dp/dM < 0$. This effect is weaker for candidates with high levels of a, that is $d^2p/dMda > 0$, because they are highly likely to be selected. The fact that M affects the probability of being selected as described above creates two offsetting effects on the lower threshold. On the one hand, higher material benefits increase U^* directly and thus lower $\underline{a}(s, M)$ for any given s; on the other

of returns to education of between 9.5 percent and 11.6 percent. His preferred estimate, used in the remainder of his paper, is 11.6 percent. The estimates are qualitatively unchanged when restricting the sample to Zambia only (results available upon request).

[18] Treatment can affect U^* in several ways that do not create trade-offs, for instance by increasing the marginal product of ability. This is due to career benefits giving high-ability individuals the chance to be promoted to higher-ranked positions where they can benefit more people or have more influence on key decisions. This appeals to high a individuals who can benefit from it and high s individuals who care about it. In this respect, treatment improves the quality of the applicant pool in all dimensions without creating a trade-off. In a related manner, treatment may also have a relationship with prosociality, since individuals who are prosocial may recognize that they can have a greater (lifetime) contribution through acquiring skills and ascending the career ladder. Since this effect would be stronger for those with higher ability, this would further result in the high-ability applicants attracted by treatment also being highly prosocial.

hand, the reduction in probability arising from more high-ability applicants discourages applying and thus raises $\underline{a}(s, M)$. Combining these effects yields the following.

Result 1: Increasing material benefits M attracts higher-ability applicants who would not apply otherwise ($\partial \overline{a}/\partial M > 0$) and either (i) lowers the ability of the lowest-ranked applicant ($\partial \underline{a}/\partial M < 0$) and increases the total number of applicants or (ii) discourages low-ability applicants ($\partial \underline{a}/\partial M > 0$) and has an ambiguous impact on the total number.[19]

To assess the effect on prosociality we note that U^* is increasing in both a and s. The threshold $\overline{a}(s_i, M)$ is increasing in s because, due to the fact that $V_a > U_a^* > 0$, higher-ability individuals need to have a high level of prosociality to meet their participation constraint. Symmetrically, the threshold $\underline{a}(s_i, M)$ is decreasing in s because more prosocial applicants have a higher expected payoff for the same probability of being chosen, thus the probability, and hence the level of ability, that makes them indifferent between applying and not is lower. This implies that talent and prosociality are positively (negatively) correlated among the highest- (lowest-) ability applicants, even though they are not correlated in the population. Therefore, we have the following.

Result 2: Under any M, the most able applicant is also the most prosocial. An increase in M leaves the prosociality of the marginal applicant unchanged and has an ambiguous effect on the prosociality of the average applicant.

Figure 2 illustrates treatment effects on ability and prosociality. There are many possible ways of modeling these functions. To demonstrate the intuition of the trade-off as simply as possible, we make the starkest possible assumption and set $H_{ae} = 0$, meaning that prosociality increases the marginal return to effort, while ability increases output directly and does not affect effort. Thus, all else equal, more prosocial people will work harder, while more able people work equally hard but produce more output directly. Note, however, that allowing for the complementary term would add an additional channel for ability, reinforcing the treatment effect in a multiplicative manner and thus allowing small changes in ability to lead to large changes in outputs.

We consider a simple example where $H = 2\sqrt{\xi}\, e + a$, where $\xi > 0$ is a parameter, and $d(e) = e^2$, which implies $e^* = \sqrt{\xi}\, s$, $H^* = 2\xi s + a$, and $U^* = \xi s^2 + as + M$. The probability of selection must satisfy the following: (i) increasing in a and s, (ii) decreasing in M, and (iii) $\partial^2 p/\partial a \partial M > 0$. For the illustration we use the following functional form $p(a, s, M) = (\gamma a s^{\beta} + \mu(M_T - M))/(1 + \mu(M_T - M))$, where $\gamma > 0$ is set to guarantee that $p \in [0, 1]$,[20] $\beta > 0$ captures the weight that the panel puts on prosociality, and $\mu > 0$ scales the decrease in probability that arises from

[19] Note that it is possible to simultaneously have a decrease in the number of applicants and have low-ability applicants being discouraged because the additional high-ability applicants have a much larger effect on the selection probability than the reduction in low-ability applicants.

[20] For ease of interpretation, note that $p = 1$ when $\gamma a s^{\beta} = 1$.

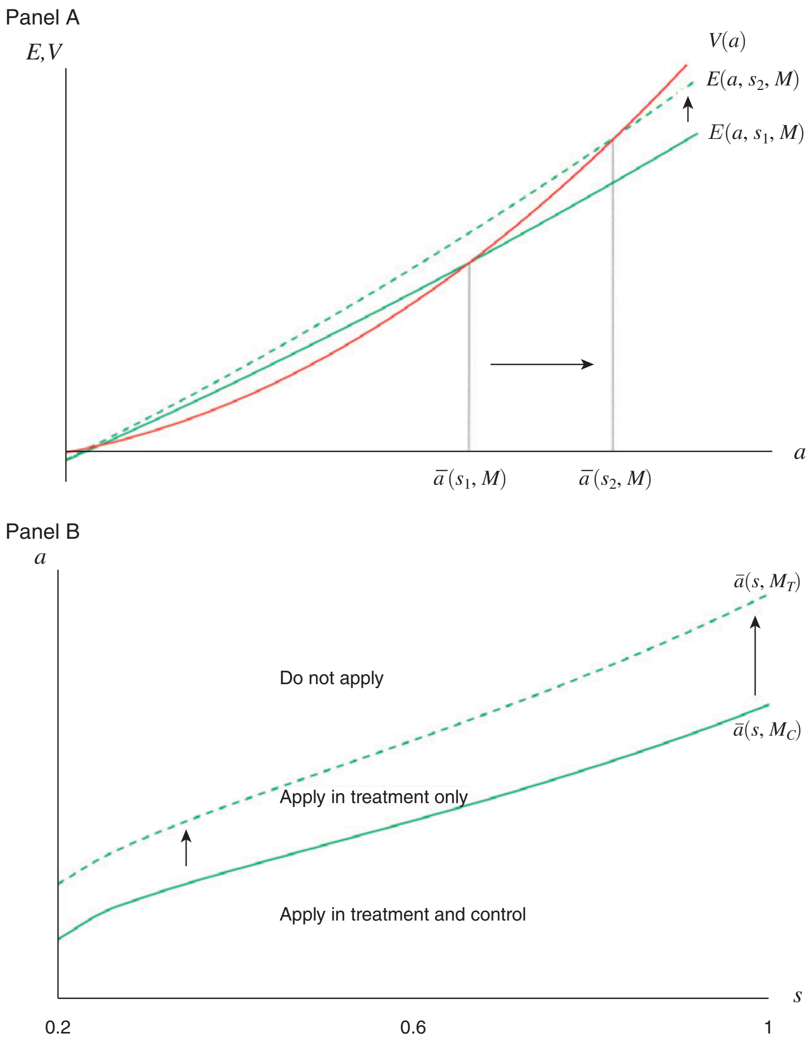

Figure 2. Treatment Effects on the Applicant Pool, in Theory

Notes: Panel A presents the expected utility to applying against the payoff of the outside option. The intersections represent the thresholds of ability at which agents are indifferent between applying and not. The upper threshold is increasing in s while the lower threshold is decreasing in s. Panel B presents the upper threshold of ability, at which agents are indifferent between applying and not, as a function of s. Higher monetary rewards shift the frontier, with a larger positive effect at higher levels of s as the negative equilibrium probability effect decreases with s. Note that applicants with ability levels never attracted in control but sometimes attracted in treatment (i.e., in the northeast of the region between the two lines) all have high s. The functional forms and equations are provided in online Appendix Section B.

more (high-ability) people applying in equilibrium when M rises.[21] Note that $dp/dM = -\mu(1-p)/1+\mu(M_T - M) \rightarrow 0$ as $p(a,s,M) \rightarrow 1$, which says that the increased applicant effect is 0 for candidates who are certain to be selected. We set $V = va^2 + a$, where we assume $v > \gamma$ which ensures that $V_a > U_a^* > 0$ everywhere and that $V_a > E_a > 0$ for all a sufficiently large. In panel A of Figure 2,

[21] In principle, some $\overline{M} > M_T$ would be more suitable in the function for general usage, but for our purposes M_T is sufficient and helps to avoid unnecessary proliferation of parameters.

we see that, at low levels of a, the cost of applying is too high; for sufficiently high levels of a, the increased return to ability in the private sector begins to dominate, thus defining $\overline{a}(s,M)$. Panel B illustrates the application frontier, that is all the combinations of a and s such that an individual is indifferent between applying and not. The frontier is positively sloped, so that the ability threshold is higher for more prosocial applicants and all individuals below the frontier will apply. Note that the main results are robust to many functional form assumptions so long as they generate $\overline{a}(s,M)$ as upward sloping. The dashed line shows that an increase in M shifts the frontier upward. The effect on average prosociality depends on the balance of two forces. First, the average prosociality of individuals whose a is sufficiently high that they would never apply without career benefits is higher than the average prosociality of those who apply without. In panel B, these individuals are in the northeast section of the region of applicants who apply in treatment but not control.[22] Intuitively, candidates with very high ability who get attracted by treatment must have very high prosociality in order for them to prefer the CHA role to the private sector. Second, the average prosociality of individuals whose a is low enough that they may still apply without career benefits is lower when these are offered. This is because U^* is increasing in M; thus for any a the level of s that makes individuals indifferent between applying or not falls. In panel B, we see this through the fact that the dashed line is always to the left of the thick line. This is the standard substitution or crowding out effect.

Given that the effect on the inframarginal and marginal applicants differ, the treatment effect on hired CHAs depends on the selection mechanism. Mechanisms that pick the highest-ability candidates from the applicant pool will produce the largest possible positive difference in ability and no difference in prosociality because the most able applicants within each pool are also the most prosocial. Mechanisms that pick randomly from the pool will still produce a positive difference in ability and a negative difference in prosociality if the average applicant under career incentives is less prosocial. Thus, to understand how differences in the applicant pool translate into differences among hired CHAs we need to understand the selection mechanism. This is the aim of the next section.

III. Treatment Effect on the Applicant Pool and Selected Candidates

A. *Treatment Effect on the Applicant Pool*

The recruitment drive yielded 2,457 applications, an average of 7.4 applicants for each position. Overall, 1,804 (73.4 percent) applicants met the eligibility requirements and were invited for interviews;[23] of these, 1,585 (87.9 percent), or 4.7 per position, reported on their interview day when we administered a questionnaire to collect information on skills, career ambition, and prosociality. These 1,585 form the applicant pool we analyze in this section.

[22] In particular, they have a such that $\underline{a}(1,M_C) < a < \overline{a}(1,M_T)$.

[23] All completed application forms were taken to the district Ministry of Health office where district health officials checked that requirements were met. No discretion was given at this stage; applicants who did not meet the objective criteria were rejected, and those who did were invited for interviews.

To measure treatment effects on the composition of the applicant pool, we collect measures of ability and prosociality at the application stage for the universe of applicants who were interviewed. To measure cognitive skills we use grade 12 final exam scores and the number of courses taken in biology and other natural sciences. These are the skills measures used as application requirements. For noncognitive skills we focus on career ambition, measured by asking applicants the job they envisage doing in five years' time, and code as career-motivated those who aim to a higher-ranked position in the Ministry. To measure prosociality we combine the applicant's self-reported willingness to stay in the community in the long term together with the "Inclusion of Others in Self (IOS)" scale that measures alignment of interests (Aron et al. 2004).[24] We do so to identify those who want to stay in the community because they care about community outcomes, as opposed to those who stay for other reasons. Guided by the framework, we estimate the effect of treatment on skills and prosociality, both on the average applicant and as a function of skill rank. Table 1 shows that treatment attracts individuals with higher cognitive skills and career motivation. The average effects are about one-fifth of a standard deviation in the control group and are all precisely estimated at the 95 percent confidence level or above. Both here, and throughout the paper, we also report *p*-values based on the effective degrees of freedom (EDF) correction procedure in Young (2016) and a randomization inference procedure (Young 2019).[25] Average prosociality is lower, albeit not precisely estimated, while age, gender, and current occupation are very similar. The latter is due to the fact that, in line with the evidence from the Census in Table A1, there is hardly any variation. Most applicants (70 percent) are farmers, a further 8 percent are housewives, 6 percent are traders, and 5 percent are teachers. Finally, the number of applicants per health post is not significantly different. Result 1 makes precise that this can happen if treatment increases both ability thresholds. To test this, panels A and B of Figure 3 report the kernel density estimate as well as the quantile treatment effects on total test scores. Both reveal a rightward shift, namely, all applicants in treatment, from the lowest- to the highest-ranked, have higher test scores.

Panels C and D report the mean levels of ability and prosociality in treatment and control across different levels of the applicant's skill rank in his or her health post. Panel C shows that the treatment group's exam scores are higher at every rank whilst panel D shows that the difference in prosociality is zero for top-ranked applicants and negative for lower-ranked applicants. Both panels C and D are thus consistent with the theory, although standard errors clustered by district are not precise enough to reject the null of equality.

The results in Figure 3 are in line with the simple theoretical framework: ability increases throughout whereas the effect on prosociality is zero for top-ranked candidates and negative for lower-ranked candidates. The figure makes clear that

[24] IOS measures the extent to which individuals perceive community and self-interest as overlapping. Applicants are asked to choose between four pictures, each showing two circles (labeled "self" and "community") with varying degrees of overlap, from non-overlapping to almost completely overlapping. This variable equals 1 if the respondent chooses the almost completely overlapping picture, 0 otherwise. IOS has been validated across a wide variety of contexts, and adapted versions are found to be strongly correlated with environmental behavior (Schultz 2012) and connectedness to the community (Mashek, Cannaday, and Tangney 2007).

[25] Specifically, we use the randomized randomization-*t* *p*-value, computed using *randcmd* in Stata with 10,000 iterations.

Table 1—The Effect of Career Opportunities on the Applicant Pool

	Treatment	Control	p-value (clustered)	p-value (EDF)	p-value (RI)
Applicants per health post	8.7	10.0	0.611	0.641	0.666
	(5.7)	(6.5)			
Cognitive skills (O-levels total exam score)	24.8	23.3	0.019	0.040	0.002
	(9.8)	(9.3)			
Cognitive skills (number of science O-levels)	1.44	1.24	0.006	0.017	0.024
	(0.86)	(0.89)			
Career motivation	0.25	0.19	0.026	0.050	0.050
	(0.43)	(0.39)			
Prosociality	2.34	2.51	0.237	0.295	0.416
	(0.79)	(0.64)			
Farmer	0.715	0.684	0.415	0.470	0.349
	(0.45)	(0.47)			
Age	25.7	26.0	0.433	0.487	0.487
	(5.5)	(5.8)			
Female	0.292	0.304	0.800	0.822	0.819
	(0.45)	(0.46)			

Notes: Columns 1 and 2 show means and standard deviations in parentheses. Sample includes the 1,585 candidates who were interviewed for the position. *Treatment* $=1$ if the candidate is interviewed in a district where career opportunities were made salient. Column 3 reports the p-values of the null hypothesis that the career treatment effect equals 0 conditional on stratification variables and with standard errors clustered at the district level. Column 4 reports the p-value using the Young (2016) effective degrees of freedom (EDF) correction, clustered at the district level. Column 5 reports the p-value using a randomization inference (RI) procedure, clustered at the district level (specifically, the randomization-t p-value from Young 2019). Ordinary levels or O-levels are administered by the Examinations Council of Zambia (ECZ) to twelfth-grade students, the highest grade in the Zambian secondary education system. O-levels total exam score is constructed as the sum of inverted O-levels scores ($1=9$, $2=8$, and so on) from all subjects in which the applicant wrote the exam, so that larger values correspond to better performance. O-levels passed in biology and other natural sciences equals the number of O-levels passed in biology, chemistry, physics, science, and agricultural science. *Career motivation* $=1$ if the candidate chooses any combination of being an "environmental health technician," "clinical officer," or "doctor" in response to the question, "When you envision yourself in 5–10 years' time, what do you envision yourself doing?" *Prosociality* is the average of "Do you see yourself in the community in 5–10 years" (yes/no) and the Inclusion of Others in Self scale (Aron et al. 2004). Applicants are asked to choose between sets of pictures, each showing two circles (labeled "self" and "community") with varying degrees of overlap, from non-overlapping to almost completely overlapping. *Farmer* $=1$ if the applicant's main occupation is self-employment or work in the family farm in agriculture. *Age* is in years. *Female* $=1$ if the applicant is female.

the effect of treatment on CHAs themselves will depend on how these are chosen among the applicants. We analyze this next.

B. *The Selection Mechanism and Treatment Effect on Selected Candidates*

Selection panels are in charge of choosing the two candidates that will serve as CHAs in the health post. Panels have five members: the district health official, a representative from the health post's associated health center, and three members of the local neighborhood health committee. Each panel was asked to nominate two top candidates and up to three reserves. The government explicitly stated a preference for women and for those who had previously worked as community health workers, but the ultimate choice was left to the panels. Overall, selection panels nominated 334 applicants as "top 2" candidates and 413 as reserves.[26]

[26] The nominations were reviewed centrally by the government of Zambia, and 334 final candidates were invited to join a yearlong training course. Of these, 314 applicants accepted the invitation and, in June 2011, moved to the

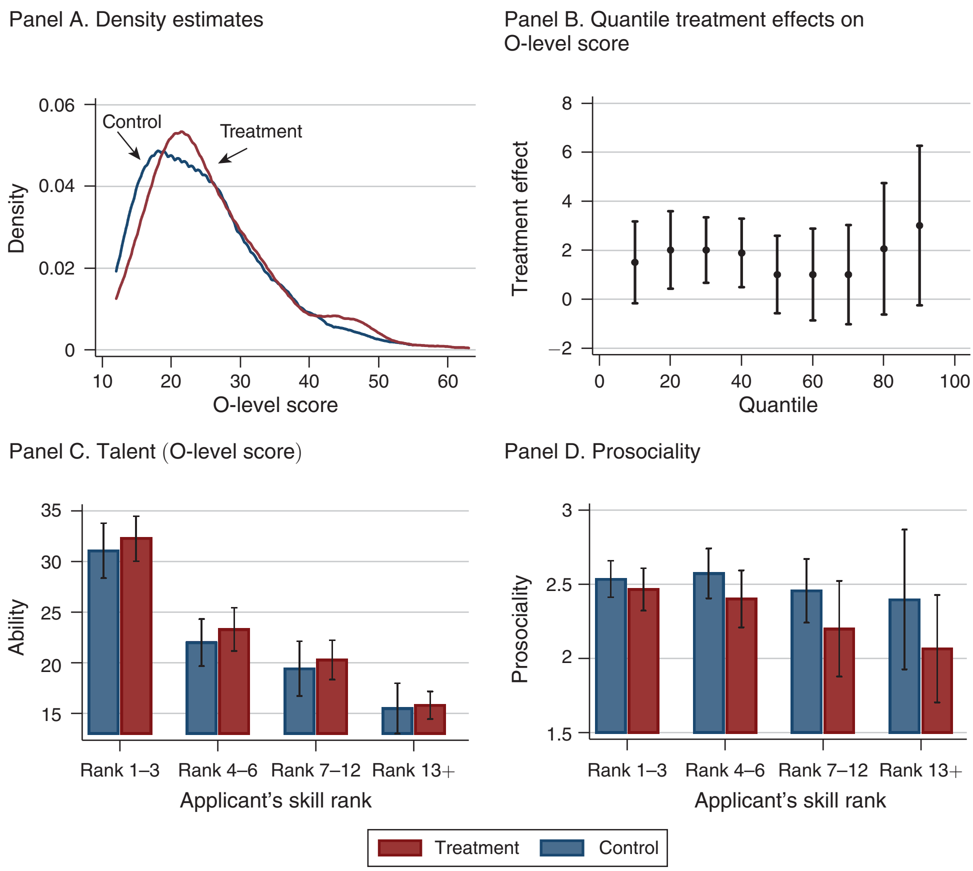

Figure 3. Treatment Effects on the Applicant Pool

Notes: Panel A reports kernel density estimates of exam scores in treatment and control. Panel B reports quantile treatment effect estimates on exam scores and 95 percent bootstrapped confidence intervals. Panels C and D present the mean levels of ability and prosociality in treatment and control across different levels of the applicant's skill rank in his or her health post. The confidence intervals are at the 95 percent level, based on standard errors clustered at the district level. Applicants are the 1,585 candidates who were interviewed for the position. *Treatment* = 1 if the health worker is recruited in a district where career opportunities were made salient. Ordinary levels or O-levels are administered by the Examinations Council of Zambia (ECZ) to twelfth-grade students, the highest grade in the Zambian secondary education system. O-levels total exam score is constructed as the sum of inverted O-levels scores (1 = 9, 2 = 8, and so on) from all subjects in which the applicant wrote the exam, so that larger values correspond to better performance. O-levels passed in biology and other natural sciences equals the number of O-levels passed in biology, chemistry, physics, science, and agricultural science. *Prosociality* is the average of "Do you see yourself in the community in 5–10 years" (yes/no) and the Inclusion of Others in Self scale (Aron et al. 2004). Applicants are asked to choose between sets of pictures, each showing two circles (labeled "self" and "community") with varying degrees of overlap, from non-overlapping to almost completely overlapping.

To understand how differences in the applicant pool translate into differences in hired CHAs we analyze how panels select candidates. This analysis also sheds light on whether treatment affects panels' choices and on which traits panel members

training school in Ndola, Zambia's second-largest city. Of the applicants who joined the program, 307 graduated and started working in August 2012. All the health workers were deployed back to their communities of origin.

TABLE 2—THE EFFECT OF CAREER OPPORTUNITIES ON CANDIDATE SELECTION BY PANELS

	= 1 if selected	*p*-value of difference	= 1 if selected	*p*-value of difference
= 1 if top 3 in skills × treatment	0.191	0.43	0.214	0.40
	(0.032)		(0.040)	
= 1 if top 3 in skills × control	0.153		0.170	
	(0.037)		(0.038)	
= 1 if top 3 prosociality × treatment	0.025	0.21	0.034	0.37
	(0.030)		(0.036)	
= 1 if top 3 prosociality × control	0.066		0.071	
	(0.027)		(0.030)	
= 1 if aims to higher rank × treatment	0.111	0.33	0.104	0.42
	(0.044)		(0.040)	
= 1 if aims to higher rank × control	0.063		0.062	
	(0.031)		(0.034)	
= 1 if connected to village leader × treatment			0.029	0.82
			(0.037)	
= 1 if connected to village leader × control			0.019	
			(0.029)	
= 1 if connected to health center staff × treatment			−0.032	0.62
			(0.066)	
= 1 if connected to health center staff × control			0.005	
			(0.044)	
Adjusted R^2	0.111		0.115	
Observations	1,468		1,242	

Notes: OLS estimates. *Treatment* = 1 if the health worker is recruited in a district where career opportunities were made salient. All regressions include the stratification variables (province dummies and share of high school graduates in the district) and standard errors clustered at the district level, using the Young (2016) effective degrees of freedom correction. Independent variables are interacted with the treatment and control dummies. Sample is the interviewed applicants as these are the ones that the panel saw and selected from. *Top 3 skills* = 1 if the applicant's exam score is one of the three highest among interviewed applicants to the same health post. *Aims to be a higher-rank health professional in 5–10 years* = 1 if the candidate chooses any combination of being an "environmental health technician," "clinical officer," or "doctor" in response to the question, "When you envision yourself in 5–10 years' time, what do you envision yourself doing?" *Prosociality* is the average of "Do you see yourself in the community in 5–10 years" (yes/no) and the Inclusion of Others in Self scale (Aron et al. 2004). Applicants are asked to choose between sets of pictures, each showing two circles (labeled "self" and "community") with varying degrees of overlap, from non-overlapping to almost completely overlapping. *Connected to village leader* (*health center staff*) = 1 if the candidate chooses "political leader" or "village committee member" ("formally trained health worker") in response to the question, "Are any of your relatives or members of your household in the following positions?"

deem important for the job. Table 2 estimates the probability that candidate i in health post h is chosen as follows:

$$s_{ih} = \sum_{j \in J} \alpha_j^T C_h X_i^j + \sum_{j \in J} \alpha_j^C (1 - C_h) X_i^j + \gamma N_h + \zeta_{ih},$$

where $s_{ih} = 1$ if i is one of the two nominated candidates and 0 otherwise. $C_h = 1$ if health post h is in treatment and 0 if it is in the control group. X_i^j are indicator variables that equal 1 if candidate i is in the top three of trait j, and the core set J includes skills, ambition, and prosociality. We also report regressions with an expanded set that includes social connections to local political leaders and health facility staff to test whether connections help with getting the job when material benefits are higher. To control for the strength of competition, we include the number of interviewed candidates in the same health post N_h. We control for the stratification variables and cluster standard errors at the district level, correcting for effective degrees of freedom using the procedure in Young (2016).

The coefficients of interest are α_j^T and α_j^C, which measure the weight given to trait *j* in the treatment and control groups, respectively. We test the null that panels use the same criteria in both groups, that is $\alpha_j^T = \alpha_j^C$. Panels are exposed to treatment as they see the posters, but in contrast to candidates, for whom the poster is the only source of information, panel members know the job attributes and who would be suitable for it. The two more senior panel members (the district health official and the health center representative) are employees of the Ministry of Health, and hence are familiar with career progression rules regardless of treatment. Thus this is likely not as powerful, or perhaps entirely moot.[27] Table 2 reports the estimates of α_j^T and α_j^C for all $j \in J$ and the *p*-value of the test of equality.

Column 1 in Table 2 shows that panels put a strong positive weight on skills and prosociality and do so equally in both treatment and control groups. The average probability of being nominated for an applicant who does not rank at the top of the skills and prosociality distributions and who has no career ambition is 0.09. This increases by 15 to 19pp for applicants at the top of the skill distribution, by 6 to 11pp for applicants with career ambitions, and by 3 to 7pp for applicants with high prosociality. The tests of equality of coefficients between treatment and control do not reject the null for any of these traits. Column 3 additionally shows that connections either to political leaders or to staff at the health facility do not affect the probability of selection in either treatment or control.

Taken together, the evidence suggests that career opportunities attract applicants who have different skills, career motivation, and prosociality and that all panels deem these traits to be valuable and are more likely to choose applicants who rank highly in all three.

Our conceptual framework makes clear that, compared to a random selection mechanism, this type of selection leads to higher skill differences and eliminates prosociality differences. To illustrate, we compare the traits of the CHAs selected by panels to 1,000 random draws of two CHAs from each health post's applicant pool. Table 3 reports the average trait for panel-selected CHAs and the tenth, fiftieth, and ninetieth percentiles of the same traits when candidates are chosen randomly. The table shows that the distribution of skills and motivation in treatment is to the right of control, and that panels choose from the top in both groups. This implies that, compared to random selection, panels select CHAs who have higher exam scores and career motivation. Indeed, panel-selected CHAs score higher than the ninetieth percentile of randomly selected CHAs on all three measures. Despite the fact that panels put the same weight on talent and career motivation in treatment and control, the average skill and career motivation of selected CHAs is higher in the treatment group because treatment attracts candidates who do not apply in control.

In contrast, whereas randomly selected CHAs have lower prosociality in the treatment group, panel selection undoes this difference because, as shown in panel D of Figure 3, the most talented applicants in each pool have the same level of prosociality, and panels select these. The fact that treatment creates a trade-off

[27] Further analysis, available upon request, shows that treatment does not affect panel composition.

Table 3—Panel Selection versus Random Selection

	Treatment	Control
Cognitive skills (*O-levels total exam score*)		
Panel selection	27.2	25.6
Random draw: median	24.9	23.1
Random draw: 10th and 90th percentile	[24.1; 25.8]	[22.2; 23.9]
Standard deviation (main sample)	(9.8)	(9.3)
Cognitive skills (*number of science O-levels*)		
Panel selection	1.55	1.46
Random draw: median	1.43	1.26
Random draw: 10th and 90th percentile	[1.35; 1.49]	[1.18; 1.33]
Standard deviation (main sample)	(0.86)	(0.89)
Career motivation (= 1 *if aims to higher rank in* 5–10 *years*)		
Panel selection	0.36	0.25
Random draw: median	0.27	0.19
Random draw: 10th and 90th percentile	[0.24; 0.31]	[0.16; 0.23]
Standard deviation (main sample)	(0.43)	(0.39)
Prosociality		
Panel selection	2.55	2.55
Random draw: median	2.46	2.54
Random draw: 10th and 90th percentile	[2.41; 2.51]	[2.49; 2.58]
Standard deviation (main sample)	(0.79)	(0.64)

Notes: Sample includes the 1,585 candidates who were interviewed for the position. *Treatment* = 1 if the candidate is interviewed in a district where career opportunities were made salient. Panel selection reports the average trait of the two CHAs chosen by the panels in each health post. Random selection reports the average trait of two CHAs chosen randomly over 1,000 draws. Ordinary levels or O-levels are administered by the Examinations Council of Zambia (ECZ) to twelfth-grade students, the highest grade in the Zambian secondary education system. O-levels total exam score is constructed as the sum of inverted O-levels scores (1 = 9, 2 = 8, and so on) from all subjects in which the applicant wrote the exam, so that larger values correspond to better performance. O-levels passed in biology and other natural sciences equals the number of O-levels passed in biology, chemistry, physics, science, and agricultural science. *Career motivation* = 1 if the candidate chooses any combination of being an "environmental health technician," "clinical officer," or "doctor" in response to the question, "When you envision yourself in 5–10 years' time, what do you envision yourself doing?" *Prosociality* is the average of "Do you see yourself in the community in 5–10 years" (yes/no) and the Inclusion of Others in Self scale (Aron et al. 2004). Applicants are asked to choose between sets of pictures, each showing two circles (labeled "self" and "community") with varying degrees of overlap, from non-overlapping to almost completely overlapping.

between ability and prosociality for low-ability applicants is of no consequence because these are not hired.

IV. Inputs in Service Delivery

A. *Measuring Inputs in Service Delivery*

The CHAs' main job task, to which they are required to devote 80 percent of their time, or 4 out of 5 days per week, is to visit households. The input part of our analysis focuses on the number of visits completed over the course of 18 months, from August 2012 (when the health workers started work) until January 2014. The number of household visits is akin to an attendance measure for teachers or nurses: the health workers are supposed to work in people's houses, and we measure how often they are there. Naturally, differences in the number of visits can be compensated for with differences in other inputs; we discuss this possibility in Section IVC after establishing the main results. Furthermore, differences in inputs ultimately are of interest only if they lead to better outcomes, which we will discuss in Section V.

Our primary measure of household visits is built by aggregating information on each visit from individual receipts. All the health workers are required to carry receipt books and issue each household a receipt for each visit, which the households are asked to sign. The health workers are required to keep the book with the copies of the receipts to send to the government when completed. They are also required to send all information on these receipts, consisting of the date, start time, and end time of the visit, as well as the client's phone number, via text message to the Ministry of Health. These text messages are collected in a central data-processing facility, which we manage.

Since visits are measured by aggregating text messages sent by the health workers themselves, identification can be compromised by the presence of measurement error that is correlated with treatment. For instance, health workers in the career treatment might put more effort in reporting visits via text messages or might report visits that never took place, leading to a positive bias in the estimated treatment effect.

We validate our visits measure by comparing it to administrative data and households' own reports of health worker activity. The administrative data is drawn from the Health Management and Information System (HMIS), which is the Ministry of Health's system for collecting routine health services data at government facilities. These are reported at the end of each month and sent electronically to the Ministry via a mobile platform, jointly by the two CHAs and the other staff working in each health post. As HMIS data are only available aggregated at the health post level (summed over the two workers in each health post) we regress these on our visit measure, also aggregated at the health post level. Columns 1 and 2 in Appendix Table A3 show that the two measures are strongly correlated ($r = 0.767$). The correlation in this measure is higher in the treatment group than the control group (0.836 versus 0.644), although the difference is not significant, which contradicts the differential reporting hypothesis.

The households' reports are collected via a survey that we administered to 16 randomly chosen households in each of 47 randomly selected communities chosen from the set of communities where the CHAs operate, stratified by district. We ask respondents whether they know each of the health workers (97 percent do), whether they have ever been visited (44 percent of them have), and their level of satisfaction with each health worker. Columns 3 and 4 show a precisely estimated correlation between our visit measure and the probability that a household reports a visit. This correlation is slightly higher in the treatment group than the control group, although the difference is not significant. Columns 5 and 6 show a precisely estimated correlation between our visit measure and the household's reported level of satisfaction with the health worker's performance. The difference between treatment and control is small and insignificant, casting doubt on the relevance of differential reporting.

Taken together, the findings in Table A3 generally validate our visits measure. Ultimately, however, we will not be able to detect a treatment effect on households' health outcomes in Section V if measured differences in visits capture differences in reporting rather than in actual visits.

B. *Treatment Effect on Household Visits*

Table 4 reports the reduced-form effects of treatment on performance: that is, the estimates of

$$v_{ihdp} = \alpha + \beta C_{id} + Z_h \gamma + \delta E_d + \rho_p + \epsilon_{ihdp}, \tag{1}$$

where v_{ihdp} is the number of visits completed by health worker i in catchment area h, district d, and province p. $C_{id} = 1$ if agent i is recruited and operates in a district assigned to the career opportunities treatment. Z_h is a vector of area characteristics, which includes the number of staff at the health post, cell network coverage, and the distribution of households between farms and villages described in Table A1, although we note that the results are qualitatively unchanged if we remove these. We control for the stratification variables, district-level high school graduation rate E_d and province indicators ρ_p, throughout. Standard errors are clustered at the level of randomization, the district, and, as mentioned before, we also report p-values from the effective degrees of freedom correction in Young (2016) and a randomization inference procedure (Young 2019).

The coefficient of interest is β, which measures the effect of making career opportunities salient at the recruitment stage on the number of visits completed over 18 months. Considering that all the health workers are given the same information on career opportunities during the year-long training, β captures the effect of career opportunities on performance through selection.

The causal effect of career opportunities on performance can be identified under the assumptions that (i) C_{id} is orthogonal to ϵ_{ihdp}, and (ii) there are no spillovers between the two groups. Orthogonality is obtained via random assignment. Spillovers via movements of health workers between treatment and control areas are ruled out by the program requirement that health workers must have been residing in the community they want to work in prior to applying. This implies that career opportunities cannot draw in talent from control areas. Spillovers of information, caused for example by potential applicants in control seeing the treatment poster, would introduce a downward bias because they would reduce the information differences between treatment and control. Information spillovers are minimized by design, as recruitment messages were randomized at the district level, which, given the travel distance between rural communities in different districts, makes it very unlikely that applicants in one group might have seen the poster assigned to the other group. Importantly, information cannot accidentally spill over through the district officials that implement the program or through the recruitment panels, as these are only exposed to one treatment.

Column 1 of Table 4 reveals a large and precisely estimated effect of career opportunities on household visits: health workers recruited by making career opportunities salient do 94 more visits (29 percent more than control) over the course of 18 months. The median treatment effect is 104.4 (bootstrapped SE 43.9), which allays the concern that the average effect is driven by outliers. The magnitude of the difference is economically meaningful: if each of the 147 health workers in control had done as many visits as their counterparts in the career treatment, 13,818 more household visits would have been performed over the 18-month period. Given that for most of these households, health workers are the only providers of health

Table 4—The Effect of Career Opportunities on the Number of Visits

Dependent variable:	Household visits			
Time horizon:	Months 1–18 (1)	Months 1–6 (2)	Months 7–12 (3)	Months 13–18 (4)
Treatment	93.86 (37.11)	33.86 (15.93)	29.57 (13.47)	30.42 (12.90)
Area characteristics	Yes	Yes	Yes	Yes
Mean of dependent variable in control	319.0	167.1	92.1	59.8
Adjusted R^2	0.113	0.116	0.064	0.106
Observations	307	307	307	307
Median treatment effect	104.4 (43.9)	51.3 (21.0)	46.6 (21.0)	31.8 (15.0)
EDF p-value	0.030	0.066	0.058	0.042
RI p-value	0.030	0.087	0.024	0.051

Notes: OLS estimates, standard errors clustered at the district level. *EDF p-value* refers to the p-value from a null hypothesis that the treatment effect is zero (in the same regression), using the Young (2016) effective degrees of freedom correction. *RI p-value* refers to the equivalent p-value using a randomization inference procedure (specifically, the randomization-t p-value from Young 2019). Standard errors for the median treatment effect are bootstrapped and clustered at the district level. The dependent variable is total number of households visited over the relevant time horizon. SMS receipts are sent by individual CHAs to MOH for each visit. *Treatment* = 1 if the health worker is recruited in a district where career opportunities were made salient. All regressions include the stratification variables (province dummies and share of high school graduates in the district). *Area characteristics* include: number of staff in the health post, geographical distribution of households in the catchment area, and an indicator variable that equals 1 if the CHA reports to have good cell network coverage most of the time or all the time.

services, the difference between treatments is likely to have implications for health outputs in these communities. We return to this issue in Section V.

Columns 2–4 divide the 18-month period into three and show that the estimated treatment effect is identical in the three semesters. This casts doubt on the alternative hypothesis that agents in the two groups have the same traits, but agents in the treatment group perceive stronger career incentives because they have known about them for longer (about 2 years versus 1 year for the control group). Such a difference should wane with time, while the difference due to stable traits should be stable.[28]

To shed light on what treatment health workers do differently, we administer a time use survey to all health workers after they have started working. The findings, reported in detail in the Appendix, indicate that treatment and control health workers work similar hours and allocate their time similarly across similar activities. This indicates that treatment health workers are more efficient at their jobs. Household visits take place in remote, low-density areas: the median 78 km^2 area has 200 households, with an interquartile range of 130 to 360. It is thus rather time-consuming to go from house to house, and this is compounded by the fact that roads are bad. In this setting, the ability to plan, e.g., by making appointments with specific households or collecting information as to whether members are likely to be home before setting out to visit them, is an important determinant of completing visits successfully.

[28] The fact that the treatment effect is stable also casts doubt on the possibility that it is driven by a negative "surprise" for agents in the control group (i.e., their effort response to finding out about career opportunities is negative and larger, in absolute value, than what it would have been had they known the career opportunities at the outset). In addition, if there really were a substantial intrinsic crowd-out of this kind, we would likely expect at least some greater attrition in the control group as participation constraints that were met ex ante would not be met ex post. Yet, as discussed later, we find that attrition is neither economically nor significantly different between treatment and control. Nonetheless, it remains possible that intrinsic crowd-out may be contributing to the results if such an effect does not fade with time and does not lead to dropping out.

In our model, we consider effort to be measuring productive inputs, thus capturing hours used conducting productive activities (e.g., household visits, community meetings, etc.) rather than just total hours recorded. These findings, of treatment workers producing more real inputs despite similar hours reported, suggest that treatment workers exert more effort than control workers. Our model is agnostic as to whether effort and ability are complementary, i.e., as to whether higher ability, for a given level of prosociality, increases output directly only or also through increased effort. Given that these two groups have similar levels of prosociality but different levels of ability, these findings, suggestive of different (productive) effort levels, appear more consistent with the version in which they are complementary.[29]

To conclude, we establish the extent to which differences in performance are due to selection on observables. We search for the vector of observables that explains the largest possible share of variation of performance in the control group and use the estimated coefficients to predict performance in the treatment group.[30] This yields the predicted difference between treatment and control on the basis of the observables that best predict performance. The best predictors explain 31 percent of the observed variation in control and the predicted difference between treatment and control is 44 visits. Given that the actual, unconditional performance gap is 101, differences in observables explain 44 percent of it. The remaining 56 percent is due to traits we do not measure.

The finding that observables have limited power in explaining performance differences echoes the well-established finding that differences in teachers' effectiveness are large and only weakly correlated with observable traits. It is also consistent with other settings where agents self-select, such as in applying for welfare programs (Alatas et al. 2016) or purchasing health products (Ashraf, Berry, and Shapiro 2010). In those settings, as in ours, self-selection cannot be mimicked by targeting on observable traits.

C. *Beyond Number of Visits: Compensation Mechanisms and Other Activities*

Table 5 investigates the hypothesis that health workers in the control group take other actions that compensate for the lower number of visits. Column 1 tests whether control health workers are more likely to be retained while career health workers leave with their newly acquired skills as soon as it is feasible to do so. Since the health workers are bonded to their position for one year,[31] we measure retention by the number of health workers who make at least one visit after the one-year commitment has elapsed. We find that, by this measure, 18 percent of health workers drop out, though some of this may be due to a combination of malfunctioning phones and the rainy season (falling between months 15–18 in our analysis window)

[29] This is also consistent with the large magnitude of the treatment effects, since the complementarity magnifies differences in ability.

[30] Specifically, we select the five best predictors, in addition to our stratification variables, using the Furnival-Wilson branch and bound algorithm (implemented using *vselect* in Stata) in the control group. We then use these to predict the mean number of visits in both treatment and control, and calculate the difference in these predicted means.

[31] The health workers were told that if they quit before one year of service, they would be required to pay monthly wages for any months not worked (rather than simply relinquishing pay) to compensate the Government for the free one-year training that they received.

Table 5—Compensation Mechanisms

Dependent variable: Source: Unit of observation:	Retention SMS receipts CHA (1)	Visit duration SMS receipts CHA (2)	Women and children visited per HH HMIS records Health post (3)	Unique HHs visited SMS receipts CHA (4)
Panel A				
Treatment	0.002	0.260	0.049	36.34
	(0.048)	(1.85)	(0.097)	(15.47)
Area characteristics	Yes	Yes	Yes	Yes
Mean of dependent variable in control	0.823	33.7	2.07	187.1
Adjusted R^2	0.041	0.012	0.064	0.121
Observations	307	304	142	307
EDF *p*-value	0.963	0.900	0.651	0.043
RI *p*-value	0.966	0.889	0.642	0.046

Dependent variable: Source: Unit of observation:	Visits per HH SMS receipts CHA (5)	Community mobilization meetings HMIS records Health post (6)	Patients seen at health post HMIS records Health post (7)	Emergency calls Time use survey CHA (8)
Panel B				
Treatment	0.487	17.12	36.90	0.047
	(0.246)	(5.23)	(261.9)	(0.058)
Area characteristics	Yes	Yes	Yes	Yes
Mean of dependent variable in control	1.76	20.32	1,126.6	0.457
Adjusted R^2	0.125	0.031	0.025	0.006
Observations	307	146	146	298
EDF *p*-value	0.085	0.006	0.899	0.473
RI *p*-value	0.029	0.012	0.907	0.517

Notes: OLS estimates, standard errors clustered at the district level. *EDF p-value* refers to the *p*-value from a null hypothesis that the treatment effect is zero (in the same regression), using the Young (2016) effective degrees of freedom correction. *RI p-value* refers to the equivalent *p*-value using a randomization inference procedure (specifically, the randomization-*t* *p*-value from Young 2019). *Treatment* = 1 if the health worker is recruited in a district where career opportunities were made salient. *Retention* = 1 if the CHA still reports visits after one year. *Visit duration* is computed as end time minus start time in minutes. *Emergency calls* = 1 if the CHA takes at least one out-of-hours call in a typical week. *SMS receipts* are sent by individual CHAs to MOH for each visit. The Health Management and Information System (HMIS) is the Zambian Ministry of Health's system for reporting health services data at government facilities. The two CHAs are required to submit monthly reports that summarize their activities at the health post/community level. The number of observations varies because some health posts do not submit the reports; these are equally distributed between treatment and control. The time use survey was administered in May 2013 during a refresher training program. All regressions include the stratification variables (province dummies and share of high school graduates in the district). Area characteristics include: number of staff in the health post, geographical distribution of households in the catchment area, and an indicator variable that equals 1 if the CHA reports to have good cell network coverage most of the time or all the time.

making travel to cell network-accessible areas difficult. This attrition rate is balanced across treatments. It is important to note that according to the Ministry's rule, health workers have to wait two years before applying for higher-ranked positions, such that none of those who left their positions did so for career progression.[32]

The number of visits can hide heterogeneity on a variety of dimensions that can make the health workers less effective in generating health outcomes, such as doing

[32] It is possible that career opportunities may affect long-term retention; we discuss the welfare implications of this in Section VI.

shorter visits, targeting the head of household rather than women and children, or targeting easier-to-reach households. We provide evidence that career health workers do not do worse on any of these dimensions. They devote the same time to a single visit (column 2) and are equally likely to target their primary clients, women and children (column 3). They also reach more households (column 4) and make more follow-up visits (column 5). The point estimates indicate that just over one-third (36/94) of the total treatment effect is due to career health workers visiting more households, and two-thirds to them visiting the same household more than once. This is consistent with the two groups of health workers having a similar number of households in their catchment area and visiting them at least once, but treatment health workers doing more follow-up visits. Note that follow-ups are considered an integral part of the health worker job, in view of which Ministry of Health guidelines state health workers should attempt to visit each household on a quarterly basis. Finally, online Appendix Table A5 shows that treatment health workers allocate their time in a similar way to control health workers during household visits. This allays the concern that health workers who see themselves as health professionals neglect "soft" tasks like counseling.

Besides household visits, the health workers are expected to assist staff at the health post by seeing patients, assisting with antenatal care, and maintaining the facility. They are also supposed to organize community meetings such as health education talks at the health post and in schools. Columns 6 and 7 of Table 5 investigate whether differences in household visits are compensated by differences in secondary tasks using HMIS data on the number of community meetings health workers organize and the number of patients they attend to at the health post. The latter should be seen as a proxy of the quantity of services delivered by the health workers at the health post, as seeing patients is mostly a nurse's job. We find that health workers recruited by making career opportunities salient organize twice as many meetings over 18 months, and the difference is precisely estimated. The effect of career opportunities on the number of patients the health workers see at the health post is also positive, but small and not precisely estimated.

V. Facility Utilization, Health Practices, and Health Outcomes

The program leads to a substantial increase in the number of health staff operating in the communities where the health workers are deployed: the number of staff associated with the community health post increases on average from 1.5 to 3.5. Given the size of the increase and the magnitude of the treatment effect on household visits and community mobilization meetings, it is reasonable to expect treatment to affect health outcomes in these communities. The health workers can directly affect facility utilization and health practices by increasing both demand, e.g., by providing information and promoting behavioral changes, and supply, e.g., by helping cover staff shortages at the health post or delivering medical treatments to households. In turn, improved facility utilization and health practices should lead to better outcomes.

Besides their intrinsic importance for the welfare of these communities, treatment effects on facility utilization and household outcomes allow us to shed light on whether health workers in the control group perform better on dimensions we cannot observe but improve outcomes. For instance, treatment health workers could target households that are more interested in health services and would use facilities when necessary anyway, while control health workers could target households that they need to persuade to change behavior, and that require more work, leading to fewer visits overall. If this were true, treatment would be uncorrelated (or even negatively correlated) with facility utilization and health outcomes.

To provide evidence on whether treatment affects facility utilization, we use data from the Ministry's HMIS administrative records; to measure effects on health practices and outcomes, we survey households residing in the communities where the health workers operate. As the main remit of the CHA job is maternal and child health, we focus on this throughout.

A. *Treatment Effect on Facility Utilization*

The Ministry's HMIS administrative records are compiled by facilities' senior staff and transmitted to the Ministry of Health via an electronic platform. Two levels of facilities serve these communities: health centers and health posts.[33] The health workers are supposed to encourage women to give birth at the closest health center and to bring in children for regular visits and immunizations at the closest facility (health center or health post). The importance of institutional deliveries in this context cannot be understated: Zambia's maternal mortality rates are very high and health centers have the equipment and medical supplies that can prevent these deaths. Regular children's visits ensure that conditions such as malnutrition are addressed before they become severe. Immunizations protect children from potentially fatal illnesses.

To test whether the treatment affected facility utilization, we obtain information on institutional deliveries, children's visits, and immunizations for the period January 2011–June 2014 and estimate the following specification:

$$y_{hdpt} = \alpha + \beta C_{hd} + \gamma A_t + \delta C_{hd} \times A_t + Z_h \theta + E_d \phi + \rho_p + \xi_{hdpt}, \tag{2}$$

where y_{hdpt} is the outcome in health facility h in district d and province p at quarter t.[34] h represents the lowest level of government facility to which the health workers can refer their patients. This is the health post if operational or the closest health center otherwise. The only exception is childbirths, which are always measured at the health center level, as that is where they are supposed to take place. $C_{hd} = 1$ if facility h is located in a district randomly assigned to the career treatment. We have data for 14 quarters, equally divided before and after the health workers' arrival, and $A_t = 1$ after the health workers' arrival (2012:IV). To minimize composition bias and to test for robustness to facility fixed effect models, we restrict the sample to the facilities for which we have at least three observations before and after the

[33] Health facilities in Zambia are structured according to a population-based hierarchy. Health posts are the first-level health facility for most rural communities and provide basic medical care (no inpatient or surgical services). Health centers, which typically serve a population encompassing four to five health posts, provide both outpatient and inpatient services, including labor and delivery and minor surgical procedures. District hospitals in turn encompass several health center catchment areas and are primarily focused on inpatient care.

[34] HMIS data should be transmitted to MOH monthly, but in practice (due to poor connectivity), reports are missing for some months and the information added to the following month. We aggregate the data at the quarterly level to smooth out monthly fluctuations due to this.

health workers' arrival.[35] Z_h is a vector of area characteristics, which includes the number of staff at the health post, cell network coverage, and the distribution of households between farms and villages described in Table A1. We control for the stratification variables, district-level high school graduation rate E_d and province indicators ρ_p, throughout. Standard errors are clustered at the level of randomization, the district, and, as mentioned before, we also report p-values from the effective degrees of freedom correction in Young (2016) and a randomization inference procedure (Young 2019).

The parameter of interest is δ, the difference-in-differences between facilities in treatment and control districts before and after the health workers' arrival. Under the parallel trend assumption, δ captures the effect of career opportunities for health workers on these outputs.

Table 6 shows that, indeed, career opportunities improve clinic utilization outputs. In particular, the number of women giving birth at a health center increases by 30 percent relative to the mean in control areas at baseline (column 1). The effect on institutional deliveries is thus the same order of magnitude as the effect of performance pay for clinics as evaluated in Rwanda (23 percent; Basinga et al. 2011) and Cambodia (25 percent; Van de Poel et al. 2014). Selection and incentive effects of similar magnitudes (22 percent each) are also found in the only firm study that identifies the two separately (Lazear 2000).

Table 6 also shows that the number of child health visits increases by 24 percent (column 3), the number of children under 5 weighed increases by 22 percent (column 4), and the number of children under 12 months receiving polio vaccination increases by 20 percent (column 6). The effects on postnatal visits for women and BCG and measles vaccinations are also positive and in the 8–22 percent range, but are not precisely estimated. The average standardized treatment effect (Kling, Liebman, and Katz 2007) over all outcomes is 0.278 (column 8), significantly different from 0 at the 1 percent level. Reassuringly, there are no significant differences between treatment and control in any of these outcomes before the health workers' arrival: all the estimated β coefficients are small and not significantly different from 0.

To provide support to our identifying assumption, in panel A of online Appendix Table A6, we run a placebo test where we split the pre-health worker period in two halves and test whether outcomes improve in treatment areas over time even in the absence of the health workers. Reassuringly, they do not. Finally, panel B estimates equation (2) with facility fixed effects; the fact that all estimated δ coefficients remain stable provides evidence that they are not biased by time-invariant facility unobservables correlated with treatment.

B. *Treatment Effect on Health Practices and Outcomes*

To provide evidence on the effect of treatment on health practices and outcomes, we survey households in 47 randomly chosen communities located in each of the

[35] This restriction keeps an average of 76 percent of the health posts and 71 percent of the health centers in the sample.

Table 6—The Effect of Career Opportunities on Facility Utilization

	Institutional deliveries (1)	Postnatal (0–6 weeks) visits (2)	Children under five visited (3)	Children under five weighed (4)
Panel A				
Treatment	−2.17	−12.88	−65.80	−72.88
	(11.27)	(9.314)	(142.7)	(133.3)
After	4.25	15.50	61.78	108.9
	(4.27)	(5.11)	(63.40)	(64.07)
Treatment × After	13.65	7.79	312.1	277.8
	(6.21)	(9.380)	(97.67)	(109.7)
Area characteristics	Yes	Yes	Yes	Yes
Mean of dependent variable in control in year 1	45.3	55.4	1,285.7	1,236.1
Adjusted R^2	0.348	0.215	0.254	0.253
Number of facilities	89	119	123	123
Observations	1,301	1,543	1,618	1,610
EDF *p*-value	0.038	0.418	0.003	0.017
RI *p*-value	0.014	0.455	0.010	0.033

	Children under one receiving BCBG vaccinations (5)	Children under one receiving polio vaccinations (6)	Children under one receiving measles vaccinations (7)	Average standardized effect (8)
Panel B				
Treatment	11.01	−0.286	1.730	−0.005
	(11.96)	(9.118)	(10.00)	(0.156)
After	−1.270	−1.124	−1.168	0.042
	(4.618)	(3.711)	(3.604)	(0.059)
Treatment × After	7.147	14.60	11.19	0.278
	(8.881)	(4.809)	(7.259)	(0.092)
Area characteristics	Yes	Yes	Yes	Yes
Mean of dependent variable in control in year 1	83.0	72.4	70.4	—
Adjusted R^2	0.151	0.151	0.118	—
Number of facilities	120	121	120	—
Observations	1,518	1,531	1,535	1,102
EDF *p*-value	0.433	0.005	0.138	—
RI *p*-value	0.463	0.003	0.134	—

Notes: OLS estimates, standard errors clustered at the district level. *EDF p-value* refers to the *p*-value from a null hypothesis that *Treatment* × *After* is 0 (in the same regression), using the Young (2016) effective degrees of freedom correction. *RI p-value* refers to the equivalent *p*-value using a randomization inference procedure (specifically, the randomization-*t p*-value from Young 2019). *Treatment* = 1 if the health worker is recruited in a district where career opportunities were made salient. Data source is the Health Management and Information System (HMIS) available monthly from January 2011 until June 2014. Health center and health post staff are required to submit monthly reports that summarize their activities at the health post/community level. These are aggregated at the quarter level in the regressions. The variable in Column 1 is defined at the health center level because health centers are equipped for child births and health posts are not. The variables in columns 2–7 are defined at the health post level if this reports data, at the health center otherwise. The average standardized treatment effect is computed using the methodology in Kling, Liebman, and Katz (2007). *After* = 1 after September 2012 (from 2012:IV onward), when CHAs started working. All regressions include the stratification variables (province dummies and share of high school graduates in the district). *Area characteristics* include: number of staff in the health post, geographical distribution of households in the catchment area, and an indicator variable that equals 1 if the CHA reports to have good cell network coverage most of the time or all the time.

47 districts where the health workers operate.[36] We randomly choose 16 households in each community, surveying 738 in total.[37] These surveys are administered by a team of enumerators who are trained by us and unconnected to the health workers or the Ministry of Health. As the main focus of the health worker job is maternal and child health, we only survey households that contain at least one child under five. The survey contains modules on health and sanitation knowledge, health practices, incidence of illnesses, and anthropometrics for the youngest child. Knowledge, practices, and illnesses are self-reported; deworming and immunization data are drawn from the child health card, and anthropometrics are measured by trained enumerators. We interview the main carer of the child, which is their mother in 90 percent of cases and either a grandparent or a sibling in the remaining 10 percent. All questions are drawn from the Zambia Demographic and Health Survey (DHS) questionnaire, with the exception of the health literacy test, which we designed based on the CHA curriculum, and mid-upper arm circumference (MUAC), which the DHS does not measure.

Table 7 reports the estimates of

$$y_{idp} = \alpha + \beta C_{id} + D_i \gamma + \delta E_d + \rho_p + \epsilon_{idp}, \tag{3}$$

where y_{idp} is the outcome of child (or respondent) i in district d and province p. $C_{id} = 1$ if child (or respondent) i lives in a district that is assigned to the career opportunities treatment. D_i is a vector of child, respondent, and household characteristics that includes child age and gender, household size and number of assets, and the education level of the respondent. As above, we control for the stratification variables, district-level high school graduation rate E_d and province indicators ρ_p, throughout, and cluster standard errors at the district level, with additional p-values reported based on the effective degrees of freedom correction procedure in Young (2016) and a randomization inference procedure (Young 2019).

Column 1 of Table 7 shows that the average respondent answers 74 percent of the health literacy questions correctly and that this does not differ by treatment status. In contrast, treatment affects all the health practices we collect information on. In particular, columns 2 and 3 show that children up to 2 years old living in treatment areas are 5pp more likely to be breastfed,[38] and their stools are 12pp more likely to be safely disposed; these effects represent an 8 percent and 20 percent increase from the control group mean, respectively. Columns 4 and 5 show that treatment also increases the number of deworming treatments by 16 percent and the likelihood that the child is on track with the immunization schedule by 4.7pp, which is

[36] Although 48 districts were randomized, one district did not participate in the recruitment process or submit nominations for CHA candidates, and hence was excluded from the survey sample.

[37] The sample frame had 752 households but we interviewed 738. The missing households are evenly spread across communities as the number of households surveyed in a community varies between 13 and 16. The difference is due to several factors. In some communities, safety concerns related to local political tensions forced the survey team to leave the community before completing surveying. In other communities, especially low-density communities where travel times between households could exceed one hour, the survey team was unable to find a sufficient number of eligible households within the allotted survey time. One household interview was lost due to malfunction of the mobile device on which the interview was recorded.

[38] WHO recommends breastfeeding until the age of two years.

Table 7—The Effect of Career Opportunities on Health Practices and Outcomes

	Percentage of correct answers on health literacy test (1)	= 1 if child is breastfed up to 2 years (2)	= 1 if child's stools are safely disposed (3)	Number of deworming treatments (4)	= 1 if child exposed to CHA is on track with immunization schedule (5)
Panel A. Health literacy and practices					
Treatment	0.002	0.051	0.121	0.225	0.047
	(0.010)	(0.023)	(0.039)	(0.129)	(0.020)
Household controls	Yes	Yes	Yes	Yes	Yes
Child controls	No	Yes	Yes	Yes	Yes
Mean of dependent variable in control	0.740	0.641	0.597	1.45	0.058
Adjusted R^2	0.057	0.561	0.161	0.263	0.024
Observations	738	613	736	659	462
EDF p-value	0.827	0.054	0.007	0.119	0.038
RI p-value	0.784	0.072	0.005	0.176	0.045

	= 1 if child experienced fever in the last two weeks (6)	= 1 if child experienced diarrhea in the last two weeks (7)	= 1 if child experienced cough in the last two weeks (8)
Panel B. Incidence of illness			
Treatment	−0.003	0.037	−0.070
	(0.037)	(0.027)	(0.033)
Household controls	Yes	Yes	Yes
Child controls	Yes	Yes	Yes
Mean of dependent variable in control	0.468	0.256	0.448
Adjusted R^2	0.077	0.017	0.021
Observations	731	731	731
EDF p-value	0.946	0.216	0.060
RI p-value	0.951	0.262	0.107

	= 1 if weight-for-age z-score < 2 SD (moderately or severely undernourished) (9)	= 1 if weight-for-age z-score < 3 SD (severely undernourished) (10)	= 1 if MUAC < 12.5 (moderately or severely wasted) (11)	= 1 if MUAC < 11.5 (severely wasted) (12)	Average standardized effect (anthro only) (13)	Average standardized effect (all) (14)
Panel C. Anthropometrics and average standardized effects						
Treatment	−0.053	−0.028	−0.023	−0.014	0.124	0.108
	(0.030)	(0.015)	(0.016)	(0.014)	(0.062)	(0.037)
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
Child controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dependent variable in control	0.210	0.051	0.036	0.015	—	—
Adjusted R^2	−0.006	0.003	0.018	0.017	—	—
Observations	582	582	581	581	579	376
EDF p-value	0.114	0.084	0.179	0.342	—	—
RI p-value	0.058	0.099	0.205	0.359	—	—

Notes: OLS estimates, standard errors clustered at the district level. *EDF p-value* refers to the p-value from a null hypothesis that the treatment effect is 0 (in the same regression), using the Young (2016) effective degrees of freedom correction. *RI p-value* refers to the equivalent p-value using a randomization Inference procedure (specifically, the randomization-t p-value from Young 2019). *Treatment* = 1 if the health worker is recruited in a district where career opportunities were made salient. The health literacy test contains 14 questions on topics that CHAs are supposed to cover; these questions were drafted by the researchers in consultation with CHA program officials and the CHA curriculum. Breastfeeding and stool disposal are self-reported. In line with UNICEF and WHO (2017), we define stools as safely disposed if flushed in toilet/latrine. Deworming and immunization data are as reported in the child health card. A child is defined as on track if they have completed all immunizations required for their age in months. The immunization sample is restricted to children who were three months or younger (including unborn) when the CHAs started working. Thresholds for weight-for-age and MUAC are taken from WHO guidelines; following these, data are restricted to children between 6–59 months. *Household controls* include size, education level of the respondent, and number of assets. *Child controls* include age and gender. All regressions include the stratification variables. The average standardized treatment effect is computed using the methodology in Kling, Liebman, and Katz (2007) after recoding all variables so that higher values indicate better outcomes. For weight-for-age z-score and MUAC we use the < 3 SD and < 11.5cm thresholds, respectively.

81 percent of the control group mean (5.8 percent).[39] Importantly, the treatment affects the incidence of immunizations for children who are young enough to have been exposed to the health workers when their immunization period started (as shown in column 5), but not for those who were too old to start the cycle when the health workers started working (coefficient −0.014, standard error 0.020). This echoes the findings in online Appendix Table A6 that show no difference in immunization rates between treatment and control areas before the health workers started working.

Panel B measures treatment effects on the incidence of three main illness symptoms: fever, diarrhea, and cough. These are fairly common, as 47 percent, 26 percent, and 45 percent of children in control areas had experienced them, respectively, in the past two weeks. As is widely acknowledged, self-reported symptoms can actually worsen as knowledge improves and individuals learn how to recognize them, so these effects are lower bounds. We find that treatment reduces the incidence of cough symptoms by 7pp while leaving the others unchanged. Finally, panel C shows treatment effects on anthropometric measurements. We report weight-for-age z-scores and mid-upper arm circumference (MUAC). The combination of these two allows us to measure both chronic and acute malnutrition.[40] Following WHO's guidelines, we use the −2SD and −3SD thresholds for weight-for-age z-scores to measure moderate and severe underweight, respectively, and 12.5 cm and 11.5 cm for MUAC to measure moderate and severe wasting, respectively (Food and Nutrition Technical Assistance Project 2011). According to these measures, 21 percent of the children in control areas are underweight, and 5 percent severely so. The incidence of wasting is much lower, with 3.6 percent of the children exhibiting some wasting and 1.5 percent severe wasting. These data, which match the corresponding DHS figures for rural Zambia (Government of Zambia 2014), suggest that these areas are characterized by high rates of chronic malnutrition but low rates of acute malnutrition.

The findings in panel C, columns 9 and 10 show that children in treatment areas are 5pp less likely to be underweight (25 percent of the control group mean) and 3pp less likely to be severely underweight (55 percent of the control group mean). In line with this, columns 11 and 12 show a large percentage reduction in wasting, but given the limited occurrence of this in our sample, the effects are not precisely estimated. The average standardized effect of the two measures is precisely estimated with p-value 0.019 for the less severe measures and 0.045 for the more severe, with the latter shown in column 13.

The average standardized treatment effect across all variables (coded so that higher values correspond to better outcomes) is 0.108 (column 14), significantly different from 0 at the 1 percent level.

[39] A child is defined to be on track if she has completed all immunizations required for her age. At age 3 months, this includes BCG, OPV 0-2, PCV 1-2, DPT-HepB-Hib 1-2, and rotavirus 1-2. At 4 months, this includes, additionally, OPV 3, PCV 3, and DPT-HepB-Hib 3. At 9 months, this includes OPV 4 if OPV 0 was not given, and measles 1. The immunization series is complete at age 18 months with measles 2. Finally, we consider a child to be on track for vitamin A supplementation if she has ever been supplemented.

[40] We did not measure weight-for-height, an alternative to MUAC for assessing acute malnutrition, for three reasons. First, compared to weight and MUAC, height measurement is more invasive, requiring, for children under two, laying the child down on a height board and having two enumerators hold the child while collecting the measurement. During survey piloting, many respondents (and the children themselves) balked at this procedure. Second, accurate height measurement is made difficult by high measurement error relative to standard effect sizes (Mwangome et al. 2012). Finally, MUAC is a more accurate predictor of mortality (Myatt, Khara, and Collins 2006).

Taken together, the findings in this and the previous section show that differences in the inputs provided by treatment and control health workers are matched by differences in facility utilization and household health practices. The selection effect of career opportunities is strong enough to generate discernible differences in household behaviors and child health outcomes.

VI. Conclusion

Attracting effective employees is a core objective for all organizations. This can be a particularly challenging objective to achieve for public organizations because both effective performance (in, for example, generating health impact) and desirable employee attributes are difficult to measure. But the stakes to getting this right are high. Our paper has shown that offering a civil service position with career opportunities for community-based work attracts agents who deliver health services with substantial impact. This significant effect on the health and well-being of communities is driven entirely by a selection effect of the types of agents drawn into the position.

The civil service job we study is one sometimes referred to as a "street-level bureaucrat" (Lipsky 1980), a job where internalizing the utility of beneficiaries could be particularly helpful. Yet it is in just such a job that offering a career in the civil service, in posters that clearly attracted ambitious types, provided large impacts. Of course, the career opportunities which attracted ambitious types, a career in the Ministry of Health, entail some social benefit, and the community-oriented nature of the job attracted a basic level of altruism across the board. But it is in precisely these types of jobs where it has been argued that adding individualistic benefits, such as career opportunities, might attract the "wrong" type of individual. Our experiment reveals that this is indeed the case, as the lower-ability applicants in the treatment group have lower prosociality. Thus, if candidates were picked by a random draw we would expect fewer prosocial recruited candidates in treatment.

In practice, however, selection mechanisms, in Zambia and elsewhere, do not choose applicants randomly. To the extent that the mechanism picks from the top of the ability distribution, the sorting equilibrium guarantees that these are the most prosocial. This allays the concern, often expressed by policymakers, that offering material rewards will crowd out prosocial applicants in education and health (World Health Organization 2006, Lehmann and Sanders 2007, Muralidharan and Sundararaman 2013). The findings also stress the importance of giving the right incentives to selectors. The two components of recruitment (sorting and selection) are equally important because good candidates cannot be hired if they do not apply and improving the applicant pool is useless unless the best candidates are selected.

As with many microeconomic studies, it can be challenging to generalize results to other contexts. Yet, we can specify three conditions under which we can expect to find a similar effect. First, the task needs to have social impact and this needs to be common knowledge. Second, the selection panels need sufficiently good information on ability so as to be able to pick from the top of the ability distribution. Third, and most importantly, the most able candidate is also the most prosocial because returns to ability on the job are lower than in the outside option. If not, all high-ability individuals would apply regardless of their prosociality. Conversely, the effect would be stronger when there is a well-developed private sector that offers alternative jobs with high rewards for ability.

Tailoring job design to attract star performers has three advantages over the more common alternative of setting higher requirements on observable characteristics. First, these characteristics are difficult to identify; in our case observables only explain 44 percent of the performance gap. Second, requirements can only screen out those who do not meet them; they do not necessarily draw in those who do. Most importantly, requirements create barriers to entry and rents, which can draw in applicants attracted by rent extraction rather than public service delivery.[41]

The findings measure the productivity gains that come from effective selection via recruitment: treatment health workers provide more inputs at the same cost, since wages are the same across both treatments.[42] The fact that the health workers are recruited locally from the communities where they are meant to serve implies that there is no competition for talent across communities: career opportunities can thus be offered in each community without losing effectiveness, as each community can only hire from their own pool, and most communities in these areas have access to a pool of skilled individuals who are either unemployed or in low-skill jobs.

While retention rates after 18 months are the same in the two groups, agents in the career incentives treatment might leave their posts for higher-ranked positions sooner than those in the control group. Whether this entails a welfare cost depends on whether they can be easily replaced and whether the government can use their skills in other jobs. In our context, replacement is straightforward; the number of applicants per post was above seven, and the government faces scarcity of health staff at all levels, such that promoting strong performers to nursing and other higher-level cadres is likely to be welfare-improving. In contexts where retention in the original post is more important, the welfare cost of attracting agents who expect to move on will be higher.

More generally, we cannot quantify the opportunity cost of the health workers' time, namely the value of the activities they give up to become full-time health workers, and the size of this difference between treatment and control. If productivity in these alternative occupations is increasing in the same qualities that make a health worker productive, the findings imply that the opportunity cost is higher in the treatment group. By revealed preference, we know that the private value of the health worker job must be at least equal to the private value of these activities. Otherwise, these individuals would not have switched occupations. To the extent that health work generates more social value than the outside option, the private and social optima coincide, and a career-oriented position for community-based public service delivery improves social welfare.

[41] In this case, pay and qualifications might end up being negatively correlated with equilibrium performance as illustrated by the evidence on effort and performance among civil-service versus contract teachers, as in, for example, Muralidharan and Sundararaman (2013); Duflo, Dupas, and Kremer (2015); and Kremer, Brannen, and Glennerster (2013). Muralidharan and Sundararaman (2013) finds that, in rural Andhra Pradesh, contract teachers (who have less education and qualifications) are at least as effective as regular civil-service teachers, suggesting that the substantial wage differential (of over a factor of 5) "is unlikely to reflect differences in productivity and mostly represents rents accruing to unionized civil-service teachers." We owe this suggestion to a constructive referee.

[42] Due to political constraints, all agents had to be paid the same amount. This implies that we cannot judge whether agents attracted by career opportunities have a higher reservation wage, such that their higher performance comes at a price; in other words, the government could get the agents in the control group to work for a lower wage. A priori, the difference in reservation wages between applicants in the two treatments is difficult to sign: that applicants to the career opportunities treatment are more skilled suggests that it might be positive, whereas the fact that they expect to move on to better-paid positions suggests that it might be negative (in the manner that interns are typically willing to forgo compensation for the sake of career opportunities).

Appendix

Table A1—Eligible Population by Treatment (Randomization Balance)

	Treatment	Control	p-value (clustered)	p-value (EDF)	p-value (RI)
Panel A. Characteristics of the eligible population					
Share of eligibles in district (18–45-year-olds with grade 12 or above)	0.044	0.043	0.917	0.919	0.961
	(0.205)	(0.203)			
Share of women among eligibles	0.371	0.391	0.241	0.258	0.142
	(0.483)	(0.488)			
Main activity of eligible candidates during past 12 months					
Not working	0.296	0.280	0.480	0.494	0.613
	(0.456)	(0.449)			
Unpaid work	0.201	0.229	0.344	0.361	0.522
	(0.401)	(0.420)			
Paid work	0.457	0.437	0.353	0.371	0.408
	(0.498)	(0.496)			
of which: mid-skill	0.240	0.230	0.705	0.714	0.724
	(0.427)	(0.421)			
of which: low-skill	0.484	0.453	0.173	0.189	0.149
	(0.500)	(0.498)			
Panel B. Catchment area characteristics					
Number of staff in health post∗	1.49	1.39	0.635	0.640	0.604
	(1.10)	(1.17)			
Geographical distribution of households in catchment area:∗					
Most people live in their farms, none in villages	0.082	0.091	0.846	0.848	0.825
	(0.277)	(0.289)			
Some people live in farms, some in small villages (5–10 HHs)	0.529	0.519	0.915	0.916	0.930
	(0.502)	(0.503)			
Most people live in medium or large villages (more than 10 HHs), a few on their farms	0.388	0.364	0.816	0.819	0.804
	(0.490)	(0.484)			
Poor cell network coverage∗	0.082	0.065	0.681	0.685	0.743
	(0.277)	(0.248)			
Panel C. Target population characteristics					
Share of district population under 5	0.187	0.187	0.915	0.916	0.877
	(0.390)	(0.390)			
Main type of toilet: pit latrine or better∗∗	0.718	0.667	0.494	0.501	0.385
	(0.450)	(0.471)			
Household water supply: protected borehole or better∗∗	0.361	0.416	0.248	0.257	0.122
	(0.480)	(0.493)			

Notes: Columns 1 and 2 show means and standard deviations in parentheses. Column 3 reports the p-value of the test of equality of means based on standard errors clustered at the district level. Column 4 reports the p-value using the Young (2016) effective degrees of freedom (EDF) correction, clustered at the district level. Column 5 reports the p-value using a randomization inference (RI) procedure, clustered at the district level (specifically, the randomization-t p-value from Young 2019). *Treatment* $= 1$ if the health worker is recruited in a district where career opportunities were made salient. Variables are drawn from the 2010 Census (10 percent PUMS sample) except those indicated by ∗, which are drawn from our surveys, and those indicated by ∗∗, which are drawn from the 2010 Living Conditions Monitoring Survey (LCMS), which covers 20,000 HHs and is representative at the district level. Activities codes follow the ILO ISCO88 convention. *Mid-skill* includes ISCO codes between 300 and 599, namely technicians, clerical workers, and services and sales workers. *Low-skill* includes ISCO codes above 600, namely agriculture, crafts, basic manufacturing, and elementary occupations. *Number of staff in health post* is the total number of nurses, environmental health technicians, and clinical officers assigned to the health post as reported by district officials surveyed by phone. Information on the geographical distribution of HHs was obtained from a survey of the deployed CHAs before deployment. CHAs were shown stylized maps accompanied by a description and asked to choose the one that most closely resembled the catchment area of their health post. Questions were asked to each CHA individually so that two CHAs from the same health post could give different answers. For the 5 out of 161 cases in which the two CHAs gave different answers, we use the information provided by supervisors to break the tie. To measure cell network coverage we attempt to call all CHAs after deployment. We make daily calls for 118 consecutive days. The health post is classified as having poor coverage if we do not manage to reach either of its two CHAs during this period. *Main type of toilet: pit latrine or better* $= 1$ if the surveyed household uses a pit latrine, ventilated improved pit (VIP), or flush toilet, and 0 if bucket, other, or no toilet. *Household water supply: protected borehole or better* $= 1$ if the water supply comes from a protected borehole or well, communal tap, or other piped water system, and 0 if it comes from an unprotected well or borehole, river/dam/stream, rain water tank, or other.

Table A2—Experimental Checks

	Treatment	Control	*p*-value (clustered)	*p*-value (EDF)	*p*-value (RI)
Panel A. Expected job benefits at entry, *June 2011*					
Good future career	0.164	0.119	0.002	0.006	0.007
	(0.157)	(0.113)			
Allows me to serve the community	0.397	0.436	0.046	0.072	0.078
	(0.226)	(0.243)			
Earns respect and status in the community	0.037	0.059	0.031	0.053	0.074
	(0.094)	(0.110)			
Interesting job	0.149	0.148	0.987	0.988	0.991
	(0.163)	(0.138)			
Allows me to acquire useful skills	0.181	0.161	0.254	0.304	0.365
	(0.169)	(0.137)			
Offers stable income	0.026	0.024	0.645	0.677	0.683
	(0.056)	(0.055)			
Pays well	0.032	0.025	0.414	0.461	0.532
	(0.093)	(0.057)			
Panel B. Expected job benefits on the job, *May 2013*					
Good future career	0.160	0.151	0.650	0.682	0.644
	(0.122)	(0.108)			
Allows me to serve the community	0.364	0.371	0.914	0.922	0.923
	(0.184)	(0.178)			
Earns respect and status in the community	0.039	0.038	0.882	0.893	0.909
	(0.069)	(0.061)			
Interesting job	0.132	0.138	0.354	0.403	0.438
	(0.103)	(0.106)			
Allows me to acquire useful skills	0.216	0.219	0.888	0.899	0.894
	(0.132)	(0.111)			
Offers stable income	0.038	0.039	0.989	0.990	0.993
	(0.069)	(0.060)			
Pays well	0.051	0.043	0.664	0.695	0.723
	(0.089)	(0.067)			

Notes: Treatment = 1 if the health worker is recruited in a district where career opportunities were made salient. CHAs were given 50 beans and asked to allocate them on cards, listing different reasons in proportion to the importance of each benefit for working as a CHA. The cards were scattered on a table in no particular order. *At entry* variables are drawn from a survey administered at the beginning of the training program. *On the job* variables are drawn from a survey administered nearly one year after the CHAs started working. We show means with standard deviations in parentheses and the *p*-value of the test of equality of means based on standard errors clustered at the district level (conditional on stratification variables). Column 4 reports the *p*-value using the Young (2016) effective degrees of freedom (EDF) correction, clustered at the district level. Column 5 reports the *p*-value using a randomization inference (RI) procedure, clustered at the district level (specifically, the randomization-*t* *p*-value from Young 2019).

Table A3—Validation of Household Visit Measures

Dependent variable:	Number of visits from HMIS records		= 1 if HH reports a visit by CHA		HH satisfaction: overall CHA's services	
Source:	HMIS		HH survey		HH survey	
Unit of observation:	Health post		HH		HH	
	(1)	(2)	(3)	(4)	(5)	(6)
Number of visits (in 00s) reported by CHA via SMS receipts	0.767	0.644	0.021	0.015	0.039	0.044
	(0.083)	(0.161)	(0.008)	(0.017)	(0.019)	(0.016)
Number of visits (in 00s) reported by CHA via SMS receipts × treatment		0.192		0.010		−0.003
		(0.187)		(0.019)		(0.036)
Mean of dependent variable	643.6		0.438		4.33	
Adjusted R^2	0.473	0.473	0.014	0.013	0.013	0.018
Observations	145	145	1,284	1,284	1,253	1,253
EDF p-value of treatment interaction	—	0.320	—	0.641	—	0.947
RI p-value of treatment interaction	—	0.358	—	0.617	—	0.950

Notes: OLS estimates, standard errors clustered at the health post level in columns 3–6. The independent variable is visits reported by SMS between September 2012 and January 2014. The dependent variable in columns 1 and 2 is the total number of visits done by the two CHAs in the health post drawn from HMIS administrative data over the period between September 2012 and January 2014. The dependent variables in columns 3–6 are drawn from a HH survey administered to 16 HHs in each of 47 communities where CHAs are active. Satisfaction measures range from 1 (very dissatisfied) to 5 (very satisfied). *EDF p-value* refers to the p-value from a null hypothesis that the treatment interaction is 0 (in the same regression), using the Young (2016) effective degrees of freedom correction. *RI p-value* refers to the equivalent p-value using a randomization inference procedure (specifically, the randomization-t p-value from Young 2019).

REFERENCES

Akerlof, George A., and Rachel E. Kranton. 2005. "Identity and the Economics of Organizations." *Journal of Economic Perspectives* 19 (1): 9–32.

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, Ririn Purnamasari, and Matthew Wai-Poi. 2016. "Self-Targeting: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 124 (2): 371–427.

Aron, Arthur, Tracy McLaughlin-Volpe, Debra Mashek, Gary Lewandowski, Stephen C. Wright, and Elaine N. Aron. 2004. "Including Others in the Self." *European Review of Social Psychology* 15 (1): 101–32.

Ashraf, Nava, Oriana Bandiera, Edward Davenport, and Scott S. Lee. 2020. "Replication Data for: Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E111683V1.

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

Ashraf, Nava, James Berry, and Jesse M. Shapiro. 2010. "Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia." *American Economic Review* 100 (5): 2383–2413.

Basinga, Paulin, Paul J. Gertler, Agnes Binagwaho, Agnes L. B. Soucat, Jennifer Sturdy, and Christel M. J. Vermeersch. 2011. "Effect on Maternal and Child Health Services in Rwanda of Payment to Primary Health-Care Providers for Performance: An Impact Evaluation." *The Lancet* 377 (9775): 1421–28.

Bénabou, Roland, and Jean Tirole. 2003. "Intrinsic and Extrinsic Motivation." *Review of Economic Studies* 70 (3): 489–520.

Bénabou, Roland, and Jean Tirole. 2006. "Incentives and Prosocial Behavior." *American Economic Review* 96 (5): 1652–78.

Bénabou, Roland, and Jean Tirole. 2011. "Identity, Morals, and Taboos: Beliefs as Assets." *Quarterly Journal of Economics* 126 (2): 805–55.

Bertrand, Marianne, Robin Burgess, Arunish Chawla, and Guo Xu. Forthcoming. "The Glittering Prizes: Caree Incentives and Bureaucrat Performance." *Review of Economic Studies*.

Besley, Timothy, and Maitreesh Ghatak. 2005. "Competition and Incentives with Motivated Agents." *American Economic Review* 95 (3): 616–36.

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

Collier, Paul. 2009. "Rethinking the Provision of Public Services in Post-Conflict States." In *Contracting Out Government Functions and Services: Emerging Lessons from Post-Conflict and Fragile Situations,* 115–22. Paris: OECD Publishing.

Dal Bó, Ernesto, Frederico Finan, and Martin A. Rossi. 2013. "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service." *Quarterly Journal of Economics* 128 (3): 1169–1218.

Deserranno, Erika. 2019. "Financial Incentives as Signals: Experimental Evidence from the Recruitment of Village Promoters in Uganda." *American Economic Journal: Applied Economics* 11 (1): 277–317.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2015. "School Governance, Teacher Incentives, and Pupil–Teacher Ratios: Experimental Evidence from Kenyan Primary Schools." *Journal of Public Economics* 123: 92–110.

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

Food and Nutrition Technical Assistance Project. 2011. "Anthropometry: Assessing Children under 5 Pocket Reference." http://www.fantaproject.org/tools/anthropometry-pocket-reference.

Fryer, Roland G. 2013. "Teacher Incentives and Student Achievement: Evidence from New York City Public Schools." *Journal of Labor Economics* 31 (2): 373–407.

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

Government of Zambia. 2014. "Zambia Demographic and Health Survey 2013–14: Preliminary Report." Lusaka: Government of Zambia.

King, William Casey. 2013. *Ambition, A History: From Vice to Virtue.* New Haven, CT: Yale University Press.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83–119.

Kremer, Michael, Conner Brannen, and Rachel Glennerster. 2013. "The Challenge of Education and Learning in the Developing World." *Science* 340 (6130): 297–300.

La Familia Sana Program. 1992. *A Historical Overview of Lay Health Worker Programs.* La Familia Sana, Inc.

Lazear, Edward P. 2000. "Performance Pay and Productivity." *American Economic Review* 90 (5): 1346–61.

Lazear, Edward P., and Paul Oyer. 2012. "Personnel Economics." In *The Handbook of Organizational Economics*, edited by Robert Gibbons and John Roberts, 479–519. Princeton, NJ: Princeton University Press.

Lehmann, Uta, and David Sanders. 2007. *Community Health Workers: What Do We Know about Them? The State of the Evidence on Programmes, Activities, Costs and Impact on Health Outcomes of Using Community Health Workers.* Geneva: World Health Organization.

Lipsky, Michael. 1980. *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services.* New York: Russell Sage Foundation.

Mashek, Debra, Lisa W. Cannaday, and June P. Tangney. 2007. "Inclusion of Community in Self Scale: A Single-Item Pictorial Measure of Community Connectedness." *Journal of Community Psychology* 35 (2): 257–75.

Miller, Grant, and Kim Babiarz. 2014. "Pay-for-Performance Incentives in Low- and Middle-Income Country Health Programs." In *Encyclopedia of Health Economics*, edited by Anthony J. Culyer, 457–466. Amsterdam: Elsevier.

Miller, Grant, Renfu Luo, Linxiu Zhang, Sean Sylvia, Yaojiang Shi, Patricia Foo, Qiran Zhao, Reynaldo Martorell, Alexis Medina, and Scott Rozelle. 2012. "Effectiveness of Provider Incentives for Anaemia Reduction in Rural China: A Cluster Randomised Trial." *BMJ* 345: e4809.

Muralidharan, Karthik, Nazmul Chaudhury, Jeffrey Hammer, Michael Kremer, and F. Halsey Rogers. 2011. "Is There a Doctor in the House? Medical Worker Absence in India." Unpublished.

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

Muralidharan, Karthik, and Venkatesh Sundararaman. 2013. "Contract Teachers: Experimental Evidence from India." Unpublished.

Mwangome, Martha K., Greg Fegan, Ronald Mbunya, Andrew M. Prentice, and James A. Berkley. 2012. "Reliability and Accuracy of Anthropometry Performed by Community Health Workers among Infants under 6 Months in Rural Kenya." *Tropical Medicine & International Health* 17 (5): 622–29.

Myatt, Mark, Tanya Khara, and Steve Collins. 2006. "A Review of Methods to Detect Cases of Severely Malnourished Children in the Community for Their Admission into Community-Based Therapeutic Care Programs." *Food and Nutrition Bulletin* 27 (3): S7–S23.

North, Douglass C. 1991. "Institutions." *Journal of Economic Perspectives* 5 (1): 97–112.

Northcote, Stafford, and C. E. Trevelyan. 1853. "Report on the Organisation of the Permanent Civil Service." Report to the House of Commons, London.

Oyer, Paul, and Scott Schaefer. 2011. "Personnel Economics: Hiring and Incentives." In *Handbook of Labor Economics*, Vol. 4B, edited by Orley Ashenfelter and David Card, 1769–1823. Elsevier.

Pérez, Leda M., and Jacqueline Martinez. 2008. "Community Health Workers: Social Justice and Policy Advocates for Community Health and Well-Being." *American Journal of Public Health* 98 (1): 11–14.

Rockoff, Jonah E., Douglas O. Staiger, Thomas J. Kane, and Eric S. Taylor. 2012. "Information and Employee Evaluation: Evidence from a Randomized Intervention in Public Schools." *American Economic Review* 102 (7): 3184–3213.

Rothstein, Jesse. 2015. "Teacher Quality Policy When Supply Matters." *American Economic Review* 105 (1): 100–30.

Roy, Andrew D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3 (2): 136–46.

Schultz, P. Wesley. 2012. "Inclusion with Nature: The Psychology of Human-Nature Relations." In *Psychology of Sustainable Development*, edited by Peter Schmuck and Wesley P. Schultz, 61–78. Dordrecht: Kluwer Academic Publishers.

Staiger, Douglas O., and Jonah E. Rockoff. 2010. "Searching for Effective Teachers with Imperfect Information." *Journal of Economic Perspectives* 24 (3): 97–118.

UNICEF and WHO. 2017. "Progress on Drinking Water, Sanitation and Hygiene: 2017 Update and SDG Baselines." Geneva: WHO and UNICEF.

US Human Resources and Service Administration. 2019. "Shortage Areas." https://data.hrsa.gov/topics/health-workforce/shortage-areas.

Van de Poel, Ellen, Gabriela Flores, Por Ir, Owen O'Donnell, and Eddy Van Doorslaer. 2014. "Can Vouchers Deliver? An Evaluation of Subsidies for Maternal Health Care in Cambodia." *Bulletin of the World Health Organization* 92 (5): 331–39.

Weber, Max. 1922. *Economy and Society.* Tübingen.

World Health Organization. 2006. *The World Health Report 2006: Working Together for Health.* Geneva: World Health Organization.

Young, Alwyn. 2012. "The African Growth Miracle." *Journal of Political Economy* 120 (4): 696–739.

Young, Alwyn. 2016. "Improved, Nearly Exact, Statistical Inference with Robust and Clustered Covariance Matrices Using Effective Degrees of Freedom Corrections." Unpublished.

Young, Alwyn. 2019. "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results." *Quarterly Journal of Economics* 134 (2): 557–98.