Performance-Based Incentives to Fight Tuberculosis

Evidence from a Randomized Experiment in Northern India

Thomas Bossuroy (World Bank)

Clara Delavallade (IFPRI and University of Cape Town)

Vincent Pons (MIT)

DRAFT VERSION – AUGUST 7, 2014

Abstract

More than nine million people around the world become ill with tuberculosis (TB) each year, making it one of the leading causes of deaths. The biggest challenges to contain the disease are to detect it early and to ensure that patients complete the entire course of the treatment. This article provides experimental evidence from India on the effect of health workers' performance-based incentives on patient detection and treatment default. 74 community health workers catering for more than 2,500 patients were hired by Operation ASHA, an NGO delegated by the Government of India to operate local TB treatment centers, and included in the experiment. For the first six months, they were randomly assigned to receive either a fixed salary or a salary dependent on the number of patients they had detected. In the following six months, they were randomly re-assigned to either a fixed or an incentivized salary scheme, based on the number of patient defaults. Preliminary results point to a large increase of reported detections induced by the provision of detection-based incentives in the first phase. However, the number of defaults also increased over the same time period. This could be due to health workers reallocating their effort towards the rewarded task (early detection) and to the detriment of other non-rewarded activities (treatment compliance), in line with the multitasking theory, or to health workers detecting patients more likely to default. There is no detectable impact of the default-based incentives introduced in the second phase.

1. Introduction

The tuberculosis (TB) epidemic is recognized as one of the key development challenges by the international community. It is directly targeted by the sixth Millennium Development Goal whose part focus is to "halt and begin to reverse the incidence of tuberculosis". The DOTS (Directly Observed Treatment – Short course) system, which stipulates that an independent observer watch the TB-infected patient swallow his or her anti-TB pill three times a week for at least two months, is reported to have led to significant progress in the treatment of TB.[1]

While the technology exists to cure cases of TB, the biggest challenge to contain the spread of the disease is to detect it in a timely manner and ensure that patients complete the entire course of treatment. Stigmatization, lack of information and remoteness of the health system prevent the timely identification of many new TB cases, and the symptoms of the disease often disappear after a few weeks of treatment, considerably weakening health workers and patients' incentives to continue undergoing treatment. Can financial incentives provided to health workers encourage them to detect new tuberculosis cases and prevent patients from defaulting on treatment?

A growing body of evidence documents that poor performance of health care in low- and middle-income countries is not only due to inadequate training or knowledge deficiencies but also to insufficient provider effort, translating into high absenteeism rates (Chaudhury and Hammer 2004, Banerjee et al. 2004) and lack of adequation between knowledge and practice (Das and Hammer 2005; Chaudhury et al. 2006; Das and Hammer 2007). However, whether external incentives are an efficient means to improve service provider effort in the public health sector is more uncertain (Miller and Babiarz, 2013). Offering monetary incentives to workers may indeed induce unintended behavioral reactions and there are several reasons why the public health sector may be more prone to those.

First, the risk that external rewards may crowd out intrinsic motivation and have little impact on effort is higher for workers with a pro-social component to their job (Benabou and Tirole (2003), Tirole and Bénabou 2006). As suggested by the Cognitive Evaluation Theory (Deci, 1975)[2], external incentives may affect individuals' intrinsic motivation by reducing their perceived autonomy (Frey and Goette (1999).[3] Workers with a social mission might perceive financial

[1] The evidence however is mixed on the impact of DOTS as opposed to self-administered treatment. While Kamolratanakul et al. (1999) report a significant increase in cure and treatment rates in Thailand, Vomik (2007), Wally, Newell and Khan (2001), Zwarenstein, Schoeman, Vundule et al. (1998) find no statistically significant difference between the two types of treatment.

[2] The Cognitive Evaluation Theory states that individuals all have fundamental needs for (i) autonomy: "the urge to be a causal agent of one's own life" and (ii) competence: "the urge to master skills and control outcomes".

[3] However, Friedman (2013) argues that performance-based incentives, as opposed to input-based incentives, may have little impact on intrinsic motivation for they also provide information on one's competence.

incentives as a signal of a "money-market" rather than a "social-market" task and reduce their effort (Heyman and Ariely, 2004) or choose their effort level based on the incentives and no longer on their intrinsic motivation: if the incentives are too low, their effort might decrease (Gneezy and Rustichini, 2000).

Second, health workers' job consists in performing several tasks, as opposed to one single repetitive task. Whether incentives work for more complex jobs or induce multitasking remains to be tested.[4] At any point of time, TB health workers are responsible for, at least, detecting new patients, monitoring their existing patients' compliance to treatment, and giving advice to their families to prevent further spread of the disease. Incentivizing one task may prompt them to neglect the dimensions of their task that are not rewarded by incentives, or "non-contracted" outcomes (Miller and Babiarz, 2013).

Third, the impact health workers' effort has on their performance may be loose. Incentives may indeed only be as efficient as individuals are able to adjust their behavior to their goals. However, community health workers may have relatively limited bandwidth for deviating from protocols and innovating in the way they work, as opposed to managers. In addition, the performance targets used here -- TB cases detection and, even more so, default prevention (i.e. adherence to treatment) -- are closely linked to the patient's behavior, leaving less room for health workers' effort to translate into performance (Loevinsohn and Harding, 2005). Low marginal returns to effort may limit the impact of incentives on performance (Miller and Babiarz, 2013).

Using a field experiment conducted in urban slums of Northern India, this study examines whether financial incentives provided to social workers and agents performing multiple tasks do increase performance. 74 TB health workers were randomly assigned to one of four treatment arms. They either received (i) financial incentives based on patient detection for six months and incentives based on treatment adherence subsequently or (ii) financial incentives based on patient detection for six months and a fixed salary subsequently or (iii) a fixed salary for six months and incentives based on treatment adherence subsequently or (iv) a fixed salary for the whole duration of the experiment.

This design allows to evaluate the impact incentives have on a series of outcomes including (i) detection - number of patients enrolled in the DOTS system; (ii) default - number of patients leaving the DOTS system during the course of their treatment; (iii) health workers' effort / motivation and job satisfaction; (iv) patients' characteristics, satisfaction and health status. Outcomes of interest are measured by a combination of administrative data and about 5,000 comprehensive health workers and patient surveys.

[4] Multitasking has been documented mostly theoretically (Holmstrom and Milgrom 1991; Petersen et al. 2006; Mannion and Davies 2008). Empirical evidence is mostly inconclusive (Mullen et al., 2010).

Preliminary results point to a large increase of reported detections due to the provision of detection-based incentives in the first phase, which seems to mostly play out shortly after the introduction of the incentives and dampen in the subsequent months. However, the number of defaults also increased over the same time period, in line with the multitasking theory. Health workers reallocated their effort towards the rewarded task (early detection) and to the detriment of other non-rewarded activities (treatment compliance). There is no detectable impact of the default-based incentives that were introduced in the second phase, probably due to low marginal returns to effort on the default prevention task.

At this preliminary stage of the analysis, uncertainties remain. First, given that we only make use of administrative data to measure our main outcomes, we cannot rule out the possibility that the positive impact on reported detections in the first phase is due to forgery. Doubts will be dispelled once we are able to exploit the monitoring data we collected. Second, the increase in the number of defaults induced by detection-based incentives in the first phase suggests some form of multitasking. But it could also be that incentivized health workers are led to detect patients more likely to default. This uncertainty will be removed using data on patient characteristics from the patient surveys.

Our results relate to the growing literature on performance-based incentives in the public health sector, showing mixed results. While Eichler and Levine (2009) and Gertler and Vermeersch (2013) show a positive impact of performance incentives on health outcomes, in several instances financial incentives have yielded limited effects (Miller and Babiarz, 2013), a positive effect only under certain conditions (Singh 2011; Basinga et al. 2011) or no effect at all (Hillman et al., 1998).

The remainder of the paper is organized as follows. Section 2 lays out the study context. Section 3 presents the experimental design. Section 3 describes the origin of the data. Section 4 presents the population sample and balance checks. Section 5 presents the results, and Section 6 concludes.

2. Background – Current directions in the fight against TB

2.1. The global strategy against TB

The tuberculosis (TB) epidemic is recognized as one of the key development challenges by the international community. It is directly targeted by the sixth Millennium Development Goal whose part focus is to "halt and reverse spread of Tuberculosis". More than nine million people around the world become ill with tuberculosis (TB) each year, making it one of the leading causes of deaths.

Excerpts from the Global Plan to Stop TB 2011-2015 (released by the WHO and the Stop TB Partnership)

Component 1: DOTS expansion and enhancement.

Objective 1: Ensure early diagnosis of all TB cases.

"Diagnosis should be easily accessible, with no or minimal financial and geographic barriers to care". "Access to care needs to be improved through strengthening and expansion of basic health-care services (especially for hard-to-reach populations, as in TBREACH projects). Particular efforts are needed to detect TB in vulnerable groups, which can include pregnant women and young children, the urban poor, contacts of TB cases…" "In addition, National TB Programmes need to establish links and collaborate with the full range of care providers through Public-Private Mixed approaches. There is good evidence that Public-Private Mixed approaches can increase the percentage of people who are diagnosed and receive high-quality treatment by between one quarter and one third, with health care providers such as pharmacists, traditional healers and private practitioners often the first point of contact for people with symptoms of TB."

Objective 2: Ensure high-quality treatment of all diagnosed cases of TB.

"High rates of treatment success also depend on provision of care and support in health care facilities and the community, including use of enablers and incentives where appropriate, effective programme management and supervision and engagement of all care providers through Public-Private Mixed and the International Standards for TB Care (ISTC). Community engagement can improve the quality of care through direct patient support, and can have a very positive and immediate impact on adherence to TB treatment."

Objective 3: Strengthen monitoring and evaluation including impact measurement.

Box 1 - Global Plan to Stop TB 2011-2015 (excerpts)

While TB can be cured in nearly all cases, the biggest challenges to contain the spread of the disease and of Multiple Drug Resistant (MDR) forms are to detect the disease in a timely manner and to ensure that patients undergo and complete the entire course of the simple though stringent treatment. First, stigmatization, lack of information and remoteness of the health system prevent the timely identification of many new TB cases. Second, the treatment requires the patients to go thrice and then once a week to the center for more than six months, while the symptoms of the disease often disappear after a few weeks of treatment. This considerably weakens a patient's incentive to continue to undergo treatment.

To address these difficulties, the WHO defined the DOTS (Directly Observed Treatment – Short course) strategy. This strategy consists in bringing drugs to a large number of small centers, and it stipulates that an independent observer watch the TB patient swallow his or her anti-TB pill three times a week for at least two months.

2.2. The DOTS system in India

India has the highest incidence rate of TB. Each year, nearly 2 million people in India develop TB and about 1,000 Indians die of TB every day, making it the leading infectious cause of death among adults in the country. Even when it is not deadly, the disease considerably weakens the patients: it is a major barrier to social and economic development.

The Indian Ministry of Health has given the priority to the eradication of TB, investing considerable efforts and money through the Revised National Tuberculosis Control Programme (RNTCP): first piloted in 1993, and designed according to the DOTS strategy, it now covers the entire country. Under the RNTCP, DOTS centers are mostly run by government hospitals or by NGOs working on delegation of the State health ministries.

2.3. The challenge of reaching urban slum populations

TB prevalence is higher in urban slums than in the rest of the country. Overcrowding increases contagion, pollution and poverty weaken the bodies and the lack of space or drainage prevents an adequate disposal of sputum. Due to poor education, many fail to take simple preventive measures which would help contain the spread of the disease, and the stigmatization that is still often associated to TB makes its detection harder.

Although the DOTS system has led to significant progress in the treatment of TB, it remains insufficient to address the challenge of curing efficiently the most remote populations. Government dispensaries which implement the DOTS model are often located far away from the slums, making it difficult and costly for TB patients there to get detected and then comply with the entire treatment. There is a blaring lack of access to health infrastructures, and of information about the detection and treatment of the disease.

The lack of interaction with health workers makes many people ignorant of the need to get detected, the process of taking a sputum test, and the possibility to get cured in a DOTS center, free of cost. Detection is therefore very low. Further, even when they get detected and enrolled in the DOTS system, slum inhabitants' poor education makes them more prone to defaulting on the treatment, and therefore developing (and spreading) Multiple Drug Resistant forms of TB (MDR-TB), which require extensive and costly chemotherapy to treat and are therefore incurable for them. MDR-TB is increasingly acknowledged as a looming threat on public health, especially in urban areas where those new forms of the disease could quickly be out of control.

To address these difficulties and "reach marginalized sections of the society", the RNTCP has now entered a stage of deepening. "Improving the case finding through an effective patient-centered approach to reach all patients, especially the poor" is listed first on their agenda which also highlights the need for "scaling-up of community TB care [and] creating demand through context-specific advocacy, communication and social mobilization" (RNTCP's webpage: www.tbcindia.org/RNTCP.asp). This includes a reflection on how to better manage DOTS centers and on the use of Public-Private Partnerships (PPPs) to have a higher efficiency and deeper reach in remote places.

2.4. Operation ASHA's programs

Operation ASHA, a Delhi-based NGO and well-established actor of the fight against TB in India, specifically targets urban slums. It has established a network of more than 100 community-based DOTS centers in five states (Delhi, Uttar Pradesh, Madhya Pradesh, Punjab, and Rajasthan) run on delegation of the State health ministries. These centers are located in small shops, pharmacies, or temples, they open at convenient hours, and they each cover a small neighborhood to minimize the distance between an average patient's house and his or her center. Each center is supervised by a community health worker (CHW), who delivers information to the community, engages in the detection of new patients through widespread community testing, and tracks patients enrolled in the center who have missed a pill to bring them back onto the regular course of treatment.

The community health workers play a critical role in improving the access to TB treatment in remote communities. When health workers are allocated to a center, their primary objective is to increase the detection in the area of coverage and to bring more patients onto the treatment. After the number of detections has reached a plateau, and before the center becomes too crowded to be manageable, the health workers need to focus on making sure that patients comply with the treatment and complete it. However, as in any organization working in remote areas, monitoring the health workers is a challenge. For instance, several studies have shown that attendance and commitment were often very low in the government health and education system in remote places. To circumvent this major issue, Operation ASHA, in collaboration with the research team, has designed an original compensation scheme aimed at improving their health workers' motivation to complete their important tasks properly and, thereby, their efficiency.

3. Experimental Design

We evaluate the impact of an enhanced Directly Observed Treatment – Short course model in which health workers are offered performance-based incentives on health workers' performance and motivation. CHWs are offered incentives based first on the number of detections of TB infected individuals, and then on the rates of default of the patients of their centers.

3.1. The salary schemes and randomization design

Operation ASHA hires community-based health workers, who are each responsible for operating two DOTS centers. During the first three months, CHWs all receive a fixed salary. Indeed, centers often take a slow start, irrespective of the involvement of the CHW. This initial 3-month period also enables the CHWs to get their bearings in a new place and for the center to be identified by the population. The experiment starts after the initial three months of a center lifespan. Between 3 and approximately 9 months, CHWs have to grow their center until they have reached the optimal size (not more than 50 patients), where they are cost-effective and where patients can be effectively followed-up. Half of the CHWs, randomly chosen, receive a fixed component and a variable amount based on their performance regarding detection of new patients (see definition in Box 2), while the other half receives a fixed salary.

Operational definition of a "new patient"

The following cases count as "new patients" for whom a counselor gets incentives:

1/ a person detected on the field by the counselor during counseling, referred to the DMC for a sputum test, tests positive and is then sent back to the Op. ASHA center with his treatment box and starts treatment from the center.

2/ a person who is referred from the DMC with a treatment box to take medicines from the Op. ASHA center and might have self reported for testing at the DMC or been detected by an Op. ASHA counselor, or any other NGO/ Govt. worker

3/ a patient transferred from another center and starts treatment on that month from the Op.ASHA center.

The following cases should not be counted as "new patients":

1/ a person detected by the counselor on the field sent to the DMC for a sputum test, tests positive but taking medicines from the DMC or another DOTS center.

2/ a person referred from the DMC to take medicines at the Op. ASHA center but not in the first month of his treatment.

Policy for incentive calculation during the leave period

For a counselor getting a salary based on incentives for detection;

- If the counselor takes a break/ leave for a period of 6 working days or less in a month, then he gets incentives for all the patients detected and started on treatment at his center whether detected by him or the leave reserve counselor.

- If the counselor takes a break/leave for a period of 7 days or more then he only gets incentives for the patients that were detected and started on treatment in his presence, and this number is scaled up for the rest of the month.

Box 2 – What is a "new patient"?

After nine months of work, CHWs should focus on preventing defaults. The CHW's compensation scheme is randomized again. Half of them receive a fixed salary for the following 6 months of their contract while the other half receive an incentivized scheme where the variable component no longer depends on the number of detections, but on the number of defaults they prevent. If the number of patients keeps growing, Operation ASHA opens a new center and the detection work is taken over by another CHW operating that newly open center.

These incentives (for detection) or penalties (for default) come in addition to a base salary, that guarantees the health workers that they will get a minimum amount for their work, whatever happens, and thereby contains the amount of risk and stress they are facing. The introduction of financial incentives explicitly linked to the outcome of their counselling work is aimed at increasing their motivation, effort, and performance, and in turn their impact on TB treatment in slums.

In brief, CHWs are randomly assigned to either a contract with fixed salary or a contract with a part linked to the number of detections between three and nine months, and one of the two types of contract 2 after nine months. Because they are re-randomized between the two phases, they end up being randomly allocated to one of the four groups described in Table 1 below.

		Incentives in second (default) phase?	
		Yes	**No**
Incentives in first (detection) phase?	**Yes**	**Group 1:** - Incentives for detection in phase 1 - Incentives for default in phase 2	**Group 2:** - Fixed salary in phase 1 - Incentives for default in phase 2
	No	**Group 3:** - Incentives for detection in phase 1 - Fixed salary in phase 2	**Group 4 (“pure control”):** - Fixed salary in phase 1 - Fixe salary in phase 2

Table 1 – Treatment arms

3.2. Design features

The structure of the incentive contracts was designed collaboratively by Operation Asha and the research team in order to minimize different possible issues.

Balancing risk and incentives. Performance incentives should ideally balance fixed and variable pay (Miller and Singer Babiarz, 2013). On one hand, the variable part must be sufficiently large to influence provider behavior (Hall and Liebman 1998), especially in health care (Hillman et al. 1998; Rosenthal and Frank 2006). On the other hand, the financial risk borne by providers increases with the variable part (Ellis and McGuire 1990). The variable part of the contract was thus limited to 25 percent of its total value.

Preventing efficiency wage effects. The computation of the variable part of contracts with incentives is done using baseline data in each city to equalize, on average, incentivized salaries and fixed salaries. This is meant to isolate the effect of incentives from any income or efficiency wage effect.

- Fixed component = 75 percent of the total fixed salary

- Variable component = computed based on the baseline data so that the total salary would equal the fixed salary in expectation.

Table 2 provides the exact parameters for salary calculation in each of the cities where the experiment took place. Amounts vary across cities for two reasons: i) so as to take into account the differences in costs of living across cities; ii) so as to reflect the relative difficulties in detecting patients or preventing defaults, based on city-specific historical data.

Preventing differential attrition of CHWs. When they are hired, CHWs are informed that their salary scheme will be randomly chosen after three months (and then after nine months), but that everyone will get the same amount on average. They are of course free to leave at this point, but this does not threaten the internal validity of our experiment. The fact that they are already three months into the job when they get the first incentive scheme, and were warned that this may happen, prevents differential drop-outs between the control and treatment groups. This will help prevent classical flaws identified by the theory of contracts such as self-selection. Further, each CHW is told about the type of contract that he/she was randomly assigned to just before this contract is implemented, so that her performance would not be affected by the anticipation of a future contract.

			DETECTION PHASE		**DEFAULT PHASE**	
		Fixed	**Base salary**	**Detection incentives**	**Base salary**	**Default incentives**
UP	Moradabad	3000	2250	100 per new patient	2250	1050 if 0 default; 750 if 1 default; 450 if 2 defaults; 150 if 3 defaults; 0 if ≥ 4 defaults
Punjab	Ludhiana	4000	3000	150 per new patient	3000	1400 if 0 default; 1000 if 1 default; 600 if 2 defaults; 200 if 3 defaults; 0 if ≥ 4 defaults
Punjab	Amritsar	4000	3000	150 per new patient	3000	1400 if 0 default; 1000 if 1 default; 600 if 2 defaults; 200 if 3 defaults; 0 if ≥ 4 defaults
Punjab	Jalandhar	4000	3000	150 per new patient	3000	1400 if 0 default; 1000 if 1 default; 600 if 2 defaults; 200 if 3 defaults; 0 if ≥ 4 defaults
MP	Bhopal	3000	2250	175 per new patient	2625	1225 if 0 default; 875 if 1 default; 525 if 2 defaults; 175 if 3 defaults; 0 if ≥ 4 defaults

MP	Gwalior	3000	2250	100 per new patient	2250	1050 if 0 default; 750 if 1 default; 450 if 2 defaults; 150 if 3 defaults; 0 if ≥ 4 defaults
MP	Jabalpur	3000	2250	175 per new patient	2250	1050 if 0 default;750 if 1 default; 450 if 2 defaults; 150 if 3 defaults; 0 if ≥ 4 defaults
MP	Indore	3500	2625	219 per new patient	2625	1225 if 0 default; 875 if 1 default; 525 if 2 defaults; 175 if 3 defaults; 0 if ≥ 4 defaults
MP	Sagar	3000	2250	150 per new patient	2250	1050 if 0 default; 750 if 1 default; 450 if 2 defaults; 150 if 3 defaults; 0 if ≥ 4 defaults
Chhattisgarh	Durg/Bhilai	4000	3000	250 per new patient	3000	1400 if 0 default; 1000 if 1 default; 600 if 2 defaults; 200 if 3 defaults; 0 if ≥ 4 defaults
Chhattisgarh	Korba	3000	2625	215 per new patient	2625	1225 if 0 default; 875 if 1 default; 525 if 2 defaults; 175 if 3 defaults; 0 if ≥ 4 defaults
Chhattisgarh	Bilaspur	3000	2625	215 per new patient	2625	1225 if 0 default; 875 if 1 default; 525 if 2 defaults; 175 if 3 defaults; 0 if ≥ 4 defaults
Delhi	Delhi		N/A		2550	1190 if 0 default; 850 if 1 default; 510 if 2 defaults; 170 if 3 defaults; 0 if ≥ 4 defaults

Table 2 - Salary schemes per phase per city

Dealing with the selection bias of patients. The characteristics of patients in each center might depend heavily on the efforts made by the CHWs to go and find the ones who would not have shown up otherwise. The first incentive scheme thus lays the ground for a selection bias in the analysis of defaults. However, the cross-randomization will allow us not only to control for this possible bias, but also to estimate its importance. The data from the patient survey will provide detailed information on the characteristics of the patients detected, and we will test whether incentives enable CHWs to find patients who are usually left out of the system and are more likely to default.

Preventing contamination. The DOTS centers are located in non-overlapping and scattered areas and the CHWs are permanently assigned to their catchment areas. This limits the scope for interactions between CHWs and spill-over effects.

4. Data sources

4.1. Program and administrative data

Salary sheets and centerwise reports were collected monthly from Operation ASHA. They were verified and used by the research team for salary computation. They record monthly detections and defaults as well as various salary components (base, incentives, allowances...) for each center and health worker.

TB registers and lab registers are kept by public health TB officers, who centralize the treatment cards generated by all centers in the area. These registers list the name and address of all enrolled patients, the dates and results of their initial and follow-up sputum tests and the outcome of the treatment. This data is not subject to forgery by the CHWs, unless they collude with senior government officials, which is unlikely. They thus provide a reliable measure of the main two outcomes of the study: detection and default.

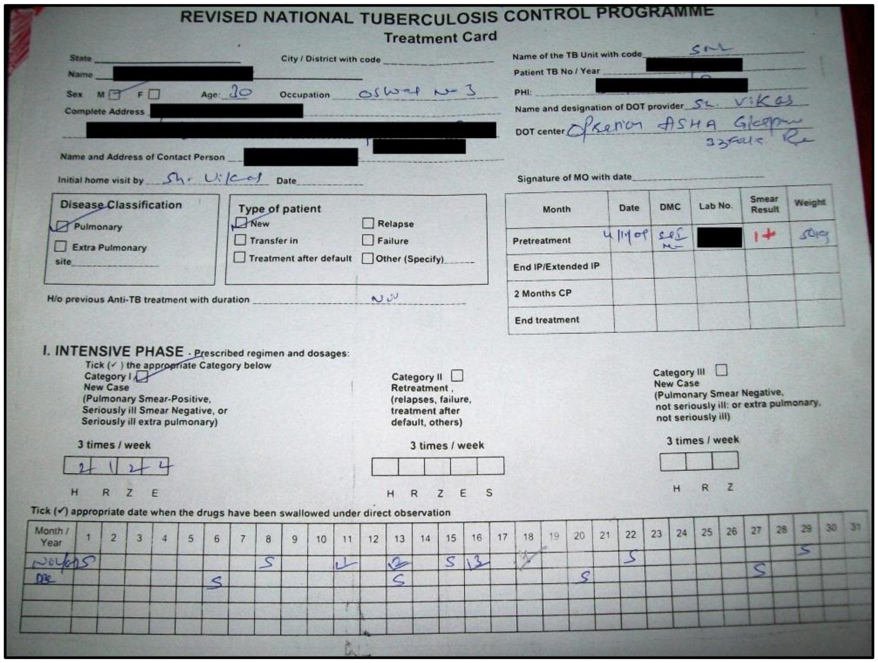
Treatment cards are generated by Operation ASHA for each patient enrolling in a DOTS center (see Figure 1). They are a second source of information about the number and identity of newly detected patients month after month. In addition to this, they give daily information about the pills taken by the patients and their sputum tests. If CHWs who get default-related incentives are able to actually reduce the number of defaults, these data will provide additional evidence on the strategy they use to do so: do they take action to any missing pill, do they wait until 2 or 3 pills or more are missed? To this end, all treatment cards were collected and entered in a specially designed dataset.

However, the treatment cards are filled by the CHWs themselves, so they may be subject to forgery, especially by incentivized CHWs. The fact that the incentives are not directly related to the number of missed pills but to the number of final defaults reduces the risk.

Dataset Name	Description	Module	Cities
Salary sheets and centerwise reports	Operation ASHA's administrative files. Contains health worker salary for each month of the experiment, along with detection/default data as reported by OA	Salary sheets and centerwise summary sheets	ALL
Treatment Cards	Contains all patient treatment information based on pictures of Operation Asha treatment cards	N/A	**Available cities:** Bhopal, Indore, Gwalior, Sagar, Durg/Bhilai, Korba, Moradabad; **Not Available Cities:** Jabalpur, Bilaspur, Amritsar, Ludhiana, Jalandar
PII	Contains all personally identifiable information of patients in the experiment sample, including surveyed and non-surveyed patients. Is kept separate for ethical reasons.	Surveyed and Non-Surveyed Patients	ALL

Patient Surveys	Contains the following modules: Personal, Work, Inactivity-Unemployment, Family TB history, Children, Current health, Healthcare, Vaccination and past TB, Detection, Current treatment, Post-treatment, Interaction with health worker, TB knowledge, Social insertion, Optimism-happiness, Tobacco, Borrowing-saving, Assets, Wealth and sanitation, Consumption, Measures	Adult Entry, Exit, Exit+, Child Entry, Exit, Exit+	ALL
Health worker Surveys	Contains the following modules: Personal information, Family information, Health, Assets, Income generating activities, HH income, Operation ASHA centers, Detection ativities, Default activities, Expectations, Job satisfaction, Job termination	Baseline, Midline, Midline+, Endline, Endline+	ALL
Monitoring Data	Contains center and health worker monitoring data	N/A	ALL
TC Tracker	Contains a subset of patient information copied from Operation Asha treatment cards during center/health worker monitoring; data reconciled with official administrative data (TB Registers)	N/A	**All data:** All - Amritsar, Ludhiana, Jalandar, Bhopal, Jabalpur, Indore, Gwalior, Sagar, Durg/Bhilai, Bilaspur, Korba, Moradabad; **Reconciled Data:** Not available for Jabalpur, Amritsar, Ludhiana, Jalandar
Backcheck Data	Contains patient survey backcheck data of 10 percent of sample, and completion status verification of non-surveyed patients	N/A	ALL

Table 3 – Data sources and description



REVISED NATIONAL TUBERCULOSIS CONTROL PROGRAMME
Treatment Card

State ____ City / District with code ____ Name of the TB Unit with code ____
Name [redacted]
Patient TB No / Year [redacted]
Sex M ☐ F ☐ Age: 20 Occupation ____ PHI: [redacted]
Complete Address [redacted]
Name and designation of DOT provider Sh. Vikas
DOT center ____
Name and Address of Contact Person [redacted]
Initial home visit by Sh. Vikas Date ____
Signature of MO with date ____

Disease Classification
☐ Pulmonary
☐ Extra Pulmonary
site ____

Type of patient
☐ New ☐ Relapse
☐ Transfer in ☐ Failure
☐ Treatment after default ☐ Other (Specify) ____

H/o previous Anti-TB treatment with duration ____

Month	Date	DMC	Lab No.	Smear Result	Weight
Pretreatment	4/11/08		[redacted]	1+	
End IP/Extended IP					
2 Months CP					
End treatment					

I. INTENSIVE PHASE - Prescribed regimen and dosages:
Tick (✓) the appropriate Category below

Category I ☐
New Case
(Pulmonary Smear-Positive, Seriously ill Smear Negative, or Seriously ill extra pulmonary)
3 times / week
H R Z E

Category II ☐
Retreatment,
(relapses, failure, treatment after default, others)
3 times / week
H R Z E S

Category III ☐
New Case
(Pulmonary Smear Negative, not seriously ill: or extra pulmonary, not seriously ill)
3 times / week
H R Z

Tick (✓) appropriate date when the drugs have been swallowed under direct observation

Month / Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31

Figure 1 – Example of a Treatment Card

4.2. Survey data

Health worker surveys are administered thrice, before the first salary randomization (baseline), before the second randomization (midline) and at the end of the default-based incentives period (endline). These surveys provide detailed information on the health workers' socio-economic background, their motivation when starting their job, and their work history. They also provide information on their daily work activities. We finally measure their satisfaction at work and the level of trust and cooperation they feel within the organization.

Patient surveys were collected all over the course of the three-year experiment, usually once at the beginning of their treatment and once towards the end. A total of 2791 patients were surveyed. We collect socio-economic characteristics, GPS data (to control for the distance to the center), and measure additional outcomes of the programs: health symptoms, daily activities, and awareness of hygiene behavior meant at reducing the risk of further contamination. This data provides evidence about the extent to which the intervention impacted the patients' lives and whether it improved the detection and treatment of the most remote, uneducated and deprived segments of the population.

The surveys are collected at two points for each patient: one at the start of the treatment, one after they complete the treatment. A patient who started treatment before the beginning of survey operations would be administered an augmented endline survey that includes some of the socio-economic characteristics normally collected at entry. The questionnaire, protocol and consent form were also adapted for children.

Potential respondents were identified through a sampling process which was carried out each time a designated enumerator visited an Operation Asha center for counselor and patient monitoring. At each visit, three sampling ranges were calculated that identified patients who were beginning treatment (Range 1), patients who had completed six months of treatment (Range 2) and patients who either defaulted or transferred out from the center (Range 3). Patient details— such as name, address, and treatment start dates— were collected from treatment cards maintained by Operation Asha counselors, and each patient was assigned an unique identification code. Range 1 patients were administered a baseline survey, while Range 2 and 3 patients were given the endline survey, provided they had received a baseline. A modified version of the endline survey was administered to patients that did not receive a baseline.

After a sample of patients was determined, appointment sheets were prepared for each patient to be interviewed. These contained details that would help locate and identify the patients (such as names and addresses), as well as information on whether not the visit was successful. Three attempts were made to locate the patient before they were declared “Not Found”. In a successful appointment, the surveyor administered an entry or exit survey to the patient, depending on their stage of treatment.

Backchecks were performed by revisiting households in our sample (both those surveyed and those reported as ‘not found’ or ‘unable to be surveyed’) to verify that they were visited by a JPAL enumerator and that they completed the survey correctly. Each city was visited approximately once a month by a designated backchecker. When re-visiting a household/patient, this person would administer a shorter version (approximately 15 questions) of the patient survey. The back-check questionnaire answers were then compared with the original survey data and inconsistencies were investigated.

The backcheck sample was determined by the outcome of the survey. For all surveys that were successfully completed, the backchecker would re-visit approximately 15% of the patients (randomly selected). For surveys that could not be completed, re-visits were done for each of these cases to confirm the reported cause behind the non-completion of the survey.

4.3. Monitoring data

Monitoring data come from visits conducted by survey staff at Operation Asha centers to observe counselor attendance, observe the number of patients visiting the center for DOTS, and to collect data from patient treatment cards. Table 4 shows the average number of visits performed per center in each city. All data was entered into a dataset. The procedure consisted of a designated field team member visiting a pre-assigned center each day. This person would stay at the center from the time it opened in the morning till the counselor finished his work, around early afternoon. During his/her time there, a systematic record of the following were maintained for each visit:

- Whether the center is open and if the counselor is present
- Arrival and departure time of the counselor

- Patient name, along with administrative details such as tb number/lab number and treatment start date
- Cases where a relative picked up medicines for the patient
- Cases where the counselor was planning to visit the patient personally

State	City	Average visits per center per month[5]
Madhya Pradesh	Bhopal	0.64
	Jabalpur	0.66
	Gwalior	0.67
	Indore	0.75
	Sagar	1
Chhattisgarh	Korba	0.75
	Durg	0.92
	Bilaspur	0.86
Punjab	Amritsar	0.58
	Ludhiana	1
	Jalandhar	1
Delhi	South Delhi	0.5
Uttar Pradesh	Moradabad	1.02

Table 4 – Number of monitoring days per center per month

The **Treatment Card Tracker** (TC Tracker) was maintained by survey staff for all the patients in the centers that the experiment ran in, and subsequently entered into a dataset. This document helped to keep track of patients who received treatment, their TB number, date of start of treatment, their test dates and the entry and exit codes that were assigned to them if they were surveyed. Most of this information was captured from the Treatment Cards maintained by the Operation ASHA counselor and was updated on a regular basis by J-PAL enumerators. The TC tracker data will be used to verify patients against the government records maintained in every city, thereby helping to identify treatment outcomes and cases of forgery.

5. Sample Description and Balance Checks

5.1. Health workers

A total of 105 health workers were enrolled in the study, but the significant turnover rate of health workers created attrition in the experimental sample. Only 90 stayed for at least a month, and 72 completed the full course of the 12-month experiment.

[5] Calculated as a simple average of total visits made/(total no: of centers*total no: of months of data collection)

State	City	Number of Health workers enrolled for at least one month
Madhya Pradesh	Bhopal	16
	Jabalpur	18
	Gwalior	13
	Indore	9
	Sagar	4
Chattisgarh	Durg	4
	Korba	4
	Bilaspur	3
Punjab	Ludhiana	4
	Amritsar	1
	Jalandar	1
Uttar Pradesh	Moradabad	6
Delhi	South Delhi	7
TOTAL		**90**

Table 5 – Geographical distribution of the participating health workers

Table 5 and Figure 2 report the geographical distribution of the 90 health workers who were enrolled for at least a month.

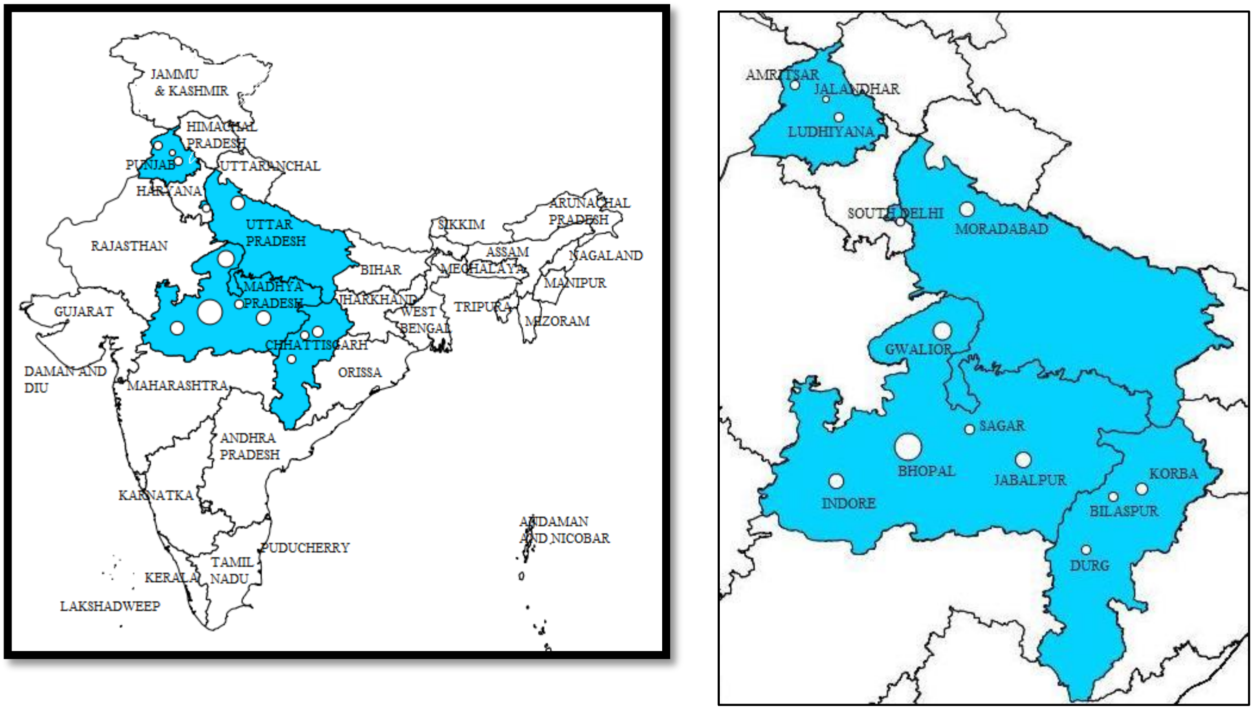


Figure 2 – Maps of study sites and sample sizes

Table 6 presents summary statistics for the community health workers. The average health worker is 32 years old. A majority (72 percent) of the health workers is made of males, 41 percent belong to general castes and 80 percent are Hindus. The health workers are well-educated on average: half have spent some time at the university and only 4 percent did not complete class 10. Three quarters of them have some previous work experience, for an average number of 8 years of previous work experience, and 16 percent of them have previous work experience in the social or NGO sector. Only 37 percent of the health workers live in a neighborhood where one of the centers that they operate is located. They mostly live in decent conditions: almost all of them have electricity at their place and two thirds have access to tap water. More than half own land, and 11 percent rent some apartment or house to a third party.

The differences between the control and treatment groups for the first-phase randomization are significant at the 10 percent level for 4 out of these 22 variables, and they are significant at the 5 percent level for two variables (Panel A). The differences between the control and treatment groups for the second-phase randomization are significant at the 10 percent level for 1 out of these variables (Panel B). Overall, only 5 differences out of 44 are significant at the 10 percent level and 2 are significant at the 5 percent level, as should be expected.

Table 6 includes only a small subset of the very detailed and comprehensive information gathered on the health workers' demographics and on their preexisting work experience. On top of this information, the baseline survey also asked the health workers to describe the different dimensions of their work, including detection and default-prevention activities. Let us briefly summarize their answers, as these two outcomes are the main focus of the results section.

At baseline, the three tasks that are most commonly highlighted by the health workers as one of their main tasks are: identify people potentially infected by TB and verbally tell them to get a sputum exam for detection and treatment (mentioned by 84 percent of the health workers) ; deliver treatment at the DOTS center (73 percent); visit patients when they miss a pill (38 percent).

Now focusing on the activities related to the detection of new patients, we asked the health workers how they proceed to identify prospective TB patients. 30 percent of them mentioned that they conduct door to door visits in the neighborhood and identify new patients on the basis of their symptoms and 36 percent answered that they ask neighbours about any sick people and follow their directions. In addition, the vast majority of health workers who undertake some detection activities mention that they go every weekday to the field for that purpose. Similarly, two thirds of the health workers mention that they go every day to the field to visit ongoing or CP patients. In sum, at baseline the health workers clearly state that their mandate involves spending a lot of time on the field rather than just providing medicine at the treatment center.

But default-preventing activities go beyond visiting ongoing patients who have missed a pill: first, respectively half and one fourth of the health workers report that when a patient is assigned to taking medicine at one of their centers, they give them advice about treatment at the center or at the patient's home. In addition, 28 percent of the health workers mention that they sometime call patients who have missed a pill on their phone. Finally, when a patient misses pills for a period longer than a month, 29 percent of health workers report going to their house and counseling them personally while 21 percent inform public health workers and ask them to counsel them.

	Control group		Treatment group		*P-value* Treatment = Control	Number of obs.
	Mean	SD	Mean	SD		
Panel A. First-phase randomization						
Male	0.71	0.46	0.74	0.44	0.78	74
Age	32.5	7.4	31.8	7.4	0.69	74
General caste	0.31	0.47	0.51	0.51	0.09	74
Hindu	0.80	0.41	0.79	0.41	0.96	74
Highest education level achieved						
Below class 10	0.03	0.17	0.05	0.22	0.63	74
Class 10	0.11	0.32	0.08	0.27	0.59	74
Class 11	0.03	0.17	0.03	0.16	0.94	74
Class 12	0.29	0.46	0.33	0.48	0.66	74
Professional course	0.03	0.17	0.03	0.16	0.94	74
First year of university	0.09	0.28	0.03	0.16	0.26	74
Second year of university	0.00	0.00	0.13	0.34	0.03	74

Third year of university	0.11	0.32	0.08	0.27	0.59	74
Graduation	0.29	0.46	0.21	0.41	0.43	74
Non formal / adult education	0.03	0.17	0.00	0.00	0.29	74
Work experience						
Any previous work experience	0.74	0.44	0.77	0.43	0.80	74
Number of years of work experience	8.0	7.0	6.8	5.4	0.44	71
Any previous experience in the social / NGO sector	0.06	0.24	0.26	0.44	0.02	74
Lives in one of the areas covered by the centers	0.30	0.47	0.43	0.50	0.27	70
Assets						
Has electricty	0.94	0.24	0.97	0.16	0.50	74
Has tap water	0.77	0.43	0.57	0.50	0.07	72
Rents an apartment or house to a third party	0.11	0.32	0.10	0.31	0.87	74
Owns land	0.51	0.51	0.59	0.50	0.50	72

Notes: For each variable, we report the means and standard deviations in both the control group and the treatment group and indicate the p-value of the difference. The unit of observation is the health worker.

Table 6 – Summary statistics for the community health workers

	Control group		Treatment group		*P-value* Treatment = Control	Number of obs.
	Mean	SD	Mean	SD		
Panel B. Second-phase randomization						
Male	0.64	0.49	0.67	0.47	0.73	76
Age	32.7	6.8	33.0	9.0	0.88	76
General caste	0.33	0.48	0.33	0.47	0.94	76
Hindu	0.73	0.45	0.84	0.37	0.25	76
Highest education level achieved						
Below class 10	0.03	0.17	0.05	0.21	0.72	76
Class 10	0.18	0.39	0.07	0.26	0.14	76
Class 11	0.03	0.17	0.05	0.21	0.72	76
Class 12	0.24	0.44	0.37	0.49	0.23	76
Professional course	0.03	0.17	0.02	0.15	0.85	76
First year of university	0.09	0.29	0.05	0.21	0.45	76
Second year of university	0.06	0.24	0.07	0.26	0.88	76
Third year of university	0.09	0.29	0.07	0.26	0.74	76
Graduation	0.21	0.42	0.21	0.41	0.98	76
Non formal / adult education	0.00	0.00	0.02	0.15	0.38	76
Work experience						
Any previous work experience	0.73	0.45	0.77	0.43	0.69	76
Number of years of work experience	7.2	6.1	9.1	6.6	0.20	73
Any previous experience in the social / NGO sector	0.09	0.29	0.19	0.39	0.25	76
Lives in one of the areas covered by the centers	0.40	0.50	0.40	0.50	0.97	72
Assets						
Has electricty	0.97	0.17	0.98	0.15	0.85	76
Has tap water	0.76	0.44	0.56	0.50	0.07	76
Rents an apartment or house to a third party	0.18	0.39	0.09	0.29	0.26	76
Owns land	0.50	0.51	0.50	0.51	1.00	74

Notes: For each variable, we report the means and standard deviations in both the control group and the treatment group and indicate the p-value of the difference. The unit of observation is the health worker.

Table 6 (cont.) – Summary statistics for the community health workers

5.2. Patients

A total of 2760 patients were surveyed over the course of the experiment, representing the vast majority of TB patients detected and enrolled in the areas placed by the Government under the responsibility of Operation ASHA. The patient surveys represent a considerable amount of information on the demographic and socio-economic characteristics of TB infected people in urban slums, their pathways to detection and treatment, their attitudes towards the public health system and the DOTS treatment, their satisfaction with the healthcare services provided by Operation ASHA, and the stigma they face because of their illness. When the process of merging the different datasets will be completed, the impact of the provision of performance-based incentives will be analyzed from the patient's perspectives. At this point, it seems appropriate to provide a description of the main characteristics of the population studied and give an example of the possible analyses that this unique dataset makes possible.

Table 7 provides a description of the socio-demographic characteristics of those patients. Men and women are almost equally represented, although there is a slight majority of men (57.7 percent). They belong to the most deprived castes in India: the Scheduled Castes, or Dalits (26.6 percent) and Other Backward Classes (40.2 percent). Only less than 20 percent belong to the better-off General category, which does not receive any government benefit. By comparison, 41 percent of health workers reported belonging to a caste classified under the General category. Patients are 76.1 percent Hindu.

Gender (n=2748)	%	**Education (n=2723)**	%
Female	42.3	No school	21.6
Male	57.7	Some primary	36.0
		Primary completed	8.6
Caste category (n=2580)	%	Secondary (completed or not)	22.4
General	19.6	Pre-university or more	11.4
OBC	40.2		
SC	26.6	**Size of household (n=2710)**	
ST	7.0	Mean	5.0
Minority	6.6	St Dev	3.4
		Median	4
Religion (n=2759)	%		
Hindu	76.1	**Migration status (n=2760)**	%
Muslim	21.7	Always lived here	59.2
Other	2.2	Lived here for more than 10 years	8.3
		Lived here for 6 to 10 years	9.1
Literacy (n=2757)	%	Lived here for 1 to 5 years	15.8
Cannot read or write	66.6	Lived here for less than a year	7.6
Can read and write	30.0		
Can read but not write	3.5		

Table 7 – Socio-demographic characteristics of patients

More than half of the sample has either never been to school or has stopped before completing primary school. Two thirds are illiterate according to their own reporting. Patients come from households of 5.5 individuals on average, although the large standard deviation and the large difference between the average and the median sizes point to the existence of a fraction of very large households in our sample. The majority of patients have been living in their community ever since they were born, but a sizable share of about a quarter (23.4 percent) are recent migrants who report living in the area for less than 5 years.

A large number of TB patients in the study sites have had early exposure to TB (see Table 8). Almost half of them (46.1 percent) report having seen at least one family member infected by TB since they were born. In 22 percent of cases, patients themselves had already had TB in the past, more than once for a small share but significant number of them (85 individuals). Three quarters of patients (73 percent) declare having received the BCG vaccine that protects them against TB, and the enumerators were indeed able to observe the mark left on their arm in a vast majority of cases.

Vaccinated against BCG (n=2437)	%	**If yes, how many times (n=621)**	%
No	25.1	1	86.3
Yes	73.2	2	8.5
Received unknown vaccine	1.8	3+	5.2
If vaccinated, mark visible (n=1820)	%	**Number of family members who had TB since respondent born (n=2729)**	%
No	15.0	0	53.9
Yes	82.6	1	29.8
Will not show	2.4	2	9.9
		3+	6.4
Has previously had TB (n=2752)	%		
No	77.3		
Yes	22.7		

Table 8 – Past exposure to TB

So as to understand the healthcare practices of the patients, a number of questions were asked about the last time they went to consult for a health-related problem (see Table 9). A total of close to 500 patients reported having consulted in the past three months. In spite of their lack of resources, they chose to consult at a private facility in two thirds of the cases – either a private doctor (more than 40 percent of consultations) or a private hospital. Only a quarter of all consultations were done using the public health system, be it a Government hospital or a Government doctor.

Have you consulted in the past 3 months? (n=2760)	%
No	81.92
Yes	18.08
What facility did you go to? (n=499)	%
Private doctor	41.88
Private hospital	23.45
Govt. referral hospital	14.23

Govt. doctor	12.02
Other	8.41
Received… (n=470)	%
Medications	92.18
An injection	39.57
A drip	17.02
Average amount spent (n=499)	Rs
Median	180
Average	595

Table 9 – Previous medical consultations

In almost all cases, patients were prescribed or received medications during that consultation. More surprisingly, but in line with previous studies on healthcare in India, a large number of them also received an injection (almost 40 percent), while 17 percent were given a drip. The median amount spent on the consultation and associated treatment purchased is very large at 180 Indian rupees (seven to ten times the price of a basic meal). The average amount is even much larger, driven by a small number of extremely high expenses.

6. Results

This section discusses the main findings obtained so far. The first subsection presents the impact of the incentives on the outcome that was rewarded: the number of newly detected patients (in phase 1) and the number of defaulting patients (in phase 2). In addition, we measure the impact of the incentives on the outcome that was NOT rewarded (the number of defaulting patients in phase 1 and the number of newly detected patients in phase 2) but may also have been affected: for instance, the CHWs may have decreased their effort on alternative dimensions of their work that were not rewarded by the incentives, due to multitasking. While both the number of detections and the number of defaults are measured using administrative data obtained from Operation ASHA, the second subsection focuses on outcomes measured through CHWs' self-reports: their reported actions undertaken to detect new patients and to prevent the default from existing patients as well as their job satisfaction.

6.1. Impact on performance

6.1.1. Impact of detection-based incentives on <u>detections</u> (phase 1)

Phase 1 incentives were computed based on the number of new patients who were detected within the month. To begin with, we examine the impact of the incentives on this outcome. We compute the average number of newly detected patients per CHW month by month, separately for the treatment and control groups, and starting with the first month following the randomization. As shown in Figure 1a, the average number of detections increases in both the control and the treatment groups as time passes, reflecting the generic growth of the treatment centers. In the treatment group, the number of detections is initially larger. However, by the third month, the difference with the control group is negligible.

To investigate the statistical significance of the overall difference between the number of detections in the control and treatment groups more systematically, we estimate the following OLS regression:

$$D_{i,t} = \alpha + \beta T_i + X_i' \lambda + \sum_c \delta_c^i + \varepsilon_{i,t}$$

where $D_{i,t}$ is the number of new detections in month t in the treatment centers operated by health worker *i*, T_i is a dummy equal to 1 if worker *i* receives phase 1 detection incentives and 0 otherwise, δ_c^i are city fixed effects (the level of stratification), and X_i is a vector of health worker characteristics. X_i includes the number of years of previous work experience and dummies indicating the health worker's education level, whether he has any prior work experience at all, and if so, whether it was in the social sector, and whether he lives in a neighborhood where one of the centers he operates is located. The key coefficient of interest is β, which estimates the differential number of detections by health workers in the treatment group. In this and all other regressions, we adjust standard errors for clustering at the health worker level since the randomization was conducted at this level.

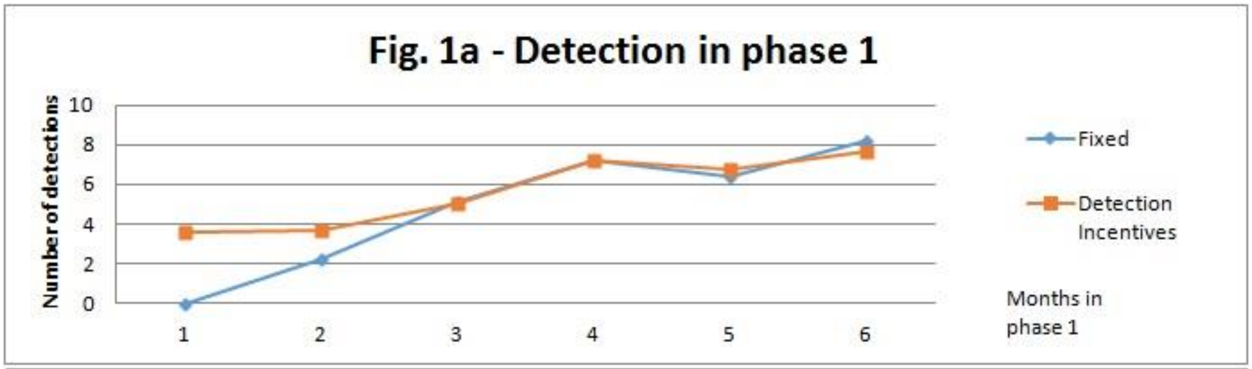


Fig. 1a - Detection in phase 1

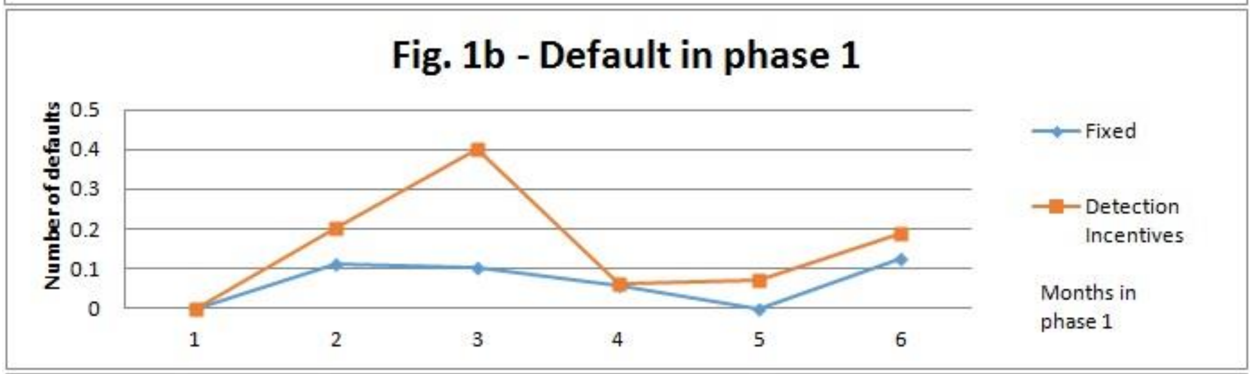


Fig. 1b - Default in phase 1

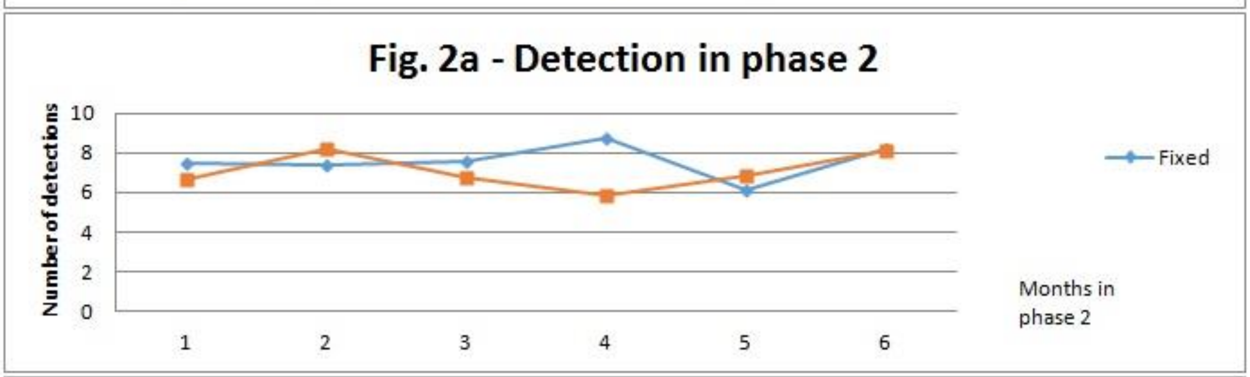


Fig. 2a - Detection in phase 2

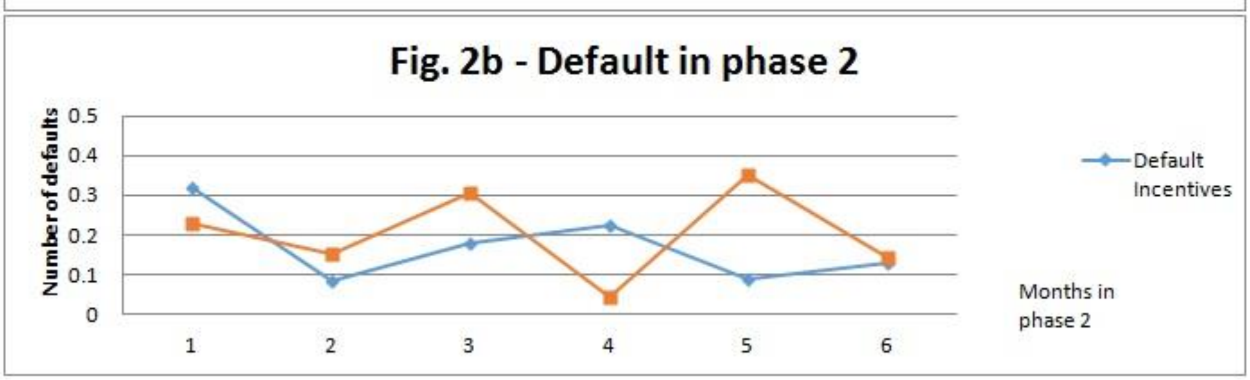


Fig. 2b - Default in phase 2

The results from Equation [1] are presented in Table 10, Panel A, columns 1 and 2. On average, the incentives increased the number of new detections by 1.58 (24.1 percent). This effect is statistically significant at the 10 percent level. It is robust to including city fixed effects and the health worker control variables: we then find an increase of the number of new detections by 2.18 (33.2 percent), significant at the 5 percent level. The detection incentives had therefore a large direct effect on the health workers' performance, which however dampens over time. Importantly, while the most likely interpretation of this result is that the incentives increased the health workers' effort to detect new patients, alternative interpretations are possible: increased forgery of the detection data, or increased effort to recruit in their centers patients who would otherwise have taken their pills at other treatment centers. Our estimates of the impact of the incentives on the health workers' reported effort in Section 6.2 will be helpful as a first attempt to disentangle these rival interpretations.

	(1)	(2)	(3)	(4)	(5)	(6)
		Detections			Defaults	
Panel A. Impact of DETECTION incentives (1st stage)						
Treatment	1.58	1.41	2.18	0.08	0.06	0.07
	(0.90)*	(0.79)*	(0.95)**	(0.04)**	(0.03)**	(0.04)*
City fixed effects		Yes	Yes		Yes	Yes
Health worker controls			Yes			Yes
Observations	507	507	426	476	476	439
R-squared	0.03	0.10	0.19	0.01	0.11	0.12
Mean in Control Group	6.56	6.56	6.56	0.08	0.08	0.08
Panel B. Impact of DEFAULT incentives (2nd stage)						
Treatment	-0.67	-0.51	0.61	-0.04	-0.02	-0.01
	(0.97)	(0.97)	(0.81)	(0.06)	(0.07)	(0.06)
City fixed effects		Yes	Yes		Yes	Yes
Health worker controls			Yes			Yes
Observations	181	181	157	340	340	311
R-squared	0.01	0.11	0.37	0.00	0.02	0.09
Mean in Control Group	7.56	7.56	7.56	0.21	0.21	0.21

Notes: Clustered standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. We take a health worker × month as the unit of observation and include all health workers. Panel A reports the impact of detection incentives (1st stage) on the number of newly detected patients in the month and the number of defaults in the month. These two outcomes are administrative data reported on the centerwise summary sheets, month after month. Panel B reports the impact of default incentives (2nd stage) on the same two outcomes.

The health worker controls include the number of years of previous work experience and dummies indicating the health worker's education level, whether he has any prior work experience at all, and if so, whether it was in the social sector, and whether he lives in a neighborhood where one of the centers he operates is located.

Table 10 – Impact of the incentives on the number of detections and defaults

6.1.2. Impact of detection-based incentives on defaults (phase 1)

In the first phase, the salary of incentivized health workers was independent on the number of defaults in their centers. To the extent that the incentives encouraged the health workers to increase their effort and time spent detecting new patients, as suggested by the results above, they may have conversely reduced their effort and time spent preventing defaults, echoing predictions from principal-agent models in a work environment characterized by multitasking.

To test this hypothesis, we plot the average number of defaults per CHW month by month, starting with the first month after the randomization. As Figure 1b makes clear, the number of defaults is very small overall, but higher in the treatment group. Columns 3 and 4 of Table 10, Panel A, report the results of Equation [1], using the number of defaults as the outcome. We find that the number of defaults is significantly larger among treated CHWs: the detection incentives led to an increase in defaults by 0.08 per month (100 percent), an effect statistically significant at the 5 percent level and robust to including city fixed effects and the health worker control variables.

This result brings empirical support for the multitasking hypothesis. An alternative interpretation is however possible: the additional patients recruited by the health workers who received incentives may have been more prone to default than the average patient. Together with our estimates of the impact of incentives on the health workers' reported effort in Section 6.2, the (forthcoming) in-depth evaluation of the impact of the detection-based incentives on patients' characteristics will enable us to disentangle the two interpretations.

6.1.3. Impact of default-based incentives (phase 2)

Phase 2 incentives were computed based on the number of defaults that occurred in the health workers' centers within the month. Again, we compute and plot the average number of newly detected patients and the average number of defaults per CHW month by month in the control and treatment groups, starting with the first month after the phase 2 randomization (Figures 2a and 2b). Unlike in phase 1, there does not seem to be any systematic difference between the control and treatment groups for any of these two outcomes.

To investigate this lack of difference more systematically, we estimate Equation [1], using the number of detections or defaults as the outcome, and redefining T_i as a dummy equal to 1 if i receives phase 2 default incentives and 0 otherwise. The results are shown in Table 10, Panel B, columns 1 through 4. We find that phase 2 incentives affected neither the number of detections nor the number of defaults : the estimates are small and not statistically significant at the standard levels. This result is robust to the inclusion of city fixed effects and health worker control variables.

6.2. Impact on health worker effort and satisfaction

We now turn to outcomes measured through CHWs' self-reports to address three questions raised by the aforementioned results.

First, how did the health workers incentivized in phase 1 manage to increase the number of detections, and for what reason did the number of defaults increase?

Second, the lack of effect of phase 2 incentives on the number of defaults is somewhat puzzling, given the large effects of phase 1 incentives. Two types of explanations are in theory possible: either the health workers did not change their default-prevention effort in response to phase 2 incentives; or their increased efforts to prevent defaults did not pay off. Which explanation(s) does our data bring empirical support for, if any?

Third, the salaries of the health workers who received incentives (whether in phase 1 or in phase 2) differed from the fixed salaries received by health workers in the control group: they had a higher variance and, in phase 1, were higher on average. How much did this affect the health workers' job satisfaction?

To address these questions, we estimate specifications of the form in Equation [2]:

$$S_i = \alpha + \beta T_i + X_i' \lambda + \sum_c \delta_c^i + \varepsilon_i$$

To measure the impact of phase 1 detection incentives (resp. phase 2 default incentives), we use self-reported outcomes S_i measured during the midline health worker survey (resp. endline health worker survey) and define T_i as a dummy equal to 1 if health worker *i* receives phase 1 detection incentives (resp. if he or she receives phase 2 default incentives). As in Equation [1], δ_c^i are city fixed effects, X_i is a vector of health worker characteristics, and β estimates the effect of the incentives.

The results are shown in Table 11. Two important caveats are in order when we interpret the results. First, all outcomes considered here are self-reported and thus potentially more prone to differential measurement bias than the administrative outcomes considered in the previous section. Second, in each regression, we can only include one observation per health worker, resulting in a lower statistical power than in the regressions using one observation per health worker per month, and increasing the likelihood of false negatives.

6.2.1. Impact of phase 1 incentives on health workers' activities

As shown in Table 11, Panel A, phase 1 incentives increased the number of methods reported by the health workers to detect new patients (columns 1 and 2), the number of types of actions undertaken with someone suspected of having TB, (columns 3 and 4) the number of types of follow-up actions undertaken after the first meeting (columns 5 and 6), and the number of sputum samples collected last week (columns 9 and 10). These estimates, however, are relatively small and not statistically significant. The impact on the number of days spent detecting new patients over the last week is negative, but not statistically significant either (columns 11 and 12). Thus, at this stage, our results do not allow us to precisely picture the dimension(s) of their daily activities on which the treated health workers increased their effort, explaining the increased number of detections.

Interestingly, however, health workers are significantly less likely to remember the name of a person suspected for tuberculosis for whom sputum was sent for testing (Table 11, Panel A, columns 7 and 8).

This effect could help explain why the detection incentives increased the number of defaults: as a result of focusing on the number of new patients, the health workers may have put less effort in building individual relationships with them, resulting in increased defaults. This evidence is only indirect, however. In our ongoing work on this, we plan to use the available data more systematically to better understand the mechanisms underlying the effects of phase 1 incentives.

6.2.2. Impact of phase 2 incentives on health workers' activities

As shown in Table 11, Panel B, phase 2 incentives did not significantly affect the activities reported by the canvassers: the estimates of the effects on the number of methods used to prevent patients from defaulting, on the number of methods used to retrieve defaulters, on the ability to recall the name of a patient counseled for missing pills, or on the number of hours worked for the job are all small, sometime positive, sometime negative, but always not statistically significant.

If anything, the available evidence suggests that the health workers did not adjust their behavior in response to the default incentives (rather than supporting the hypothesis that they did adjust their behavior but with no effect). One potential explanation for the contrasted effect of phase 1 and phase 2 incentives is that the number of defaults may have been perceived by the health workers as a dimension on which there was less margin of improvement than the number of detections: indeed, the number of defaults was very low to begin with.

6.2.3. Impact of incentives on job satisfaction

Both phase 1 incentives and phase 2 incentives seem to have affected the health workers' job satisfaction in a negative way: health workers who received incentives were several percentage points less likely to report that they were satisfied with their job or to have recommended the job to someone else. This is consistent with the fact that incentives increase the variance of the salaries, with no effect (for phase 2 incentives) or a small effect only (for phase 1 incentives) on the average amount received. However, the estimates are not statistically significant. On this question, our analysis will become more conclusive and more precise when we also take into account the health workers of our second partner NGO.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
	# of methods used to detect new patients		# of types of actions undertaken with someone suspected of having TB		# of types of follow-up actions undertaken after the first meeting		Remembers the name of a person suspected for TB for whom sputum was sent for testing		# of sputum samples collected last week		# of days spent detecting new patients last week		Satisfaction with the compensation scheme		# of hours worked for the job per day		Job satisfaction		Has recommended the job to someone else	
Panel A. Impact of DETECTION incentives (1st stage)																				
Treatment	0.15	0.07	0.19	0.30	0.02	0.15	-0.14	-0.25	0.55	-0.60	-0.39	-0.22	0.02	-0.06	-0.18	-0.14	-0.10	0.02	-0.12	-0.05
	(0.29)	(0.41)	(0.18)	(0.23)	(0.27)	(0.29)	(0.08)*	(0.10)**	(0.93)	(1.10)	(0.28)	(0.33)	(0.26)	(0.36)	(0.22)	(0.27)	(0.11)	(0.15)	(0.10)	(0.13)
City fixed effects		Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes
Health worker controls		Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes
Observations	64	55	64	55	64	55	63	55	64	55	63	55	61	53	62	54	62	54	62	54
R-squared	0.00	0.44	0.02	0.47	0.00	0.64	0.05	0.60	0.01	0.60	0.03	0.65	0.00	0.53	0.01	0.63	0.01	0.51	0.02	0.58
Mean in Control Group	3.10	3.10	1.55	1.55	2.24	2.24	0.96	0.96	1.62	1.62	5.82	5.82	3.33	3.33	7.70	7.70	0.81	0.81	0.89	0.89

Panel B. Impact of DEFAULT incentives (2nd stage)														
	# of methods used to prevent default		# of methods used to retrieve defaulters		Remembers the name of a person counseled for missing pills		Satisfaction with compensation scheme		# of hours worked for the job per day		Job satisfaction		Has recommended the job to someone else	
Treatment	-0.32	-0.25	-0.25	-0.41	0.01	0.01	0.34	0.17	0.11	0.39	-0.10	-0.16	-0.06	-0.17
	(0.23)	(0.25)	(0.28)	(0.31)	(0.10)	(0.16)	(0.26)	(0.31)	(0.23)	(0.26)	(0.11)	(0.14)	(0.12)	(0.10)
City fixed effects		Yes		Yes		Yes		Yes		Yes		Yes		Yes
Health worker controls		Yes		Yes		Yes		Yes		Yes		Yes		Yes
Observations	71	64	70	63	45	40	71	64	70	63	71	64	70	63
R-squared	0.03	0.53	0.01	0.49	0.00	0.50	0.02	0.40	0.00	0.50	0.01	0.40	0.00	0.69
Mean in Control Group	3.66	3.66	2.61	2.61	0.86	0.86	2.84	2.84	7.31	7.31	0.72	0.72	0.68	0.68

Notes: Clustered standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. We take a health worker as the unit of observation and include all health workers. Panel A reports the impact of detection incentives (1st stage) on a number of variables self-reported by the health workers during the midline survey, at the end of the first stage of the study. Panel B reports the impact of default incentives (2nd stage) on identical or similar variables, self-reported by the health workers during the endline survey, at the end of the second stage of the study. The health worker controls include the number of years of previous work experience and dummies indicating the health worker's education level, whether he has any prior work experience at all, and if so, whether it was in the social sector, and whether he lives in a neighborhood where one of the centers he operates is located.

Table 11 - Impact of the incentives on the health workers' description of their work and on their satisfaction

7. Concluding comments

The main results at this point can be summarized as follows: i) introducing performance-based incentives related to detections generates a strong but seemingly short-lived boost to the number of new patients registered in an Operation ASHA center, which can be due to new detections, transfer-ins from other centers, or possibly forgery; ii) the number of defaults increases simultaneously, which likely reflects a reallocation of health workers' effort towards the incentivized activity and to the detriment of the non-incentivized one (referred to as 'multitasking' in the economics literature), but may also be due to incentivized health workers detecting patients more prone to default; iii) default-based incentives in phase 2 do not cause any detectable impact.

Our estimates of the impact of the incentives on health workers' self-reported effort provide some support for the interpretation that the detection-based incentives led to a reallocation of the health workers' effort towards this activity and away from the prevention of default, but the identification of causal mechanisms, time effects and interaction effects between the two incentive schemes is hampered by the small sample size (only 76 health workers).

A more comprehensive analysis will be conducted when the remaining data is collected and cleaned, and when the different data sources are matched. It will include a study of the impact of performance-based incentives for health workers on their interactions with patients, patients' satisfaction, health behavior and health status. The matched patient-provider data will allow to test several hypotheses: (i) financial incentives may lead to selection against the most remote, hard-to-detect TB-infected individuals, or "cherry-picking", as suggested in Shen (2003) and Oxman and Fretheim (2008); (ii) social distance between patient and provider, as measured by sharing the same language or religion, may strengthen the impact of incentives, as shown in Singh and Mitra (2013). In addition, our data will allow for investigating potential forgery in the data reported by CHWs and their managers by cross-checking detection and default figures with information from random visits and administrative data.

References

Arora, V. K., Sarin, R., K., L., 2003. Feasibility and effectiveness of a public-private mix project for improved TBcontrol in Delhi, India. International Journal of Tuberculosis and Lung Disease 7, 1131-1138.

Banerjee, A., Deaton, A., Duflo, E., May 2004. Wealth, Health, and Health Services in Rural Rajasthan. American Economic Review 94 (2), 326-330.

Basinga, P., Gertler, P. J., Binagwaho, A., Soucat, A. L., Sturdy, J., Vermeersch, C. M., 2011. Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation. The Lancet 377 (9775), 1421-1428.

Benabou, R., Tirole, J., 07 2003. Intrinsic and Extrinsic Motivation. Review of Economic Studies 70 (3), 489-520.

Chaudhury, N., Hammer, J., Kremer, M., Muralidharan, K., Rogers, F. H., Winter 2006. Missing in Action: Teacher and Health Worker Absence in Developing Countries. Journal of Economic Perspectives 20 (1), 91-116.

Chaudhury, N., Hammer, J. S., 2004. Ghost Doctors: Absenteeism in Rural Bangladeshi Health Facilities. World Bank Economic Review 18 (3), 423-441.

Das, J., Hammer, J., December 2005. Which doctor? Combining vignettes and item response to measure clinical competence. Journal of Development Economics 78 (2), 348-383.

Das, J., Hammer, J., May 2007. Money for nothing: The dire straits of medical practice in Delhi, India. Journal of Development Economics 83 (1), 1-36.

Deci, E. L., 1975. Intrinsic motivation. New York: Plenum

Eichler, R., Levine, R. (Eds.), 2009. Performance Incentives for Global Health: Potential and Pitfalls. Washington DC: Center for Global Development.

Frey, B. S., Goette, L., 1999. Does Pay Motivate Volunteers? IEW – Working Papers 007, Institute for Empirical Research in Economics - University of Zurich.

Friedman, J., May 2013. Do financial incentives undermine the motivation of public sector workers? maybe, but where is the evidence from the field?, world Bank Blogs. Development Impact.

Gertler, P., Vermeersch, C., May 2013. Using Performance Incentives to Improve Medical Care Productivity and Health Outcomes. NBER Working Papers 19046, National Bureau of Economic Research, Inc.

Gneezy, U., Rustichini, A., August 2000. Pay Enough Or Don'T Pay At All. The Quarterly Journal of Economics 115 (3), 791-810.

Heyman, J., Ariely, D., 2004. Effort for payment: A tale of two markets. Psychological Science 15 (11), 787-793.

Hillman, A. L., Ripley, K., Goldfarb, N., Nuamah, I., Weiner, J., Lusk, E., 1998. Physician financial incentives and feedback: failure to increase cancer screening in medicaid managed care. American Journal of Public Health 88 (11), 1699-1701.

Holmstrom, B., Milgrom, P., Special I 1991. Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. Journal of Law, Economics and Organization 7 (0), 24-52.

Kamolratanakul, P., Hiransuthikul, N., Singhadong, N., Kasetjaroen, Y., Akksilp, S., Lertmaharit, S., 2002. Cost analysis of different types of tuberculosis patient at tuberculosis centers in Thailand. Southeast Asian Journal of Tropical Medicine and Public Health 33 (2), 321-30.

Laffont, J.-J., Martimort, D., 2002. The Theory of Incentives: The Principal- Agent Model. Princeton University Press.

Lazear, E. P., December 2000. Performance Pay and Productivity. American Economic Review 90 (5), 1346-1361.

Leonard, K. L., Masatu, M. C., December 2010. Professionalism and the knowdo gap: exploring intrinsic motivation among health workers in Tanzania. Health Economics 19 (12), 1461-1477.

Lewis, M., Pettersson, G., Oct. 2009. Governance in health care delivery : raising performance. Policy Research Working Paper Series 5074, The World Bank.

Loevinsohn, B., Harding, A., 2005. Buying results? Contracting for health service delivery in developing countries. The Lancet 366 (9486), 676-681.

Maung, M., Kluge, H., Aye, T., Maung, W., Noe, P., Zaw, M., Jost, S. P., Uplekar, M., Lönnroth, K., 2006. Private gps contribute to TB control in Myanmar: evaluation of a PPM initiative in Mandalay division. The International Journal of Tuberculosis and Lung Disease 10 (9), 982-7.

Miller, G., Babiarz, K. S., Apr. 2013. Pay-for-Performance Incentives in Lowand Middle-Income Country Health Programs. NBER Working Papers 18932, National Bureau of Economic Research, Inc.

Mullen, K. J., Frank, R. G., Rosenthal, M. B., 2010. Can you get what you pay for? Pay-for-performance and the quality of healthcare providers. RAND Journal of Economics 41 (1), 64-91.

Oxman, A.D., & Fretheim, A. (2008) An overview of research on the effects of results-based financing. Report from Norwegian Knowledge Centre for the Health Services, Systematic Reviews, nr 16 –2008.

Shearer, B., 2004. Piece rates, fixed wages and incentives: Evidence from a field experiment. Review of Economic Studies.

Shen, Y. (2003). Selection Incentives in a Performance-Based Contracting System. Health Services Research, 38(2), 535–552.

Singh, P., Jul. 2011. Performance Pay and Information: Reducing Child Malnutrition in Urban Slums. MPRA Paper 29403, University Library of Munich, Germany.

Singh, P. and S. Mitra (2013) Absolute Versus Relative Performance Pay: Evidence from an Experiment Targeting Child Malnutrition in West Bengal. Working paper available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2332896 .

Tirole, J., Bénabou, R., December 2006. Incentives and Prosocial Behavior. American Economic Review 96 (5), 1652-1678.

Volmink, J., Gamer, P., 1997. Directly observed therapy. The Lancet 349 (9062), 1399-1400.

Walley, J. D., Khan, A., Newell, J. N., Khan, H., 2001. Effectiveness of the direct observation component of DOTS for tuberculosis: a randomised controlled trial in Pakistan. The Lancet 357 (9257), 664-669.

Zwarenstein, M., Schoeman, J. H., Vundule, C., Lombard, C. J., Tatley, M., 1998. Randomised controlled trial of self-supervised and directly observed treatment for tuberculosis. The Lancet 352 (9137), 1340-1343.