Misplaced Effort: The Effect of Pay-for-Performance in the Health Sector in the DRC *

Elise Huillery† and Juliette Seban‡

December 22, 2015

Abstract

The use of financial incentives for service providers is increasing in developing countries. Using a field experiment in the DRC, we show that introducing a pay-for-performance mechanism in the health sector reduced the revenue of the facilities and more importantly service utilization and child health in the catchment areas. The explanations sometimes proposed in the literature for the detrimental effect of incentives, such as motivational crowding out or switching away from non incentivized actions, do not seem to play a role. In fact, the workers provided more effort, but this effort was evidently misplaced, suggesting that incentives can have detrimental effects in environments where performing is difficult relative to worker capacity.

JEL Codes: H51, I18, 012

1 Introduction

Whether governments should incentivize service providers to improve service delivery and utilization is a crucial question in both developing and developed countries. Incentives are beneficial under two conditions: (i) incentives should encourage more effort and (ii) greater effort should translate

*We thank Yann Algan, Elizabeth Beasley, Nicolas de Borman, Julia Cagé, Vera Chiodi, Damien De Walque, Pascaline Dupas, Gyuri Fritsche, Roberto Galbiati, Emeric Henry, Paul Jacob Robyn, Hadia Samaha, Mark Stabile and Christel Vermeesch for helpful comments and discussions, as well as Emmanuelle Auriol, Roland Benabou, Denis Cogneau, Marc Gurgand, Guy Laroque, Alice Mésnard, Pieter Serneels, Philippe de Vreyer and numerous seminar participants at the World Bank, EUDN 2013 scientific conference, Paris School of Economics, Toulouse School of Economics, Sciences Po, Navarra Center For International Development, Journées Louis-André Gérard-Varet 2015, 2015 iHEA World Congress, Paris I University, and the Blavatnik School of Government, Oxford University. We wish to acknowledge the support of the World Bank and J-PAL Europe. Jean-Jacques Frère and Hadia Samaha led the implementation of this project and we are grateful to them for their leadership. We thank the Health Authorities of Haut-Katanga and the Health Sector Rehabilitation and Support project (Projet d'Appui à la Réhabilitation du Secteur de la Santé - PARSS) in Lubumbahi and Kinshasa for their collaboration. We are deeply indebted to Julie Berthet-Valdois for outstanding project management, and to the Ecole de Santé Publique de Lubumbashi for excellent data collection. We also thank Saidou Ouedraogo for excellent assistance with data entry. Finally, we gratefully acknowledge the staff and patients of the health facilities as well as the households for the time and information they gave. All errors remain our own, and the opinions expressed in this paper are ours alone and should not be attributed to the institutions with which we are associated, the World Bank or the Government of the Democratic Republic of Congo.

†Sciences Po, Department of Economics (elise.huillery@sciencespo.fr) and J-PAL

‡International Rescue Committee (seban.juliette@gmail.com)

into higher performances. This paper shows that a pay-for-performance mechanism introduced in the health sector in the Democratic Republic of Congo (DRC) decreased performances despite greater effort from the health workers. The incentive scheme reduced service utilization, health outcomes, and providers' revenue, showing that motivated health workers may not always be good entrepreneurs. These findings provides first evidence in the field of previous observations in the lab that people who are offered a reward for performing at some tasks may perform worse at difficult tasks (Glucksberg, 1962) and that larger stakes can cause "big mistakes" (Ariely et al. 2009).

The pay-for-performance scheme implemented in the DRC is a particular form of performance-based financing (henceforth PBF) by which the government allocates its budget to the health facilities on the basis of the number of patients who visit the facility for specific services relative to the other facilities. Different models have been implemented in many developed and developing countries using various performance criteria. Here, the model is a team and relative incentive scheme rewarding an output (the number of patients) rather than an input (e.g. daily attendance or technical quality[1]), hence pushing health workers to develop appropriate strategies to increase service uptake. Many different obstacles can hinder the demand for health services: prices, information, service quality, or behavioral issues. Since local health workers should be in a better position than the central government to identify the relevant obstacles in a specific area, PBF is a contract that decentralizes the task of finding the appropriate strategies to increase health service uptake. There are other PBF models (see Miller and Barbiaz (2013) for a review) so the results will thus be discussed in light of this particular PBF approach. This paper compares the efficiency of this incentive scheme to a model where the governement allocates its budget on a fixed basis, irrespectively of health facilities' activity.

Our empirical strategy relies on a field experiment conducted in the Haut-Katanga district of the DRC between 2009 and 2013. The 96 health areas of the Haut-Katanga district were randomly assigned to performance-based or fixed governmental payments, while ensuring that the same amount of resources was allocated to each group to neutralize any resource effect. All of the 152 public, private or religious health facilities in these health areas except the four biggest hospitals participated in the experiment. Unannounced visits to the facilities were performed to measure worker attendance, and an independent survey was administered a few months after the payments had been

[1] Two studies provide evidence that rewards contingent on a specific input (respectively attendance and service quality) do motivate health workers to provide more of this input (at least in the short run), but did not lead to any increase in health service utilization (the output) (Banerjee and Duflo, 2008; Peabody et al., 2011).

withdrawn to collect data on the supply and price of health services, health worker motivation, service utilization, and population health, during and after the PBF implementation. The analysis distinguishes targeted and non-targeted services in order to test the potential disruptive effect of incentives on non-targeted services.

We find that the incentive scheme led to an overall decrease in utilization of health services by the population, in particular for curative and prenatal care services. The incentivized facilities suffered from a 42% decrease in their total revenue (even though the two groups received the same budget from the government), and a 34% reduction in worker revenue. The loss in revenue translated in lower quantity and quality of equipment and infrastructure. Even more critical, we find a deterioration in child health outcomes.

The reduced performances do not result from a reduction in worker effort and motivation. The introduction of PBF spurred health workers into greater effort to attract patients: (1) they were more present in facilities; (2) they organized more preventive health sessions at facilities; (3) they conducted more community-based outreach activities to inform the population about the services offered at the facility. Overall, the financial incentives thus induced an intensification of effort to increase utilization of targeted health services. Equally important, we find that the increased effort invested in the targeted services did not happen at the expense of the effort invested in the non-targeted services. Also, the reward did not induce significant score manipulation. Finally, we did not find evidence that the collective nature of the incentive induced significant free-riding. Overall, none of the perverse behavioral effects that could be dreaded were realized. This result contrasts with the finding of Ashraf et al. (2014) in Zambia in which financial incentives did not induce more effort by hairdressers to sell condoms, as well as Rasul and Rogger (2014) in which incentives to infrastructure project managers induced perverse effects on service supply.

However, workers' strategies to attract more patients were evidently counterproductive. In addition to the more intense direct selling through preventive sessions and outreach activities, workers also significantly reduced fees for targeted services and did not change service technical quality. The higher request to visit from the health workers combined with the lower prices may have been perceived as aggressive marketing and signaled low quality of the supply. In fact, the decrease in demand is observed along with an increase in the proportion of non-users who declare that the service is "of little interest" or "too far away", suggesting lower perceived service utility. This result indicates that the population needs more than eased access and logistical information,

may be a better understanding of health service benefit, which was not anticipated by the health workers. The lower quantity and quality of equipment and infrastructure resulting from loss in revenue could also be an additional explanation for the decreased demand for health services and lower perceived service utility.

The broad empirical literature on incentives in for-profit organizations shows that rewards reinforce agents' willingness to achieve the rewarded action and increase the output (Lazear, 2000; Bandiera et al. 2007; Bandiera et al. 2013). In service delivery, several empirical studies advocate that performance-based financing improves accountability, efficiency, quality and quantity of service delivery (see Loevinsohn and Harding (2005) and Eichler and Levine (2009) for an overview). However, the presence of confounding factors[2] and the fact that it is often not possible to isolate the effects of financial incentives from other elements[3] make the question of the impact of PBF largely unanswered (Christianson et al. 2008; Eldridge and Palmer 2009; Oxman and Fretheim 2009). Olken et al. (2014) report on a pay-for-performance mechanism applied to village committees in Indonesia, testing whether incentivized community members can monitor efficiently teachers and health workers. They find that incentives to village committees led to an increase in health workers' attendance, better health outcomes, and an absence of negative spillovers on untargeted outcomes. Basinga et al. (2011) conducted a quasi-experimental study on the effect of PBF in Rwanda that is the closest to our study. The study uses a difference-in-difference strategy in order to control for potential selection effects[4]. It finds that PBF is an efficient way to increase utilization of some of the targeted services as well as worker productivity, and to improve some targeted health outcomes (Basinga et al. 2011; De Walque et al. 2013; Gertler and Vermeesch 2013). The literature on the effect of PBF using clean identification is thus very limited, and the lack of information on precise worker responses and strategies still needs to be addressed[5] (Miller and Babiarz, 2013).

[2]Until 2011, the studies of the impact of PBF did not use credible comparison groups: they compare very small groups (generally 2-3 districts) which were not randomly assigned to the different treatments (Soeters, 2011; Rusa et al., 2009; Soeters et al., 2005; Eicher et al., 2007; Soeters and Griffths, 2003; Forsberg, 2001), or the situation before and after the introduction of PBF (Sondorp et al., 2008; Eicher et al., 2007; Meessen et al., 2007).

[3]PBF has commonly been a part of a package that may include increased funding, technical support, training, changes in management, and new information systems. In most studies, the level of resources allocated to the health facilities in different treatments is not similar, as well as the level of technical supervision and information system.

[4]166 facilities were grouped into 8 pairs and one side of each pair was randomly assigned to pay-for-performance funding, while the other side continued with the traditional input-based funding until 23 months after study baseline. The paper uses a difference-in-difference strategy in order to control for potential selection effects since the number of units of randomization was very small and some post-randomization reassignment of some districts happened because of administrative boundaries' reorganization.

[5]The literature is not very developed in the context of high income countries either, and identification issues also limit the scope of many studies. One recent study on the effect of pay-for-performance mechanism is Mullen, Franck and Rosenthal (2010), which uses a difference-in-difference strategy on US data and show that pay-for-performance targeted on service quality did not lead to any major improvement in quality of targeted services, nor notable effect on the quality of non-targeted services. Note that pay-for-performance in high income countries tends to reward

This paper makes several contributions to the literature on improving health service delivery. First, this paper constitutes one of the few studies using the random assignment of a large number of health areas to estimate the effects of a performance-based mechanism as a way to allocate governmental resources to health facilities, following Olken et al. (2014) and Basinga et al. (2011). Second, this paper provides first empirical evidence that a pay-for-performance scheme may lead to counterproductive results for both the workers and the population. Third, this paper explores in detail worker responses, strategies and motivation to test the potential adverse effects of financial incentives found in the theoretical and behavioral literatures: (i) that incentives may be negative motivational reinforcers (Lepper et al. (1973), Deci (1975), Deci and Ryan (1985), Benabou and Tirole 2003, Benabou and Tirole 2006, Gneezy et al. 2011); (ii) that agents may concentrate their effort on the actions attached to the reward at the expense of non incentivized actions (Holmström and Milgrom 1991); (iii) that PBF may induce a reduction in effort due to free-riding problems since rewards are collective and not individual (Bandiera et al. (2013)); (iv) and that incentivized agents may manipulate performance measures to obtain more of the reward. We show that in the context of the health sector in the DRC, none of these adverse behavioral effects happen. While the existing evidence in Indonesia and Rwanda demonstrated positive effects of PBF, this paper shows that financial incentives can also generate misplaced effort.

There are key policy implications of our findings for governments considering performance-based mechanisms as a way to allocate public resources to the health sector. First, financial incentives increase health worker motivation without detrimental effect on non-incentivized actions, score manipulation, or free-riding. Second, the increased motivation can be accompanied by reduced performance when the task requires complex strategies, for instance when demand is delicate and users do not respond as expected. The translation of motivation into performance may be better in contexts where demand is classic and the rewarded task is easy relative to worker capacities.

The remainder of the paper is organized as follows. Section 2 presents the context in which the experiment was set up and the experimental design. Section 3 examines the data and econometric approach. Section 4 presents the effects of PBF compared to a fixed payment approach, and Section 5 concludes.

quality measures instead of service volume. This might be because the policy concern is more about service quality than about service utilization in rich countries relative to poor countries. See Stabile and Thomson (2014) for a review.

2 Experimental Set-Up

2.1 Background on Health in DRC and Haut-Katanga

The Democratic Republic of Congo (DRC) is the second largest country in Africa by area, with the fourth largest population at 66 million (World Bank, 2012). It is also among the poorest countries in the world: the country is ranked second from the bottom of the Human Development Index (186 out of 187 in 2012) (UNDP, 2012), with an estimated per capita income of US$ 220 (current) in 2012 (World Bank, 2012). Impoverished by decades of war, instability and bad governance, DRC is not on track to reach the health-related Millenium Development Goals. Since the democratic elections in 2006, the country has started a slow reconstruction phase and a decentralization process, with the election of provincial governments, including provincial ministers of health. Performance-based Financing (PBF) is a strategy for improving health outcomes among the population which has been developed and implemented to promote effective service delivery.

The district of Haut-Katanga entails 1.26 million people in the province of Katanga in the south-eastern corner of the DRC. From September to November 2009, a survey was conducted to better understand the health situation in Haut-Katanga by providing a description of the functioning of the health facilities as well as the characteristics and behavior of the health workers, patients and households in the district. The survey sample entailed 152 health facilities (5% referral centers, 71% health centers and 24% health posts)[6]. In regards to health services coverage, 87% of patients lived 10km or less from facilities, 70% spent less than one hour to travel to the facility, and there was one health worker for every 1860 individuals[7]. Coverage for basic health services was thus not so worrying.

However, the poor quality of infrastructure was striking: only one out of four facilities had access to a water tap or electricity. The majority of facilities had only low-cost basic equipment. Most health workers were not public agents: one worker out of four did not receive any fixed wage from the government. Worker payment thus came from facility revenue, mainly user fees and drug sales, but also public grants and -sometimes- funds from NGOs and private donors. Health workers spent on average 52 hours per week working in the health facility. They received 35 patients the week before the survey, equating approximately 7 patients per working day per health worker, which means that health workers were far from overworked. Patients reported quite short consultation

[6] 161 health facilities were recognized as part of the government health system in the district, among which 5 hospitals were excluded from the study and 4 health centers could not be reached.

[7] The Ministry of Health considers that there should be at least one health worker for every 1500 individuals.

time (16 minutes on average), and twice as much waiting time before the consultation (30 minutes on average)[8]. 56% of patients had to pay a fee for the service, although the median fee for a visit was quite low 800FC (0.88$).

In 2009, the health status of the population was found to be poor: 25% of the sample had been sick in the last four weeks, with malaria and diarrhea being the most prevalent diseases. Concerning maternal health, 31% of births in the last 12 months were not attended in a formal health facility. Mothers used more prenatal than postnatal health services: 76% of women pregnant in the last 12 months had at least one prenatal visit while only 10% attended a postnatal visit. However, according to women's recall, only a third of prenatal visits included the minimum tests. Despite frequent immunization campaigns, only 13% of children under 5 years-old were able to present an immunization card (although based on mothers' declaration a majority of children got immunized at least once). Finally, we found low exposure to prevention campaigns other than immunization, with around two thirds of the households never exposed to any HIV prevention, child nutrition, or maternal health campaign.

2.2 Experimental Design

Payment Calculation

In the Haut-Katanga district, the 96 health areas (totalizing 152 health facilities) were randomly assigned to one of two payment systems. *In the fixed payment group*, the amount allocated to each facility was calculated based on the staff in the facility: a list of eligible workers was established at the beginning of the pilot by the Ministry of Health. Each worker was entitled to a given amount of governmental payment depending on his/her grade and experience. *In the performance-based payment group*, payments were made based on declaration of service volumes by facilities. The fact that payment was not attached to specific workers in the PBF system led to a significantly more egalitarian distribution of payments among workers: in the fixed payment group, 77% of health workers received a share of the payment, whereas 93% of workers in the PBF group[9].

The targeted services included seven services at the primary care level (outpatient first curative consultations, prenatal consultations, deliveries, obstetric referral, children completely vaccinated,

[8] This survey did not allow for assessing the technical quality of medical procedures.

[9] This information was collected at endline from the facility heads. The facility heads listed the workers in the facility, indicated whether each of them received a share of the last payment, and the corresponding amount. PBF benefited especially to non-technical workers (pharmacists, managers, secretaries, receptionists and maintenance workers) who are not in the governmental payroll and are therefore less likely to receive a share of the fixed payment.

tetanus toxoid vaccination, and family planning consultations) and three additional services at the secondary care level (C-section, blood transfusion, and obstetric referrals to hospitals). Relative prices for each service are presented in Appendix Table 1.

Formally, payments to health facilities can be written as:

$$P_{i,m} = \alpha_i + \beta_m Q_{i,m}$$

where $P_{i,m}$ is the payment to facility i in month m, α_i represents a fixed component, $Q_{i,m}$ is the vector of targeted service quantities provided by facility i in month m, and β_m is the vector of prices that the government attach to each targeted service in month m. The PBF group was characterized by a pure performance-based mechanism ($\alpha_i = 0$ and $\beta_m > 0$), whereas the comparison payment group had a pure fixed payment ($\alpha_i > 0$ and $\beta_m = 0$). In order to ensure neutrality in the level of funds received by both groups and to isolate the incentive effect from the resource effect, the total budget allocated to health facilities in the PBF group was the same as the total budget allocated to health facilities in the fixed payment groupHence, noting $\overline{Q_m}$ the average service provision in the PBF group in month m and $\overline{\alpha}$ the average payment in the fixed payment group:

$$\overline{\alpha} = \beta_m \overline{Q_m}$$

In practice, $\overline{\alpha}$ was fixed and β_m was adjusted accordingly at $\frac{\overline{\alpha}}{\overline{Q_m}}$[10]. Although relative prices attached to the targeted services were constant , absolute prices and facility payments were thus determined by the quantity of services provided by the facility *relative to* the quantity of services provided by the other incentivized health facilities[11].

The budget used in this experiment estimated at $0.43 per capita per year (average monthly facility payments were $550 and the average catchment area population was 12,900)[12]. The average monthly payment by facility from June 2010 and September 2012 did not differ in the fixed payment and in the PBF group. This confirms that the experimental design was respected and that the study isolates the incentive effect from any resource effect.

[10] The other way to equalize the two total budgets is to fix $\beta_m = \beta$ and adjust $\overline{\alpha}$ accordingly at $\beta\overline{Q_m}$. This technique was used in the Rwanda experiment where the governmental budget could increase according to the average service provision in the incentivized group.

[11] As discussed in Bandiera et al. (2005), relative incentives might yield lower effort from the health workers than piece rates because effort imposes a negative externality on others, in particular when others are friends. In the context of this PBF program, we do not have measures of interpersonnal connections between workers of different health facilities. However, health facilities are generally distant one from another and it seems unlikely that health workers from different health facilities live in the same neighborhood and are close friends.

[12] This is lower than in other contexts where output budgets range between $2 and $3 per capita per year.

Score Manipulation

Service volumes were measured using monthly reports submitted by facilities, in which the number of patients for each targeted service was reported. These numbers were verified by public agents at the beginning of the following month by comparing reported volumes with those found in health facility registers[13]. Payments were calculated and paid as soon as the register verification was done, generally during the following month. The same payment lag applied to the fixed payment group since all payments happened at the same time. Subsequent verification of the information noted in the registers was also conducted: a random sample of 30 patients[14] from the registers were selected and visited by independent associations to check the accuracy of the registers[15]. A system of retroactive financial sanctions was integrated in order to reduce providers' incentives to submit fraudulent reports and register phantom patients.

In reality, the community verification system proved weak: PBF facilities only received 3 community verifications on average throughout the experiment and there was no effective financial sanction associated with being caught for fraudulent over-reporting. Specifically, the reductions in payments were proportionally equal to the percentage of patients not being identified through community verification. For example, if 18% of patients were not found through community verification, the facility would only receive a reduction of 18% in their corresponding payment and no additional sanctions were enforced. Despite the weak verification process, we did not find any significant difference in the propensity to report phantom patients in the registers[16]: the average proportion of missing patients was found to be 17% in the fixed payment group and 21% in the PBF group, this difference being non significant.

[13]Register verification was also meant to take place in health facilities under the fixed payment mechanism since the government wanted to improve the accountability of health facilities in general, not only as an element of PBF. At endline, the average number of register verifications in the last 12 months is 7 in both in the PBF and in the fixed-payment group (p-value of the test of equality of means in the two group = 0.48).

[14]The 30 patients were chosen such that each targeted service is present in the sample, but none of the non-targeted services.

[15]Community verifications were meant to take place only in the PBF group as part of the financing mechanism. However, we conducted community verifications in the fixed payment health facilities for impact evaluation purposes (1 community verification by facility in the comparison group). The fixed payment health facilities had no incentive to cheat on service volumes so the comparison of discrepancy rates between the PBF and the fixed payment groups allow for differentiating cheating from natural -unavoidable- discrepancies due to the fact that some patients moved or were absent at the time of the verification.

[16]However, the health workers in the PBF group were significantly more likely to fill out consultation reports for their patients than in the fixed payment group, so service utilization in registers was under-reported in the fixed-payment group. For that reason, it is crucial to rely on an independent source of information about service utilization, like we do in this paper, since registers not give an accurate measure service utilization in the fixed-payment group.

Pay-for-Performance and the Structure of Worker Motivation

In the context of this specific incentive scheme, the task workers have to perform is attracting more patients. The treatment changes the structure of worker motivation by adding a financial benefit of attracting patients in a context where workers already have a financial benefit of attracting more patients: in the fixed payment group, worker utility of attracting patients entails not only the intrinsic value they attribute to this task, but also the user fees. Table shows that user fees account for two thirds of facilities' revenue on average, which means that workers' incentive to attract patients is already large[17]. Importantly, as long as utility is not too concave in total revenue, the utility of attracting patients is unchanged by the presence or absence of the governmental fixed payments.

In the PBF group, the introduction of a new contingent reward from the government adds a financial benefit of engaging into attracting patients. This is likely to increase worker utility of attracting patients as long as the potential decrease in intrinsic motives and signalling effects are not too large (Benabou and Tirole 2003, 2006). After government payments are withdrawn, worker utility of attracting patients is unambiguously reduced in the PBF group compared to their past situation with incentives. Whether it ends up below, equal or above worker utility of attracting more patients in the ex-fixed payment group depends on how incentives affected the intrinsic value they attribute to the task, as well as on the resulting level of user fees.

3 Data and Empirical Strategy

3.1 Data Sources

Five sources of data are used for the impact evaluation.

Baseline Survey A survey was administered between September and November 2009. Only 85% of health facilities involved in the experiment (129 out of 152) were interviewed in this survey. As a result, we perform the balance checks on this subsample of our experimental sample.

Administrative Data Administrative data are available every month from January 2010 to December 2012 for all 152 health facilities. This data includes the number of targeted services provided, the payment due to the health facility, the actual payment made to the health facility,

[17] Workers' remuneration is provided by the facility

whether a performance verification occurred and related indicators (e.g., % missing patients and consequent financial sanctions). We use this data to examine payments received by the facilities but we do not rely on it to measure service provision and utilization since it can be both manipulated and not evenly reported in the PBF and fixed payment groups as a consequence of the incentive.

Qualitative Data In April and June 2012, qualitative interviews were conducted in 31 health facilities randomly selected in 4 out of the 8 health zones (Kafubu, Kipushi, Kasenga and Lukafu). In each facility, one interview was done with the facility head and another one with a health worker (on a voluntary basis). In total, 29 facility heads and 31 health workers were interviewed, all by the same person. They were equally distributed between the PBF group and the fixed payment group. Questions were all open and dealt with the perception of the payment (transparency, fairness, understanding of the calculation), the general functioning of the health facility, recent changes that might have occurred in the facility, and obstacles to improve the number of patients and the quality of services.

Attendance Spot Checks Unannounced spotchecks were performed in July, August and September 2012 to collect data on worker attendance in the health facilities that is impervious to gaming.

Endline Survey A final survey was administered between December 2012 and February 2013, four months after the PBF mechanism was withdrawn. The endline survey was administered in 87 out of the 96 health areas involved in the experiment. The rainy season and the insecurity created by the Maï Maï insurgency made it impossible to reach the other 9 health areas. Attrition occurred at the same rate in both groups, with 44 health areas in the PBF group and 43 in the fixed payment group included.

The endline survey included four different questionnaires for facility heads, health workers, patients straight out of consultation, and households living in the catchment area. Appendix Table 2 reports the endline sample size by questionnaire and treatment status. All facilities in the 87 health areas that could be reached were interviewed, totalizing 123 health facilities. All the technical staff in each health facility was interviewed up to 10 persons[18], totalizing 332 health workers. A sample

[18] In the facilities staffing more than 10 health workers, 10 were randomly chosen from the list of all health workers during the facility head interview. The health workers who were present the day when the interviewer visited the health facility were interviewed on-site, whereas the others were visited at home. Only those health workers who were out of the neighborhood at the time of the survey (because they were on vacation or because they temporarily migrated) could not be interviewed.

of 10 patients per facility was randomly selected for exit interviews, or the maximum available if fewer are present, totalizing 1,014 patients. Finally, the household questionnaire was administered to 1,708 households: 20 households were interviewed in each of the 87 health areas, among which 10 households randomly chosen in the population and 10 randomly chosen among the households with a pregnancy in the last 12 months[19]. Appendix Table 3 shows basic descriptive statistics of the endline sample.

3.2 Outcomes of Interest

Service Utilization

First, we measure overall health service utilization by asking each household member whether s/he visited a health facility in the last 12 months. Second, we disentangle utilization of different services: curative services, child immunization, maternal health services and family planning. For curative services we examine whether each household member visited a health facility in the last 12 months to use curative services.

For 0-5 child immunization we look at whether the mother declares that her child had at least one immunization shot and whether a scar from TB immunization could be observed on the child's shoulder. To focus on immunization when payments were in place, we restrict the sample to children aged at least 15 months at endline (at least 1 year-old when payments were withdrawn).

For maternal health services we look at whether women who have been pregnant (gave birth) in the last 12 months used prenatal (postnatal) services as well as the number of prenatal (postnatal) visits, whether delivery (if any) was attended, whether delivery (if any) was done with a c-section[20]. We focus on utilization when payments were in place by restricting the sample to women who gave birth before September 2012.

Finally, for family planning we asked each woman aged 15-49 whether she was using a modern contraceptive method: IUD, daily pill or implant. We also use whether each woman aged 15-49

[19]The selection of the 20 households was done as follows: four axes in the locality were randomly drawn from a central point, then one household was visited every five houses on each axis. - On two axes, all households were eligible and took the survey if it consented to (otherwise the next household was visited). After each interview, the interviewer went five houses further and continued the selection until he could interview 5 household on each axis. - On the two other axes, only households where a woman had been pregnant in the last 12 months were eligible. If the household did not meet the criteria, then the next household was visited etc. until an eligible household was found. After each interview, the interviewer went five houses further and continued the selection until he could interview 5 household on each axis.

[20]We also examined utilization of traditional healers and den mothers services in order to take into account potential substitution effects between modern and traditional maternal health services. However, utilization of traditional maternal services was found very low and not affected by PBF so we do not report these results in the paper for the sake of space (they are available upon request).

has had a pregnancy in the last 12 months as a direct measure of utilization of family planning. Pregnancy rate was calculated on the representative sample (randomly selected households) only since, by construction, all women in the other sample have been pregnant in the last 12 months.

Population Health Status

We use mortality rates as well as standard under-5 weight-for-age and height-for-age z-scores to assess health status. Mortality rates are measured using the number of persons who died in the last 12 months in the household, in particular the number of women who died for perinatal reasons, and the number of children under 5. We also use the proportion of new-born in the last 12 months that are still alive. To focus on impact when payments were in place, we show results restricting the sample to children born before September 2012.

Health Facility Revenue

Depending on the strategies used by the health workers and on the responses from the population, it is unclear what the effect of PBF on total resources in health facilities is. We thus examine all sources of revenue at the facility level the month before the endline survey as reported by facility heads, as well as workers' payment the month before the survey as reported by facility heads and health workers themselves.

The enumerators also observed the quantity and quality of equipment and infrastructure during their visit, which reflect both total revenue and management decisions made at the facility level. We constructed three indices, each index being the first component of a principal component analysis. The *quality index* is based on direct observation by the enumerator when s/he arrived at the facility for the endline survey of twelve items: building quality, waiting room, consultation room, lavabo, soap, clean towels, bathrooms, sterilization material, permanent display of user fees and drugs' costs, use of an examination table and ordinogram. The *infrastructure index* includes six items: phone ownership, motorized transportation mean ownership, access to clean water, toilet and electricity, and hard roof. Finally, the *equipment index* includes the quantity of fifteen types of medical equipment owned by the health facility: generator, sterilizer, tensiometer, stethoscope, baby-scales, weighing scale, height gauge, microscope, gynecological examination table, fridge, delivery boxes, fuel, kerosene, bed and solar panel.

Worker Effort

The facility's opening hours, the number of service varieties offered, and the number and qualification of workers were collected from facility heads. To examine access, patients and household members were also asked whether they could consult every time they visited. Worker attendance (number of health workers present at the facility) and on-the-job effort (number of health workers actually working) were collected from the unannounced spotchecks done by independent research assistants[21]. Regular preventive sessions at the facility help service utilization by giving greater opportunity to users to access preventive services. The number of preventive sessions organized at the facility in the last 12 months was collected from facility heads. Also, outreach activities in communities are made to inform the population about the preventive sessions (topic, day and hour). The number of outreach activities in the community in the last 12 months was collected from health workers. Using the service related to each preventive session and outreach activity, we can separate the number of activities related to targeted services (prenatal care, immunization and family planing) from the number of activites related to non-targeted services (postnatal care and HIV prevention).

Since free-riding is a concern when incentives are collective, we present some statistics on the distribution of effort within the facility using the number of outreach activities for targeted services in the last 12 months per agent. First, we show the proportion of agents who did not do any outreach activities in order to assess whether some workers changed their effort on the extensive margin. Second, we present the 25th, 50th and 75th percentiles among agents who did some outreach activities to assess whether workers changed their effort on the intensive margin, and where. Finally, for facilities with at least two agents, we present the standard deviation of the number of outreach activities per agent at the facility level to test whether the incentives changed the dispersion of effort among workers, and not simply induced a homogenous translation.

The Structure of Worker Motivation

The effect of financial incentives on the nature of worker motivation is measured using worker attendance after the payments were withdrawn on the one hand, and worker motive elicitation on the other hand.

[21] Note that the interviewer reported the number of workers present and working without telling the facility heads and the workers. The purpose of the visit was officially related to administrative matters and not attendance checks in order to avoid any interference with worker behavior at a later point. Observational data on workers' attendance and on-the-job effort was anonymous and aggregated at the facility level.

The interviewers did not announce the day they would arrive in the facility for the endline survey to avoid manipulation of staff attendance. At the time of the endline survey, workers are no longer incentivized in the PBF group so the incentive structure does no longer differ between the two groups: workers' behavior is driven by intrinsic motivation (perceived value of the job) and extrinsic motives (job remuneration). Any difference in worker behavior therefore reflects persistent effects of PBF on either intrinsic motivation, or job remuneration. Staff attendance provides a measure of workers' total motivation (intrinsic plus extrinsic).

We also elicit workers' motives in order to assess the effect of the incentives on the nature of motivation. We posit that financial incentives may draw worker attention on financial motives at the expense of non-financial motives, therefore changing the nature of motivation. To test this hypothesis, workers were asked first about the main advantage of their occupation, then about the main disadvantage. These questions were open to not induce any type of response and capture the most salient motives, those that come at the top of their mind. We classified the responses into seven categories of advantages (social recognition, remuneration, material comfort, care about others' health and life, power, interest in the activity) and six categories of disadvantages (lack of social recognition, low remuneration, low material comfort, responsibility over others' life, too much pressure and responsibility, risk of being sick due to the contact with patients). We calculate the proportion of workers who mention either remuneration or material comfort as the main advantage, or low remuneration or low material comfort as the main disadvantage. We use this proportion as a measure of the relative importance of extrinsic versus intrinsic motives in workers' total motivation.

Service Prices

A strategy to increase utilization of targeted services may be to reduce service prices. The reverse effect may happen on non-targeted services as a way to discourage utilization of those services or compensate for the loss in revenue from targeted services. User fees were collected from the facility heads at endline and from users in the last 12 months. In order to compare fees declared by facility heads across the largest number of health facilities, we used the fees of the most commonly offered services: curative consultations, birth delivery, prenatal visits, postnatal visits, and preschool consultations. To improve statistical power to detect effects that go in the same direction within a domain, we also present findings for a Fee Summary Index that aggregates information over all these user fees (following Kling et al, 2007), as well as a Fee Summary Index for targeted services

(curative and prenatal consultations, and birth delivery) and a Fee Summary Index for non-targeted services (postnatal and preschool consultations). We also collected user fees from users in the last 12 months to examine price levels when payments were in place. For preventive services, we present user fees reported by users in the last 12 months on the one hand and users before September 2012 -when PBF was implemented- on the other hand. For curative services, we were not able to apply the same strategy as we only asked about the last visit which mostly happened after September 2012.

Service Quality

Service quality is primarily measured by technical quality. Consultation time is considered as a component of service technical quality, although we consider compliance with standard medical procedures as the main indicator. Compliance was assessed on patients immediately following the consultation who consulted for illness: they were asked whether three basic procedures were followed during the consultation (being weighted, examined and having his tension checked). Compliance was also assessed on women who gave birth in the last 12 months who were asked about standard procedures applied during prenatal visits (weighing, stomach palpation, tension check, stomach measure, HIV test, tetanus shot, blood test, urine analysis and information on immunization schedule) and postnatal visits (stomach palpation, child weighing, child examination, child immunization and child immunization card). We also measure the proportion of patients straight out of consultation who visited for illness whether they were prescribed drugs without being examined, as well as the number of days women attended the facility after giving birth. Finally, as complementary measures of service quality, we use the proportion of patients who understood the diagnosis and prescriptions, as well as the proportion of patients and household members who were satisfied with the visit.

Perceived Benefit of Health Services

Perception of the benefit of health services is captured by eliciting the reason why people did not use health services: why women do not use family planning, why pregnant women do not use prenatal services, and why mothers do not use attended delivery, postnatal services and immunization[22]. The question "why don't you use this service?" was left open and the interviewer classified the

[22]Note that we failed at asking why people do not use curative services so we cannot provide evidence on their perceived benefit.

response within one of five pre-determined categories: 1) it is too expensive, 2) the waiting time is too long, 3) it takes too much time to go to the health facility, 4) I don't see the benefit, and 5) the service is poor quality. We then examine the proportion of the total population in each category (individuals who use the service receive a zero).

3.3 Empirical Strategy

Validation of the Experimental Protocol The internal validity of the impact evaluation relies on the comparability of the fixed payment and the PBF groups as observed at endline. With a large number of units of randomization, the law of large numbers insures that the characteristics in both groups are balanced. Here randomization was done on 96 health areas and it is preferable to check whether the pre-program characteristics of the fixed payment and the PBF groups are similar.

This comparison was done using the 2009 survey administered to health facilities, health workers, and randomly chosen households in the catchment area. As explained earlier, only 85% of health facilities involved in the experiment took the 2009 survey. As a result, 129 out of the 152 pilot health facilities can be observed to check how characteristics were initially balanced between the fixed payment and the PBF groups. Most initial characteristics are balanced, although the urban health facilities (17% of the sample) were not equally distributed in the PBF and fixed payment groups: they represent 12% of the PBF health facilities while 23% of the fixed payments ones. Since the urban health facilities, staff, patients and households are likely to differ from the rural ones, Appendix Table 4 presents the means of observables collected in 2009 in the PBF and fixed payment groups as well as t-tests for the following null hypothesis: the difference is zero controlling for a dummy indicating whether the unit of observation is located in a urban area. 2 differences in means are significant at the 10% level out of 59 tests, which is consistent with what would be expected with random sampling variations. It is particularly important to note that our main outcomes - utilization and health ouctomes - are balanced at baseline. We are therefore confident that differences in outcomes at endline between the two groups are not driven by initial conditions as long as we control for urban location.

Estimation Strategy For each outcome of interest, we show the estimation results of an equation of the form:

$$Y_i = \alpha + \beta PBF_i + X_i^{'}\gamma + \varepsilon_i$$

where PBF is a dummy for being in the PBF group. Because the treatment was randomly assigned, it is in expectation uncorrelated with the error term and can therefore be estimated through OLS. Coefficient β estimates the average local effect of PBF and is presented in the third column of our result tables after the unit and number of observations. We show the p-value for a test that this coefficient is equal to zero in the fourth column of the result tables.

The unit of observation i varies: it stands either for a health area, a health facility, a health worker, a patient straight out of consultation, a household, or a household member. Following the results of the balance checks discussed above, we control for a dummy indicating whether the facility is urban. To improve the precision of the estimation of the average treatment effect, we also use a small set of controls X_i which varies according to the unit of observation i: At the *health area level*, it includes a dummy variables for the health zone (the Haut-Katanga province entails eight health zones) and whether the majority of health facilities in a specific geographic area are religious. At the *health facility level*, it includes dummies indicating the health zone, and whether the health facility is religious. At the *health worker level* it also includes dummies indicating that the health worker is a female, a doctor, a nurse, as well as the age and number of years of experience of the health worker. At the *patient level* it includes a dummy indicating that the patient is a female, the age of the patient, and the reason for the visit. At the *household level*, it includes the sex and age of the household member, and for women a dummy indicating that the woman is literate. The results are robust whether or not these controls are included in the regressions. We favor the results controlling for these characteristics since it improves the precision of the estimates. Finally, we clustered error terms at the health area level to take into account potential correlation between units in the same assignment unit.

4 Results

In this section we present the impact of PBF on first facilities' performances, and then on worker effort and motivation. Third, we show and discuss the strategies that workers used to attract patients.

4.1 Performances

In this section we present the effect of PBF on facilities' performances in terms of service utilization, health outcomes and providers' revenue.

Utilization

Table 1 presents the effects of PBF on service utilization. In the control group, 50% of people visited a health facility in the last 12 months, and PBF reduced this proportion to 45% (the difference is significant at the 1% level). The detailed analysis by type of service shows that this overall decrease in utilization comes from the services whose price was reduced: curative and prenatal services. In the control group, 41% of household members used curative services at the facility in the last 12 months, while 36% in the PBF group (the difference is significant at the 5% level). Although self-declaration of sickness is generally not reliable because it is endogenous to consultation (see Akin et al. 1998), we check that the decrease in utilization is not due to a decrease in needs by focusing on people who say that they have been sick in the last 12 months: the proportion of sick people who visited a facility was reduced from 62% in the fixed payment group to 55% in the PBF group (significant at the 10% level, result not reported in Table 4). Also, take-up for prenatal visits was found to be 79% in the control group and 69% in the PBF group, resulting in a decrease of 0.4 prenatal visit per pregnant woman on a basis of 3.2 visits[23] (these differences are both significant at the 1% level). The reduction in utilization of prenatal services is similar when we restrict the sample to women who gave birth before PBF stopped, showing that the impact appeared during the exposure to PBF. We thus find a clear reduction in prenatal service take-up due to PBF, whereas at baseline there was no significant difference in utilization of this service (Appendix Table 4).

We do not find any effect of PBF on utilization of the other services: immunization, attended delivery, postnatal services, and family planning. 85% of children aged 0-5 received at least one immunization shot based on mother declaration and the enumerators could see the TB immunization scar on the shoulder of 60% of children. 82% of births were attended in a health facility over the last 12 months, and 47% of mothers used postnatal services for an average of one postnatal consultation per mother. Only 5% of women aged 15-49 were using a modern contraceptive method[24] and the fertility rate is very high: 35% of women aged 15-49 had been pregnant in the last 12 months in both the control and PBF groups (here we consider only women from the representative sample since all women in the other sample have been pregnant in the last 12 months by construction). This result indicates that the differential utilization of prenatal care services is not a composition

[23] We don't find any substitution of modern medicine for traditional medicine: the same reduction in the number of prenatal visits is found when we include visits to healers and den mothers. Note that women report a very small utilization of traditional medicine for prenatal care.

[24] Modern contraceptive methods are pill, shot, condom, IUD, spermicidal, implant and sterilization.

effect due to differential fertility. Fertility is indeed exactly similar in both groups.

Finally, women who gave birth before September 2012, during the pilot, were more likely to have a c-section. The effect size is large: while 1.1% of women had a c-section in the fixed payment group, 4% of women had a c-section in the PBF group. This result suggests that doctors performed strategically more c-sections in response to the incentive, which should be related to the fact that c-section brings a huge amount of points in this particular PBF system (see Appendix Table 1). The consequence in terms of public health is not clear since c-sections should be used to avoid difficult deliveries, although it might also be unjustified and cause negative outcomes for the mother and the child. Health outcomes are covered in the next section.

Population Health Status

Table 2 presents the effect of PBF on health outcomes. We find a deterioration in newborn and child health: height-for-age and weight-for-age z-scores were found to be lower in the in the catchment area of a PBF facility than in the catchment area of a fixed payment facility, as well as the proportion of children born in the last 12 months that were still alive, while these outcomes were identical at baseline (Appendix Table 4). The effect size is substantial: the means of weight-for-age and height-for-age in the PBF group are 0.18 standard deviations below the mean in the fixed payment group, and the proportion of newborns who did not survive in the catchment area of a PBF facility is twice as big as in the in the catchment area of a fixed payment facility. The estimates are less precise when we focus on children born before September 2012, when PBF was in place, but the point estimates are of similar magnitudes. We don't find that PBF affected the overall mortality rate in the households, nor mortality of women who gave birth in the last 12 months and children aged 0-5. The negative impact PBF had on mortality thus concentrated on children born in the last 12 months. The negative effects on newborn mortality and child weight and height may be related to the negative impact PBF on prenatal care utilization.

Health Facility Revenue

Table 3 presents the effects of PBF on facility resources, worker's payment, and the overall quality of facility infrastructure and equipment.

Total Resources at the Facility Level and Worker Payment We find 45% less total resources in the hands of PBF health facilities than fixed payment health facilities the month before

the survey (the difference is significant at the 5% level). The reduction in facility resources comes from a 56% reduction in revenue from users (significant at the 10% level). In contrast, we don't observe any difference in revenue from the government or NGOs. This result is consistent with our previous findings that PBF led to lower user fees and price of drugs, and lower service utilization, than fixed payments. According to the qualitative interviews, incentivized health workers who reduced their fees to increase demand found themselves in a situation where they were not able to re-adjust their price schedule and raise prices back to their initial values as the population had become accustomed to the reduced prices and they were fearful of reducing demand to even lower levels. As a consequence, salary to health workers was significantly lower in PBF health facilities than in fixed payment ones. We find a 34% reduction in workers' total payment in the last month as reported by the facility head, and a 28% decrease as reported by the health workers (significant at the 10% and 5% level respectively). Payment from the government are identical in the PBF and the fixed payment group, but payment from the facility itself is lowe, which is consistent with the reduced user fees and drug prices observed in the PBF health facilities' revenue.

Quality of the Facility Infrastructure and Equipment We find a significant negative impact of PBF on the quantity and quality of equipment and infrastructure. The mean *quality index* in the PBF group is 0.35 standard deviations below the mean in the fixed payment group. Most of the twelve items included in this index indicate a lower quality of equipment in the PBF facilities - negative differences are significant for four items: sink, clean towels, sterilization material and the availability of an examination table[25]. Furthermore, the mean *equipment index* in the PBF group is 0.29 standard deviations below the mean in the fixed payment group. The components of this index show that PBF facilities have consistently less equipment than the comparison ones. The differences are significant for four medical equipments: microscope, gynecological examination table, fridge and fuel. The day of the survey, the enumerator also checked the availability of five common vaccines[26] and nine common drugs[27]. We find less-than perfect –although not so bad- availability of these products: four out of five vaccines and seven out of nine drugs were available in the health facility the day of the survey. The PBF had a negative impact on the availability of vaccines the day of the survey, with fewer than 3.5 out of five vaccines available in the PBF group,

[25] However, it is worth noting that PBF facilities are more likely to permanently display the user fees and drugs' costs in the facility.

[26] Vaccines: DTP, Poliomyelitis, BCG, Measles and Yellow Fever.

[27] Drugs: oral rehydration salts, paracetamol, co-trimoxazole, ampicillin, metronidazole, quinine sulfate, mebendazole, tetracycline and Ringer's solution.

but no impact on the availability of drugs, nor on the *infrastructure index*. The negative effects of PBF on the quality index, the equipment index and the availability of vaccines are likely to be related to the reduced revenue in the PBF group. Because of the lack of resources, PBF health facilities had difficulties in investing in new equipment and repairing existing equipment.

4.2 Effort and Motivation

In this section, we examine whether the reduced performances are due to a reduction in worker effort. Facilities count on average seven workers, among which two-thirds are technical workers (doctors –only 3% of staff – nurses and birth-assistants) and one third is non-technical workers (pharmacists, managers, secretaries, receptionists and maintenance workers). It is important to notice first that contrary to the result of Bandiera et al. (2013), here the team incentive did change team composition: the number of workers, the levels of qualification, and the turnover are identical in the PBF and in the fixed payment groups (results available upon request). The finding that staff composition remained stable suggests that worker mobility was low and that financial incentives were not able to spur health workers into assortative matching by ability.

Worker Effort

Table 4 presents the impact of PBF on worker effort.

Health Facility Opening and Services Offered We find that PBF did not change the extent to which health facilities are open: according to the facility heads, facilities open on average 30 days per month and 139 hours per week. Ninety-four percent of patients and 86% of households report that they could consult every time they visited the facility. These results suggest that health facilities are generally open and that the margin of improvement in this domain is almost nonexistent. Out of a list of 23 potentially offered health services, the typical health facility offers 14 services. PBF health facilities offer the same number of targeted and non-targeted services as in the fixed payment group. PBF did thus not induce changes in the extensive margin of service supply.

Attendance We find higher staff attendance under PBF than under fixed-payment in the unannounced visits in July, August and September 2012: 58% in the fixed payment group versus 65% in the PBF group, a 14% increase, significant at the 10% level (Figure 9 in the Online Appendix

shows the distribution of staff attendance at facilities by treatment status). The higher attendance due to PBF echoes workers' statements in the qualitative interviews: "If we work a lot, we will have more money", or "We need to work many days and hours in order to have more patients." Even if the pay-for-performance mechanism reduced worker payment, workers were thus more motivated in the PBF group than in the fixed-payment group.

Preventive Sessions at Facilities Incentivized workers organized more preventive sessions at facilities in the last 12 months than non-incentivized workers (120 instead of 100, although the difference is not significant). The difference is actually driven by targeted services (immunization, prenatal care, and family planning) for which 74 preventive sessions were offered in the fixed payment group versus 106 in the PBF group (a 43% increase significant at the 5% level). For non-targeted services (postnatal care and VIH prevention), the number of preventive sessions is also higher in the PBF group but the difference is not significant[28] (see Figure 5 and 6 in the Online Appendix for the change in distributions). As a result, access to targeted health services is easier in the PBF group since a larger number of preventive sessions gives more opportunities to use the service, while access to non-targeted services remained equivalent in both groups.

Number of Outreach Activities in the Community The number of outreach activities by health workers in the community is higher in the PBF group: health workers performed an average of 22 visits to the community in the last 12 months, versus 15 in the fixed payment group, but this difference is not significant. In fact, the difference in the number of outreach activities is driven by targeted services: health workers made 16 visits to communities for these services in the PBF group, versus 10 in the fixed payment group (a 60% significant increase significant at the 10% level). In contrast, the difference in the number of outreach activities for non-targeted services is small and not significant. (see Figure 7 and 8 in the Online Appendix for the change in distributions). The population in the catchement area of a PBF facility was thus better informed about the time and day of the preventive sessions organized at that facility than the population in the catchement area in a fixed payment facility.

[28] The supply for preventive sessions is already higher for targeted services than for non-targeted services (out of 100 preventive sessions in the last 12 months, 74 were devoted targeted services and 26 to non-targeted ones). PBF thus widened this gap.

Free-riding We don't find evidence that the collective nature of the incentive led to free-riding. First, the proportion of agents who did not do any outreach activities remained equal in both groups (48%). Second, the 25th, 50th and 75th percentiles of the number of outreach activities per health worker all increased in the incentivized group, which means that effort increased over the whole counterfactual distribution. Finally, the standard deviation of the number of outreach activities per agent at the facility level is larger in the PBF group, but not statistically different from the fixed payment group. We cannot reject the hypothesis that the effect of PBF on worker effort is a pure upward translation for all workers. Altogether, these findings suggest that workers did not free-ride on others' effort.

The Structure of Worker Motivation

Table 5 presents the impact of PBF on the structure of worker motivation in terms of intrinsic versus extrinsic motives.

Attention Paid to Material versus Non-Material Benefits In the fixed payment group, 38% of workers mention spontaneously remuneration or material comfort as the main advantage or disadvantage of their position, as opposed to non material benefits (like social recognition or health benefits to the population). This proportion increases dramatically to 51% in the PBF group (a 34% increase significant at the 5% level). This finding suggests that exposure to PBF changed the salience of financial motives in health workers' mind. Importantly, this change is also unlikely to be driven by the decrease in worker salary since we observe an increase from 11% to 17% (significant at the 10% level) in the proportion of workers who mention financial benefits as the main *advantage*, while a smaller and insignificant increase in the proportion of workers who mention financial benefits as the main *disadvantage* (from 29% to 35%, p-value 0.15). This finding gives evidence of a shift in attention from the intrinsic value that the worker attributes to her job to its material benefits. We interpret this effect as evidence that incentives change the locus of control from internal to external by increasing the weight of external motives relative to intrinsic motives in worker utility.

Staff Attendance After the Incentives were Removed The positive effect of the incentive on staff attendance when incentives were in place reversed after incentives were withdrawn. The attendance rate in the fixed payment group was 57% at endline, similar to before the governmental payment was withdrawn, which confirms that the termination of fixed payments did not affect staff

effort and left worker motivation intact. In contrast, a striking reversal happened in the PBF group: the attendance rate was at 65% before the incentive was withdrawn while only 45% after. This represents a substantial and statistically significant (at the 5% level) difference in the number of present workers between the ex-incentivized and ex-non-incentivized facilities: 3.8 in the comparison group while only 2.5 in the PBF group[29] (see Figure 10 in the Online Appendix for the change in distribution). The financial incentive thus induced higher worker motivation compared to fixed payments as long as the incentives were in place, but lower motivation after the incentives were withdrawn. It is important to keep in mind that payments from the government stopped in both the PBF and the fixed payment groups at the same time, which represents the same average reduction in health facilities' revenue by design. However, payment termination equalizes back worker incentives of attracting patients in the ex-incentivized group and in the ex-fixed payment group, which should lead to a convergence of worker motivation. The fact that attendance is found lower rather than equal can result from two effects of PBF: a loss in intrinsic motivation, or the reduced user fees resulting from worker past strategy. Our previous results suggest that both mechanisms are in play.

4.3 Misplaced Strategies: Choking Under Pressure or Low Capacity?

Reduced performances when PBF was in place did thus not result from reduced effort and motivation. This finding echoes Glucksberg (1962) who observe that people who are offered a reward for performing at some tasks perform better at simple tasks but worse at tasks calling for cognitive skills. It does also echoe Ariely et al. (2009) who find that when larger stakes may lead to big mistakes. Both results are explained by the *choking under pressure* effect: rewards generate negative stress limiting one's creative thinking (Baumeister, 1984; Kamenica, 2012). McGraw and McCullers (1979) explain that rewards lead to underachievement when the task requires open-minded thinking because the focus of attention limits one's capacity to draw unusual connections between elements. Does choking under pressure explain the counterproductive effort of the incentivized health worker in the context of this experiment? The health workers experienced the incentive scheme during 28 months, so the choking-under-pressure concept needs to be extended to long-lasting situations. In this section, we review the strategies used by the health workers to better understand how the incentive scheme could have resulted in counterproductive effort.

[29] This result is consistent with declarative data from the workers: self-declared worker attendance rate in the last seven days is found 78% in the fixed payment group while 71% in the PBF group (p-value of the test of equality 0.04).

Prices of Health Services

Table 6 presents the effects of PBF on user fees and drug prices. We find consistent evidence that PBF induced a reduction in user fees for some targeted services. The mean Summary Fee Index for all services is not statistically different between both groups, but the mean Summary Fee Index for *targeted* services it is 0.81 standard deviations below the mean in the fixed payment group (significant at the 10% level). The mean Summary Fee Index for *non-targeted services* of the PBF group is 0.4 standard deviations below the mean of the fixed payment group, a sizeable difference which is not statistically significant (probably due to the lack of statistical power at the health facility level). The detailed analysis of fees by service shows that the average fee is lower for all services in PBF facilities, although the difference is significant only for prenatal[30] and curative visits[31]. Overall, the data consistently suggest that PBF encouraged health facilities to decrease the price of some targeted services, e.g. curative care and prenatal care, probably a strategy to attract more patients.

Service Quality Table 7 presents the impact of PBF on service quality. On average, patients straight out of consultation reported 16-minute consultations in both groups, while household members who visited in the last 12 months reported a slightly longer consultation time for their last consultation in the PBF group (19 minutes) than in the fixed payment group (16 minutes) (the difference is significant at the 1% level). This finding dispels the fear that incentives based on the quantity of health services would imply maximizing the number of patients at the expense of time spent with each of them. More importantly, the average compliance rates with standard medical procedures were low in both groups: 32% for curative visits, 67% for prenatal visits, and 62% for postnatal visits. 49% of patients straight out of consultation also reported that drugs were prescribed without them having been examined. On average, women stayed 2.3 days in the health facility after giving birth. These measures of technical quality were not affected by PBF. Our data thus show that technical quality is quite poor and that the incentive scheme had no impact on it.

[30] The first prenatal visit costs 442 FC in the PBF group instead of 850 FC in the fixed-payment group (a 48% decrease), and the second prenatal visit costs 51 FC in the PBF group and 132 FC in the fixed-payment group (a 61% decrease). Moreover, women who have been pregnant in the last 12 months and visited for a prenatal visit report a 21% lower fee in the PBF group than in the fixed-payment group (significant at the 5% level). This effect is even more pronounced on pregnant women who visited before September 2012 when PBF was in place: we find a 31% decrease in fees for prenatal visits (significant at the 1% level).

[31] Patients straight out of consultation who visited for a curative consultation paid 26% less for the consultation and 49% less for the drugs in the PBF health facilities than in the fixed payment facilities (these differences are significant respectively at the 5% and 1% levels)

Both users' understanding of diagnosis and medication and users' satisfaction are high: 82% of patients straight out of consultation and 94% of household members who visited in the last 12 months understood the diagnosis, and 90% of patients were aware of what drugs they were supposed to take; 94% among patients straight out of consultation and 91% among household members who visited in the last 12 months are satisfied with their visit. The main reason for being satisfied or dissatisfied of the visit is care quality, way above the second factor which is welcome quality, all other factors being negligeable[32]. Patients' satisfaction was not affected by the incentive scheme. Together, these results suggest that health workers did not reduce quality in response to the volume-based incentive, nor improved quality to attract more patients.

Perceived Benefit of Health Services

Table 8 presents the impact of PBF on perceived benefit of health services by eliciting the reasons why non-users do not use the services. We find a clear change in perceived benefit of prenatal services, which sheds light on the mechanism that led PBF to result in lower utilization of prenatal services. Indeed, we see an increase in the proportions of women who do not use prenatal services because "it is too far away" (by 155%) and "it is of little interest to me" (by 79%). Given that PBF was randomized across health areas, the actual distance of households to health facilities should be the same for the control and PBF households. Therefore, we interpret the difference in the proportion of individuals complaining about distance as a difference in motivation to get the service. We also find a small increase in the proportion who do not use because "the service is poor quality" from 0% in the control group to 1% in the PBF group. In contrast, there is no change in the proportion of women who do not use the service because of the price ("Too expensive"), nor because health workers are not available ("Long waiting time"). All in all, we interpret these findings as evidence that PBF reduced the value that women attribute to prenatal services. For the other services, we see some changes in the reasons why people do not use, although it is less clear that these changes express a change perceived benefit[33]. Among the services that we observe here, prenatal care services are thus the only ones that suffered from a clear loss in perceived benefit.

[32]Table 6 reports only reasons related to service quality. Other reasons for being satisfied or disatisfied by the visit were price and distance, which both account for very small proportions of satisfaction, and even smaller proportions of dissatisfaction.

[33]For family planning, a smaller proportion report that it is of little interest to them and a higher proportion reports that they are opposed to it. For postnatal services, a smaller proportion reports that it is of little interest to them and a higher proportion that it is too far away, a sign of lower motivation. For child immunization, we observe a decrease in the proportion of women who report that waiting time is too long, which might reflect the increase in staff attendance and in the number of preventive sessions at facility.

Misunderstanding of Users' Behavior?

In our data, prices do not appear as an important parameter neither in the decision to visit, nor in the decision not to visit. Only 3% of patients in exit interviews declared that they chose the facility because prices were attractive, while 59% mentioned proximity and 34% service quality. Similarly, 6% of individuals who visited a facility in the last 12 months declared that they chose the facility because prices were attractive, 60% because of proximity, and 26% because of service quality. Finally, 5% of women who gave birth in a facility declared that they chose the facility because prices were attractive, while 61% mentioned proximity and 32% service quality[34] (these statistics are available upon request). As for non-users, Table 8 suggests that prices is not the main reason why people do not use. The main reasons for not using the services are always "It is of little interest to me" or "It is too far away", while "It is too expensive" never accounts for more than 3% of the population. On the other hand, low technical quality suggests that there were margins for improvement in that domain, in a context where patients seem to appreciate care quality more than anything else (Table 7) and where non-users mostly ignore the benefit of health service benefits (Table 8).

In this context, selling hard and decreasing prices may not be the most adequate strategy to attract more patients, at least until more important barriers like the lack of perceived benefit are not addressed. The positive effects of hard selling and price reduction may not realize if quality is the main concern and perceived benefit is fragile. In the health sector, marketing and price reduction may have to come with an effort to increase awareness of health service benefit[35]. Conversely, it is surprising that the health workers did not invest in service quality and awareness campaigns. The important lesson of this paper is that pay-for-performance may be counterproductive when performing is difficult and capacity is low: the combination of low worker capacity and a delicate demand for health is not appropriate for an incentive scheme which requires to understand users' behavior, experiment some strategies, possibly fail, think of new strategies, and finally succeed.

[34] These numbers do not sum at exactly 100% because a handful of responses did not enter into these three categories. The question "why did you choose this facility?" was an open question, and we coded the answer into the categories "prices", "proximity", "quality" and "others".

[35] We would urge that future research examines the informational effects of marketing and prices in sectors like health, especially in contexts where the population may not be fully aware of the benefit of health services. Dupas (2014) shows that lower prices of bednets lead to higher adoption when households are offered a subsidy *and* informed about the market value of the bednet: the market value of the bednet was printed on the voucher so households were made aware of both the value of the bednet and the price reduction associated with the subsidy. Prices could thus not work as a signal for quality.

5 Conclusion

This study compares a performance-based payment mechanism to a fixed payment mechanism for health care providers in the district of Haut-Katanga, DRC. The performance-based payment studied was conditional on the number of patients for some pre-determined services, which is one specific approach of PBF. The findings show that the performance-based mechanism led to more effort by health workers to attract patients for the services included in the performance measure, without crowding out non-targeted services and service quality, nor generating score manipulation or free-riding within the facilities. However, the increased effort from the health workers was associated with a smaller utilization of some health services, and lower perceived benefit of these services by the non-user population. It led to a substantial reduction in facility revenue and worker income, as well as a deterioration in newborn and child health outcomes. These findings suggest that existing health workers cannot be treated as entrepreneurs as they were not always able to identify the successful strategies to increase demand for health services[36]. Finally, we also find that PBF created a shift in workers' attention from non-financial to financial motives apart from the reduction in worker income, and that workers decreased their effort after the removal of the incentives below its non-incentivized level.

In terms of policy lessons, these findings suggest that financial incentives should be used as a permanent policy rather than a temporary policy in order to limit the adverse effects of the motivational shift, and only in situations where the task is easy so that workers have the capacity to carry out the rewarded output, which might be a challenge in the health sector where demand is delicate. Specific interventions to stimulate demand for health should be combined with an incentive scheme, in particular interventions to improve awareness about the benefit of health services.

[36] New health workers informed about the payment system who would have self-selected in the health sector might be more able to develop appropriate strategies, as suggested by the results of Ashraf, Bandiera and Lee (2014) on the effects of carreer incentives on selection in the public health sector.

References

[1] Akin, John, David Guilkey, Paul Hutchinson and Michael Mcintosh. 1998. "Price Elasticities of demand for curative health care with control for sample selectivity on endogenous illness: an analysis from Sri Lanka", Health Economics, 7(6): 509-531.

[2] Ariely D., U. Gneezy, G. Loewenstein and N. Mazar. 2009. "Large Stakes and Big Mistakes", Review of Economic Studies, 76: 451–469.

[3] Ashraf, Nava, Oriana Bandiera and Kesley Jack. 2014. "No margin, no mission? A field experiment on incentives for public service delivery", *forthcoming* Journal of Public Economics.

[4] Ashraf, Nava, Oriana Bandiera, Scott Lee. "Do-gooders and Go-getters: carreer incentives, selection and performance in public service delivery", Working Paper, July 2014.

[5] Bandiera, Oriana, Iwan Barankay and Imran Rasul. 2005. "Social Preferences and the Response to Incentives: Evidence from Personnel Data", The Quarterly Journal of Economics, 120(3): 917-962.

[6] Bandiera, Oriana, Iwan Barankay and Imran Rasul. 2007. "Incentives for Managers and Inequality among Workers: Evidence from a Firm-Level Experiment", The Quarterly Journal of Economics, 122(2): 729-773.

[7] Bandiera, Oriana, Iwan Barankay and Imran Rasul. 2013. "Team Incentives: Evidence from a Firm Level Experiment", The Journal of the European Economic Association, 11(5): 1079-1114.

[8] Banerjee A., E. Duflo and R. Glennester, 2008, "Putting a Band-Aid on a Corpse: Incentives for Nurses in the Indian Public Health Care System", Journal of European Economic Association, 6(2-3), 487-500.

[9] Basinga P., P. Gertler, A. Binagwaho, A.s Soucat, J. Sturdy, C. 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 - 1428.

[10] Baumeister, R. F. 1984. "Choking Under Pressure: Self-Consciousness and Paradoxical Effects of Incentives on Skillful Performance", Journal of Personality and Social Psychology, 46 (3), 610–620.

[11] Benabou, R. and J. Tirole. 2003. "Intrinsic and Extrinsic Motivation", The Review of Economic Studies, 70(3): 489-520.

[12] Benabou, R. and J. Tirole. 2006. "Incentives and Prosocial Behavior", The American Economic Review, 96(5): 1652-1678.

[13] Christianson, J.B., S. Leatherman and K. Sutherland. 2008. "Lessons from Evaluations of Purchaser Pay-for-Performance Programs: A Review of the Evidence", Medical Care Research and Review, 65(6supp): 5S-35S.

[14] Deci ,E. 1975. Intrinsic Motivation. New York: Plenum Press.

[15] Deci, E. and R. Ryan. 1985. Intrinsic Motivation and Self-Determination in Human Behavior, New York; Plenum Press.

[16] De Walque, D., Gertler P., Baitista-Arredondo S., Kwan A., Vermeesch C., de Dieu Bizimina J., Binagwaho A., Condo J. 2013. "Using Provider Performance Rewards to Increase HIV testing and Counseling Services in Rwanda", World Bank Policy Research Working Paper n° 6364, Impact Evaluation Series 84, Washington DC: The World Bank.

[17] Dupas, Pascaline. 2014. "Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence from a Field Experiment", Econometrica, 82(1): 197-228.

[18] Eichler, R. and R. Levine. 2009. Performance Incentives for Global Health: Potential and Pitfalls. Washington DC: Center for Global Development.

[19] Eldridge C. and N. Palmer, 2009, "Performance-based payment:some reflections on the discourse evidence and unanswered questions", Health Policy and Planning, 24:160–166.

[20] Forsberg E, R Axelsson, B. Arnetz, 2001, "Financial incentives in health care. The impact of performance-based reimbursement", Health Policy, 58.

[21] Gertler P. and C. Vermeesch, 2013, "Using Performance Incentives to Improve Medical Care Productivity and Health Outcomes", unpublished manuscript.

[22] Glucksberg, S. 1962. "The influence of strength of drive on functional fixedness and perceptual recognition", Journal of Experimental Psychology, 63: 36–41

[23] Gneezy, U., S. Meier and P. Rey-Biel. 2011. "When and Why Incentives (Don't) Work to Modify Behavior". The Journal of Economic Perspectives, 25(4): 191-209.

[24] Holmström B. and P. Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design", Journal of Law, Economics, & Organization, 7: 24-52.

[25] Kamenica, E. 2012. "Behavioral Economics and Psychology of Incentives", Annual Review of Economics, 4: 427-452.

[26] Lazear, E. 2000. "Performance, Pay and Productivity", American Economic Review, 90(5): 1346-1361.

[27] Lepper, M, D. Greene and R. Nisbett. 1973. "Undermining Children's Interest with Extrinsic Rewards: A Test of the "Overjustification Hypothesis"", Journal of Personality and social Psychology, 28: 129-137.

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

[29] Miller, G. and K. Singer Babiarz. 2013. "Pay-For-Performance Incentives in Low and Middle-Income Country Health Programs", NBER Working Paper 18932.

[30] Meessen B, Kashala JP, Musango L, 2007, "Output-based payment to boost staff productivity in public health centres: contracting in Kabutare district, Rwanda", 85(2):108-15.

[31] Mullen, K., R. Franck and M. Rosenthal. 2010. "Can You Get What You Paid for? Pay-For-Performance and the Quality of Healthcare Providers", The RAND Journal of Economics, 41(1): 64-91.

[32] Olken Ben, Junko Onishi and Susan Wong, 2014. "Should Aid Reward Performance? Evidence From a Field Experiment on Health and Education in Indonesia", American Economic Journal: Applied Economics, 6(4): 1-34.

[33] Oxman, A.D. and A. Fretheim, 2009, "Can paying for results help to achieve the Millennium Development Goals? A critical review of selected evaluations of results-based financing", Journal of Evidence Based Medicine.

[34] Peabody JW, Shimkhada R, Quimbo S, Florentino J, Bacate MF, McCulloch C, Solon O., 2011, "Financial Incentives and Measurement Improved Physicians' Quality of Care in the Philippines." , Health Affairs 10 (4) 773-81.

[35] Rasul, Imran and Daniel Rogger. 2014. "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerien Civil Service", Working Paper.

[36] Rusa L, JD Ngirabega, W Janssen, 2009, "Performance-based financing for better quality of services in Rwandan health centres: 3 year experience", Tropical Medicine & International Health, Volume 14, Issue 7, pages 830–837.

[37] Stabile, Mark and Sarah Thomson. 2014. "The Changing Role of Government in Financing Health Care: An International Perspective", Journal of Economic Literature, 52(2): 480-518.

[38] Soeters R, F Griffths, 2003, "Improving government health services through contract management: a case from Cambodia", Health Policy and Planning, Oxford Univ Press.

[39] Soeters, R. Musango, L.Meessen, B., 2005, "Comparison of two output based schemes in Butare and Cyangugu provinces with two control provinces in Rwanda", rapport pour le GPOBA, WB Ministry of Health Rwanda.

[40] Soeters Robert, P.-B. Peerenboom, et al. 2011, "Performance Based Health Financing Experiment Improves Care in a Failed State.Performance-Based Financing Experiment Improved Health Care In The Democratic Republic Of Congo Health", Health Affairs 30(:8): 1518-1527.

[41] Sondorp E, Palmer N, Strong L, Wali E. 2008. "Afghanistan: Paying NGOs for performance". Washington, DC: Centre for Global Development.

Table 1: Impact on Service Utilization

	Unit of Observation	Number of Observations	Average Treatment Effect (ATE)	p-value (ATE=0)	Mean of Dep. Var. (Control)	St.dev. of Dep. Var. (Control)
Overall Utilization						
Visited a health facility in the last 12 months	Household Member	9113	-.051	0.006***	.4961274	.5000388
Curative Services						
Used curative care services at the health facility in the last 12 months	Household Member	9124	-.052	0.014**	.4124597	.49233
Family Planning						
Uses a modern contraceptive method	Women 15-49	1873	.005	0.69	.0505263	.2191437
Has been pregnant in the last 12 months (representative sample only)	Women 15-49	902	-.005	0.882	.3522976	.4782096
Prenatal Services						
Used prenatal care services at the health facility	Pregnant Women	1121	-.099	0.003***	.7910714	.4069067
Number of prenatal visits at the health facility	Pregnant Women	1098	-.401	0.007***	3.23133	1.899684
If gave birth before Sept. 2012 , used prenatal services at the health facility	Pregnant Women	624	-.079	0.003***	.9365079	.2442339
If gave birth before Sept. 2012 , number of prenatal visits at a health facility	Pregnant Women	603	-.373	0.054*	3.467105	1.751634
Delivery						
The mother delivered in a health facility	Mother	961	-.015	0.684	.8241309	.3810987
If delivery attended, had a C-section	Mother	773	.018	0.121	.0173697	.130807
If gave birth before Sept. 2012 , delivered at a health facility	Mother	624	-.021	0.623	.8285714	.3774827
If gave birth before Sept. 2012 and delivery attended , had a C-section	Mother	500	.029	0.071*	.0114943	.1067981
Postnatal Services						
Used postnatal care services at the health facility	Mother	960	.012	0.772	.4662577	.499371
Number of postnatal visits at the health facility	Mother	945	0	0.998	.7805383	1.193635
If gave birth before Sept. 2012 , used postnatal services at the health facility	Mother	623	-.036	0.532	.4924812	.5008858
If gave birth before Sept. 2012 , number of postnatal visits at a health facility	Mother	623	-.059	0.647	.8769231	1.282585
Child Immunization						
Ever had an immunization shot	Children 0-5	2448	-.002	0.94	.8486739	.3585063
Has a scar from tuberculosis immunization	Children 0-5	2441	.016	0.677	.6	.4900902
If aged 15 months or older , ever had an immunization shot	Children 0-5	1415	.016	0.359	.9282759	.2582087
If aged 15 months or older , has a scar from tuberculosis immunization	Children 0-5	1411	.041	0.322	.6546463	.4758135

Data Source: Endline survey. ***, **, * indicate significance at 1, 5, and 10%. Error terms are clustered at the health area level. We control for the urban/rural location of the health facility.
Unit of Observation: Pregnant Women = Women 15-49 who were pregnant in the last 12 months ; Mother = Women who gave birth in the last 12 months.

Table 2: Impact on Health Outcomes

	Unit of Observation	Number of Observations	Average Treatment Effect (ATE)	p-value (ATE=0)	Mean of Dep. Var. (Control)	St.dev. of Dep. Var. (Control)
Newborn and Child Health						
Weight-for-age z-score	Children 0-5	2402	-.178	0.044**	-1.08942	1.627851
Height-for-age z-score	Children 0-5	2376	-.157	0.094*	-1.672742	1.718997
If gave birth in the last 12 months, her child is still alive	Mother	961	-.01	0.093*	.9897751	.1007032
If born before September 2012 , weight-for-age z-score	Children 0-5	2109	-.126	0.184	-1.305793	1.577741
If born before September 2012 , height-for-age z-score	Children 0-5	2087	-.177	0.073*	-1.842894	1.782993
If gave birth before Sept. 2012 , her child is still alive	Mother	624	-.012	0.203	.9936508	.0795549
Mortality						
Number of persons in the household who died in the last 12 months	Household	1708	.007	0.732	.1366313	.4006933
Number of women in the household who died for perinatal reasons in the last 12 months	Household	1707	-.004	0.427	.009434	.0967264
Number of children under 5 in the household who died in the last 12 month	Household	1707	.012	0.55	.0896226	.3171387

Data Source: Endline survey. ***, **, * indicate significance at 1, 5, and 10%. Error terms are clustered at the health area level. We control for the urban/rural location of the health facility.
Unit of Observation: Mother = Women who gave birth in the last 12 months.

Table 3: Impact on Facility Ressources

	Unit of Observation	Number of Observations	Average Treatment Effect (ATE)	p-value (ATE=0)	Mean of Dep. Var. (Control)	St.dev. of Dep. Var. (Control)
Total Resources at the Facility Level						
Revenue from users	Health Facility	116	-258312.7	0.051*	463328.7	945376.8
Revenue from the government	Health Facility	121	6917.788	0.903	121674.5	250612.5
Revenue from NGOs and private donors	Health Facility	121	-121.757	0.555	435.4839	3429.003
Total revenue	Health Facility	116	-290462.7	0.023**	642700.7	1028920
Workers' Payment						
Payment to the Workers (reported by the Facility Head)						
Average total payment per worker in the last month (FC)	Health Facility	116	-17553.24	0.084*	50679.49	57950.96
Average wage from the government per worker in the last month (FC)	Health Facility	118	4660.341	0.083*	1731.591	5442.607
Average payment from the facility per worker in the last month (FC)	Health Facility	119	-12444.67	0.154	42580.55	47536.47
Payment to the Health Workers (reported by the Health Workers)						
Total payment in the last month (FC)	Health Worker	282	-35885.75	0.031**	127139.5	174494.9
Wage received from the government in the last month (FC)	Health Worker	326	-4999.407	0.5	23654.04	88004.44
Payment received from the facility in the last month (FC)	Health Worker	285	-28682.54	0.061*	102552.8	153866.8
Quality of the Facility Infrastructure and Equipment						
Quality index^ based on interviewers' observation (Principal Component Analysis)	Health Facility	116	-.525	0.014**	.1990995	1.511479
Infrastructure index^^ (Principal Component Analysis)	Health Facility	110	.184	0.372	-.1715342	1.425423
Equipment index^^^ (Principal Component Analysis)	Health Facility	116	-.639	0.026**	.052816	2.226755
Number of types of vaccine currently available (between 0 and 5)	Health Facility	118	-.744	0.034**	4.16129	1.738603
Number of types of vaccine that have been unavailable at some point in the last 12 months (between 0 and 5)	Health Facility	118	.036	0.929	1.52381	1.740014
Number of types of drug currently available (between 0 and 9)	Health Facility	117	.236	0.646	6.7	3.185241
Number of types of drug that have been missing once in the last 12 months (between 0 and 9)	Health Facility	111	-.276	0.589	5.333333	3.445148

Data Source: Endline survey. ***, **, * indicate significance at 1, 5, and 10%. Error terms are clustered at the health area level. We control for the urban/rural location of the unit of observation.

^The quality index includes observation on building quality, waiting room, consultation room, lavabo, soap, clean towels, bathrooms, sterilization material, permanent display of user fees and drugs' costs, use of an examination table and ordinogram.

^^The infrastructure index includes six items: phone ownership, motorized transportation mean ownership, access to clean water, toilet and electricity, and hard roof.

^^^The equipment index includes the quantity of fifteen types of medical equipment owned by the health facility: generator, sterilizer, tensiometer, stethoscope, baby-scales, weighing scale, height gauge, microscope, gynecological examination table, fridge, delivery boxes, fuel, kerosene, bed and solar panel.

Table 4: Impact on Worker Effort

	Unit of Observation	Number of Observations	Average Treatment Effect (ATE)	p-value (ATE=0)	Mean of Dep. Var. (Control)	St.dev. of Dep. Var. (Control)
Health Facility Opening and Services Offered						
Opening						
Number of opening hours per week (as reported by the facility head)	Health Facility	116	-6.522	0.524	138.9262	47.86586
Number of opening days in the last month (as reported by the facility head)	Health Facility	119	-.139	0.816	29.73016	1.885482
The patient could consult each time s/he visited	Patient	993	-.019	0.322	.9375	.2422843
The household member could consult each time s/he visited	Household Member	4323	.016	0.351	.857081	.3500661
Services Offered at the Facility						
Number of services offered by the facility (between 0 and 23)	Health Facility	123	-.492	0.35	13.55556	3.644606
Number of targeted services offered by the facility (between 0 and 10)	Health Facility	123	-.141	0.606	7.730159	1.715267
Number of non-targeted health services offered by the facility (between 0 and 13)	Health Facility	123	-.351	0.329	5.825397	2.393133
Attendance						
Av. % workers present in the facility on unannounced visits 1, 2 and 3	Health Facility	138	.074	0.067*	.5807223	.2924829
Preventive Sessions Organized at the Facility in the Last 12 Months						
Number of preventive sessions at facility provided in the last 12 months	Health Facility	118	20.084	0.291	100.4426	82.87933
Number of preventive sessions at facility provided in the last 12 months for targeted services	Health Facility	119	31.542	0.044**	73.91803	57.09679
Number of preventive sessions at facility provided in the last 12 months for non-targeted services	Health Facility	120	10.808	0.107	26.87097	31.89197
Outreach Activities by Health Workers in the Last 12 Months						
Number of outreach activities in the last 12 months	Health Worker	326	7.184	0.171	15.23295	44.47532
Number of outreach activities in the last 12 months for targeted services	Health Worker	326	5.976	0.096*	9.829545	26.42281
Number of outreach activities in the last 12 months for non-targeted services	Health Worker	326	1.208	0.523	5.403409	19.53698
Free-riding						
% agents who did not do any outreach activity for targeted services in the last 12 months	Health Worker	326	-.004	0.947	0.482954	0.501135
Among agents who did some outreach activities for targeted services, 25th percentile	Health Worker	172	3	0.006***	2	na
Among agents who did some outreach activities for targeted services, 50th percentile	Health Worker	172	6	0.092*	6	na
Among agents who did some outreach activities for targeted services, 75th percentile	Health Worker	172	9	0.077*	24	na
Facility level stand. dev. of the number of outreach activities for targeted services per agent (if more than 1 agent)	Health Facility	87	9.083	0.136	14.74143	21.81978

Data Source: Endline survey. ***, **, * indicate significance at 1, 5, and 10%. Error terms are clustered at the health area level. We control for the urban/rural location of the unit of observation.

Preventive sessions include: immunization, prenatal care and family planning (targeted services), postnatal care and HIV prevention (non-targeted services).

Outreach activities include: immunization, prenatal care and family planning (targeted services), postnatal care and HIV prevention (non-targeted services).

Table 5: Impact on Staff Intrinsic Motivation

	Unit of Observation	Number of Observations	Average Treatment Effect (ATE)	p-value (ATE=0)	Mean of Dep. Var. (Control)	St.dev. of Dep. Var. (Control)
Importance Attached to Job Remuneration						
Facility staff elicits financial benefits as the main advantage or disadvantage of his position	Facility Staff	454	.117	0.025**	.3833333	.4872145
Facility staff elicits financial benefits as the main advantage of his job	Facility Staff	452	.065	0.075*	.1087866	.3120247
Facility staff elicits financial benefits as the main disadvantage of his job	Facility Staff	454	.063	0.155	.2916667	.4554796
Staff Attendance after PBF was withdrawn						
Number of workers in the facility on unnanounced visit 4 (enline survey)	Health Facility	123	-1.354	0.032**	3.84127	3.418198
% workers present in the facility on unannounced visit 4 (endline survey)	Health Facility	123	-.121	0.099*	.5741979	.3109018
Attendance rate in the facility in the last 7 days	Health Worker	331	-.067	0.042**	.7799358	.1429585

Data Source: Endline survey. ***, **, * indicate significance at 1, 5, and 10%. Error terms are clustered at the health area level. We control for the urban/rural location of the health facility.

Table 6: Impact on User Fees

	Unit of Observation	Number of Observations	Average Treatment Effect (ATE)	p-value (ATE=0)	Mean of Dep. Var. (Control)	St.dev. of Dep. Var. (Control)
User fees as reported by facility heads at endline						
Fee Summary Index	Health Facility	93	-1.077	0.141	.166473	4.212105
Fee Summary Index, targeted services	Health Facility	109	-.807	0.061*	.0366889	2.866472
Fee Summary Index, non-targeted services	Health Facility	95	-.398	0.346	.1007338	2.064238
Targeted Services						
User fee for the first curative consultation (FC)	Health Facility	123	-692.45	0.281	1263.492	4557.316
User fee for delivery (FC)	Health Facility	113	-224.185	0.655	2747.414	2423.25
User fee for the first prenatal visit (FC)	Health Facility	118	-407.873	0.095*	850	1741.42
User fee for the second prenatal visit (FC)	Health Facility	115	-80.801	0.053*	132.2034	264.8622
Non-Targeted Services						
User fee for the second curative consultation (FC)	Health Facility	112	-178.082	0.18	459.4828	799.0377
User fee for postnatal visit (FC)	Health Facility	111	-57.43	0.386	105.3571	430.8215
User fee for preschool consultation (FC)	Health Facility	112	-6.718	0.838	86.66666	154.8281
User fees as reported by users in the last 12 months						
Fee the last postnatal visit (if any in the last 12 months) (FC)	Mother	388	-40.458	0.544	349.2684	585.9773
Fee for attended delivery (if any in the last 12 months) (FC)	Mother	762	-234.426	0.732	8768.171	6910.282
Fee the last prenatal visit (if any in the last 12 months) (FC)	Pregnant Women	918	-120.798	0.028**	583.4368	721.459
Fee the last immunization visit (if any in the last 12 months) (FC)	Children 0-5	2039	-22.096	0.237	87.71028	316.9161
Fee for the present curative visit (if the reason for visit was illness) (FC)	Patient	718	-1897.282	0.034**	7311.323	16030.53
Cost of drugs at the facility for the present curative visit (if the reason for visit was illness) (FC)	Patient	549	-1581.249	0.002***	3628.322	6160.657
User fees as reported by users *before September 2012*						
Fee the last postnatal visit (if any before September 2012) (FC)	Mother	227	-32.896	0.637	315.7232	539.3551
Fee for attended delivery (if any before September 2012) (FC)	Mother	493	463.057	0.546	8726.834	6926.556
Fee the last prenatal visit (if any before September 2012) (FC)	Pregnant Women	581	-187.611	0.001***	598.0456	706.0172
Fee the last immunization visit (if any before September 2012) (FC)	Children 0-5	508	18.014	0.489	73.16177	211.3834

Data Source: Endline survey. ***, **, * indicate significance at 1, 5, and 10%. Error terms are clustered at the health area level. We control for the urban/rural location of the unit of observation.
Fee Summary Index is the equally weighted average of z-scores of its components. The z-scores are calculated by subtracting the control group mean and dividing by the control group standard deviation. The components of the index are fees paid for first and second curative consultations, delivery, prenatal and postnatal visits, and preschool consultation.
Targeted services are: first curative consultation, delivery, and prenatal visits.
Non-targeted services are: second curative consultation, postnatal visit, and preschool consultation.
Unit of Observation: Pregnant Women = Women 15-49 who were pregnant in the last 12 months ; Mother = Women who gave birth in the last 12 months.

Table 7: Impact on Service Quality

	Unit of Observation	Number of Observations	Average Treatment Effect (ATE)	p-value (ATE=0)	Mean of Dep. Var. (Control)	St.dev. of Dep. Var. (Control)
Technical Quality						
Consultation time (minutes)	Patient	974	1.028	0.422	16.09263	15.51822
If visited a health facility, consultation time (minutes)	Household Member	4265	2.506	0***	16.125	11.96238
If visited for illness, compliance rate with medical procedure for the present curative consultation	Patient	713	-.008	0.805	.3239075	.2992397
If visited for illness, drugs were prescribed to the patient and the patient was not examined	Patient	719	-.036	0.495	.4936387	.5005968
If delivery attended, number of days in the health facility after the delivery	Mother	746	.052	0.733	2.123711	1.170328
If any postnatal visit, compliance rate with medical procedures for the last postnatal consultation	Mother	389	.048	0.123	.6166667	.258334
If any prenatal visit, compliance rate with medical procedures for the last prenatal consultation	Pregnant Women	923	.004	0.818	.6657578	.1680248
Users' Understanding						
If visited for illness, patient understands diagnosis and next steps	Patient	720	-.001	0.971	.822335	.3827164
If visited for illness, patient knows what drugs to be taken	Patient	718	-.046	0.187	.8982188	.3027457
If visited a facility, household member understands diagnosis	Household Member	4258	.017	0.241	.9372237	.2426138
Users' Satisfaction						
The Patient reports that s/he was…						
satisfied	Patient	994	.013	0.359	.9430147	.2320279
satisfied thanks to care quality	Patient	990	.003	0.937	.5722222	.4952152
satisfied thanks to welcome quality	Patient	990	-.027	0.442	.2796296	.4492334
satisfied thanks to equipment quality	Patient	990	0	0.997	.0333333	.1796719
dissatisfied because of care quality	Patient	993	-.005	0.671	.0349265	.1837626
dissatisfied because of welcome quality	Patient	993	0	0.946	.0073529	.0855121
dissatisfied because of equipment quality	Patient	993	-.006	0.359	.0110294	.1045364
If visited a health facility, the Household Member reports that s/he was…						
satisfied	Household Member	4326	.004	0.778	.9142857	.2800023
satisfied thanks to care quality	Household Member	4318	-.005	0.857	.7417678	.4377572
satisfied thanks to welcome quality	Household Member	4318	-.008	0.547	.0836222	.2768804
satisfied thanks to equipment quality	Household Member	4318	.001	0.855	.0186308	.1352467
dissatisfied because of care quality	Household Member	4312	-.002	0.853	.0487593	.2154112
dissatisfied because of welcome quality	Household Member	4312	-.001	0.844	.0104484	.1017042
dissatisfied because of equipment quality	Household Member	4312	.001	0.76	.008707	.0929245

Data Source: Endline survey. ***, **, * indicate significance at 1, 5, and 10%. Error terms are clustered at the health area level. We control for the urban/rural location of the health facility.

Unit of Observation: Household Member = any person in the sampled households; Pregnant Women = Women 15-49 who were pregnant in the last 12 months; Mother = Women who gave birth in the last 12

Table 8: Impact on Perception of Health Service Benefit

Do not use the service because...	Unit of Observation	Number of Observations	Average Treatment Effect (ATE)	p-value (ATE=0)	Mean of Dep. Var. (Control)	St.dev. of Dep. Var. (Control)
Family Planning						
… It is too expensive	Women 15-49	1876	-.002	0.407	.0052521	.0723188
… It is too far away	Women 15-49	1876	-.008	0.462	.0346639	.1830231
… It is of little interest to me / I don't know what it is / how it works	Women 15-49	1876	-.061	0.089*	.4705882	.4993966
… I am opposed to it	Women 15-49	1876	.114	0.003***	.2079832	.4060781
… I want to get pregnant	Women 15-49	1876	-.007	0.622	.0577731	.2334365
Prenatal Services						
… It is too expensive	Pregnant Women	1119	.004	0.722	.0268336	.1617417
… Waiting time for consultation is too long	Pregnant Women	1119	.001	0.613	.0017889	.0422955
… It is too far away	Pregnant Women	1119	.042	0.009***	.0268336	.1617417
… It is of little interest to me / I don't know what it is / how it works	Pregnant Women	1119	.033	0.045**	.0572451	.2325184
… The service is poor quality	Pregnant Women	1119	.006	0.075*	0	0
Attended Delivery						
… It is too expensive	Mother	960	.015	0.127	.0184426	.1346836
… Waiting time for consultation is too long	Mother	960	0	0.	0	0
… It is too far away	Mother	960	-.01	0.74	.1331967	.3401359
… It is of little interest to me / I don't know what it is / how it works	Mother	960	.007	0.451	.022541	.1485871
… The service is poor quality	Mother	960	.004	0.238	0	0
Postnatal Services						
… It is too expensive	Mother	954	-.006	0.626	.0349076	.1837345
… Waiting time for consultation is too long	Mother	954	-.006	0.287	.0102669	.1009081
… It is too far away	Mother	954	.089	0.059*	.1765914	.3817146
… It is of little interest to me / I don't know what it is / how it works	Mother	954	-.089	0.006***	.2956879	.4568206
… The service is poor quality	Mother	954	-.004	0.661	.0143737	.1191481
Child Immunization						
… It is too expensive	Children 0-5	2446	-.006	0.174	.0101404	.1002268
… Waiting time for consultation is too long	Children 0-5	2446	-.018	0.01**	.024181	.1536707
… It is too far away	Children 0-5	2446	.005	0.5	.0218409	.1462208
… It is of little interest to me / I don't know what it is / how it works	Children 0-5	2446	.017	0.43	.0951638	.2935556
… The service is poor quality	Children 0-5	2446	.002	0.29	0	0

Data Source: Endline survey. ***, **, * indicate significance at 1, 5, and 10%. Error terms are clustered at the health area level. We control for the urban/rural location of the health facility.
Unit of Observation: Pregnant Women = Women 15-49 who were pregnant in the last 12 months ; Mother = Women who gave birth in the last 12 months.

Appendix Table 1: Relative Prices of Targeted Health Services

Service	Indicator	Relative Price (USD)
Services targeted at health centers and referral health centers		
Curative care	Per new curative consultation	$0.6
Institutional delivery	Per delivery at the health center	$5
Obstetric referral	Per pregnant woman referred to the referral center/hospital	$5
Full childhood immunization	Per fully immunized child	$3.5
Prenatal care	Per prenatal care consultation	$1.2
Tetanus toxoid vaccination	Per 5th dose of tetanus toxoid vaccination	$2
Family planning	Per woman that uses a modern method of family planning	$4.5
Additional services targeted only at referral health centers:		
Caesarean section	Per caesarean section delivery (and decision-tree has been followed)	$30
Blood transfusion, when appropriate	Per transfusion episode	$5
Obstetric referral	Per delivery referred to the referral center/ hospital"	$5

Appendix Table 2: Endline Sample

Endline Sample, by Payment Status

	PBF Group	Comparison Group	Total
Health areas	44	43	87
Health Facilities	60	63	123
Facility Staff	154	178	332
Patients	470	544	1,014
Households	859	849	1,708
Household members	4,578	4,656	9,234
Women 15-49	939	957	1,896
Pregnant Women*	571	560	1,131
Mothers**	479	489	968
Children 0-5	1,228	1,285	2,513

*Pregnant Women = Women who have been pregnant in the last 12 months
**Mothers = Women who gave birth in the last 12 months

Appendix Table 3: Descriptive Statistics at Endline

	Mean	Standard Deviation	Nb. of Observations
A. HEALTH FACILITY			
The facility is a "Centre de Santé de Référence"	0.11	0.31	123
The facility is a "Centre de Santé"	0.69	0.46	123
The facility is a "Poste de Santé"	0.20	0.40	123
The facility is public	0.66	0.48	123
The facility is religious	0.15	0.36	123
The facility is private	0.19	0.39	123
The facility is urban or semi-urban	0.17	0.38	123
The facility is rural	0.83	0.38	123
Served population size	12872.76	11570.57	123
B. HEALTH WORKERS			
The health worker is a female	0.57	0.50	332
Age of the health worker (years)	42.14	11.20	332
The health worker is a doctor	0.06	0.23	332
The health worker is a nurse	0.57	0.50	332
Number of years of experience	12.56	10.13	331
C. PATIENT			
The patient is a female	0.67	0.47	1006
Age of the patient (years)	18.61	17.39	1002
D. HOUSEHOLDS			
The household member is a female	0.50	0.50	9225
Age of the household member (years)	17.17	16.13	9135
The household member is literate (if aged 15 or more)	0.57	0.49	4166

Appendix Table 4: Balance Checks

	Unit of Observation	Number of Observations	Mean in the control group	Standard Deviation in the control group	Difference in mean in the treatment group	p-value (difference=0)
Health Facilities						
Health Center (versus Health Post)	Health Facility	129	.78125	.4166667	-.038	0.623
Public	Health Facility	129	.59375	.4950148	-.029	0.734
Private	Health Facility	129	.28125	.4531635	-.086	0.276
Religious	Health Facility	129	.125	.3333333	.115	0.119
Number of years of activity	Health Facility	122	20.18333	22.42539	.266	0.948
Catchement population	Health Facility	122	11129.3	15802.48	1255.156	0.669
Catchement area (km2)	Health Facility	109	368.963	826.58	-19.957	0.892
Number of beds	Health Facility	129	8.953125	13.43229	1.23	0.536
Number of workers	Health Facility	129	6.359375	5.524454	-.162	0.866
Infrastructure Index^	Health Facility	129	1.00e-08	.5614322	.057	0.525
Equipment Index^	Health Facility	128	-.0031828	.524324	.116	0.352
Medical Material Index^	Health Facility	129	-5.01e-09	.4828278	-.164	0.147
Stock of Vaccines Index^	Health Facility	125	-7.71e-10	.8745002	-.17	0.144
Health Workers						
Female	Health Worker	457	.4810127	.5006968	-.101	0.042**
Age (years)	Health Worker	456	40.31224	10.94959	-.101	0.932
Doctor	Health Worker	457	.0421941	.2014572	.016	0.29
Qualified Nurse	Health Worker	457	.2362869	.4256995	-.013	0.708
Non-qualified Nurse	Health Worker	457	.3122363	.4643864	-.011	0.801
Midwife	Health Worker	457	.1561181	.3637355	.006	0.849
Adjunct	Health Worker	457	.1687764	.375347	.003	0.952
No education	Health Worker	457	.0759494	.2654777	-.014	0.747
Primary Education	Health Worker	457	.0801688	.272129	-.016	0.612
Secondary Education	Health Worker	457	.3122363	.4643864	.061	0.262
Higher Education	Health Worker	457	.2278481	.4203318	.012	0.822
Experience (years)	Health Worker	455	10.64255	10.16824	-1.489	0.137
Income (Francs Congolais)	Health Worker	304	69508.57	69909.81	2084.831	0.837
Satisfied in the current position	Health Worker	457	.4767933	.5005182	.057	0.262

Appendix Table 4 (continued): Balance Checks

	Unit of Observation	Number of Observations	Mean in the control group	Standard Deviation in the control group	Difference in mean in the treatment group	p-value (difference=0)
Households						
Muslim	Household	1059	.0288462	.1675352	.001	0.931
Christian	Household	1059	.9076923	.2897385	-.039	0.065*
Animist	Household	1059	.0192308	.1374674	.005	0.668
Housing Index^	Household	1059	-.0003307	.6473751	-.076	0.349
Female	Household Member	6816	.4944524	.5000404	-.005	0.593
Maried	Adult 15+	3845	.5153217	.4998929	.032	0.154
Single person	Adult 15+	3845	.4147089	.4927976	-.017	0.445
Attended some school	Adult 15+	5431	.7348323	.4415013	-.033	0.266
Completed primary school	Adult 15+	2091	.4817245	.4999002	-.031	0.349
If was ever pregnant, visited a health facility during the last pregnancy	Pregnant Women	1017	.7447217	.4364366	-.015	0.691
If was ever pregnant, number of visits to a health facility during the last pregnancy	Pregnant Women	989	2.84466	2.235882	-.119	0.529
If ever had a child, the last delivery was attended	Mother	862	.7112676	.4537061	-.046	0.379
If ever had a child, visited a health facility after the last delivery	Mother	846	.2805755	.4498201	.011	0.785
If ever had a child, the last child is still alive	Mother	859	.9270588	.2603464	-.005	0.777
If ever had a child, weight of the last child at birth (grams)	Mother	521	3262.886	623.8203	10.781	0.862
Age in months	Children 0-10	2654	57.73658	36.12342	-.756	0.537
Immunization card seen by the interviewer	Children 0-5	1351	.122093	.3276314	.02	0.409
If immunization card seen, had BCG	Children 0-5	177	.797619	.4041878	-.034	0.653
If immunization card seen, had Polio	Children 0-5	172	.5432099	.501233	.017	0.859
If immunization card seen, had DTC	Children 0-5	177	.7261904	.4485906	.08	0.348
If immunization card seen, had Measles	Children 0-5	175	.4938272	.503077	-.015	0.878
If immunization card seen, had Vitamins	Children 0-5	178	.2738095	.4485906	-.05	0.489

Data Source: Baseline survey. ***, **, * indicate significance at 1, 5, and 10%. Error terms are clustered at the health area level. We control for the urban/rural location of the health facility.
^Each Summary Index is the equally weighted average of z-scores of its components. The z-scores are calculated by subtracting the control group mean and dividing by the control group standard deviation. The components of the Infrastructure Index are dummies of whether the health facility has a water tap, electricity, a waste disposal, a sewage disposal, and a pharmacy. The components of the Equipment Index are dummies for whether the health facility has a phone, a radio, the number of electricity generators, the number of examination tables, the number of refrigerators, the number of fuel liters, and the number of kerosene liters. The components for the Medical Material Index are dummies for whether the health facility has autoclaves, tensiometers, sthetoscopes, scales, gauges, microscope, and a delivery kit. The Stock of Vaccines Index components are the number of BCG vaccines, DTC vaccines, Polio vaccines, measles vaccines and anti-amariale vaccines in the facility. The Housing Index components are dummies for whether the household housing has a water tap, sanitation, garbage collection, and some energy (fuel or electricity).