

WORLD BANK POLICY NOTE

**Impact Evaluation on Performance-
Based Financing In Haut-Katanga District,
Democratic Republic Of The Congo**

TABLE OF CONTENTS

INTRODUCTION.....	1
BACKGROUND	3
Description of the PARSS project and impact evaluation	3
The design of the intervention and impact evaluation	4
RESULTS	9
LINKS BETWEEN PAYMENT SCHEME DESIGN CHARACTERISTICS AND IMPLEMENTATION OUTCOMES.....	15
Well-balanced benefit package at all levels	17
Rigorous results verification	17
Use of client satisfaction surveys.....	18
Use of a quantified quality checklist with the result tied to payments	18
Use of a fee-for-service provider payment mechanism	18
Individual fees and total earnings that are significant and paid regularly	18
RECOMMENDATIONS	21
REFERENCES	23

INTRODUCTION

This policy note provides an overview of policy issues pertaining to the experience of a Performance-based Financing (PBF) pilot in the district of Haut-Katanga in the Democratic Republic of the Congo (DRC). It gives an overview of the main results of the pilot and its impact evaluation and discusses design and implementation issues that emerged during the implementation of the program. The note is structured as follows; Section 2 provides a background to the Haut-Katanga PBF pilot methodology for the impact

evaluation; Section 3 presents the findings and recommendations of the impact evaluation, and Section 4 gives an assessment of the design attributes of the pilot and the link between these attributes and results observed. The final section (Conclusion) highlights the design attributes recommended based on experiences in DRC and elsewhere for improved results for Performance-based Financing schemes in DRC and other similar contexts.

BACKGROUND

The Democratic Republic of Congo is the second largest country in Africa by area, with the fourth largest population at 66 million (World Bank, 2013). 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), with an estimated per capita income of US\$ 220 (current) in 2012 (World Bank, 2013). Impoverished by decades of war, instability and bad governance, it is not surprising that DRC is not on track to reach the health-related MDGs.

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. The government's 2006 health sector strategy emphasizes development of the health zone system, indicating that contractual approaches and public-private partnerships will increasingly be options for health service financing and delivery. Developing and putting in place effective service delivery models such as Results-based Financing (RBF) would be a strategy for improving health outcomes among the population.

Results-Based Financing (RBF) is an instrument that links financing to pre-determined results, with payment made only upon verification that the agreed-upon results have actually been delivered. Different RBF mechanisms include Output-Based Aid (OBA), Performance-Based Financing (PBF), or other provider payment mechanisms that link payments to service volume and quality, and conditional cash transfers. The central theme in all of these is that a principle entity provides a reward, conditional on the recipient undertaking a set of actions to produce a desired outcome.

Even though performance-related payment models have been implemented in developed and developing countries in various settings and forms, the scientific evidence base on the impact of these mechanisms on specific outcomes remains thin. In Rwanda, PBF proved an efficient way to increase health service quality and utilization, resulting in improved child health outcomes (Basinga et al., 2011). In Argentina, RBF led to improved birth weights, a reduction in neonatal mortality, and an increase in the use of prenatal services (Gertler et al., 2013).

Description of the PARSS project and impact evaluation

The district of Haut-Katanga (1.26 million people in the province of Katanga in the south-eastern corner of the DRC) is part of the \$335 million World Bank supported Health Sector Rehabilitation and Support project (*Projet d'Appui à la Réhabilitation du Secteur de la Santé – PARSS*) which covers 83 health zones in ten districts across five provinces of DRC. The objective of the project is to ensure that the target population of selected health zones has access to a well-defined package of quality essential health services by supporting, inter alia, construction, the provision of equipment and medicine, training and salary top-ups. In addition, health facilities were subject to a user fee reduction policy in which prices of essential health services were either reduced or eliminated. Descriptive statistics of Haut-Katanga district relative to national statistics are provided in Table 1.

From 2009 to 2013, an impact evaluation was conducted in the Haut-Katanga district of DRC to study the effect of different types of provider payment

TABLE 1: Socio-economic and health characteristics at baseline, Haut-Katanga district vs. national level

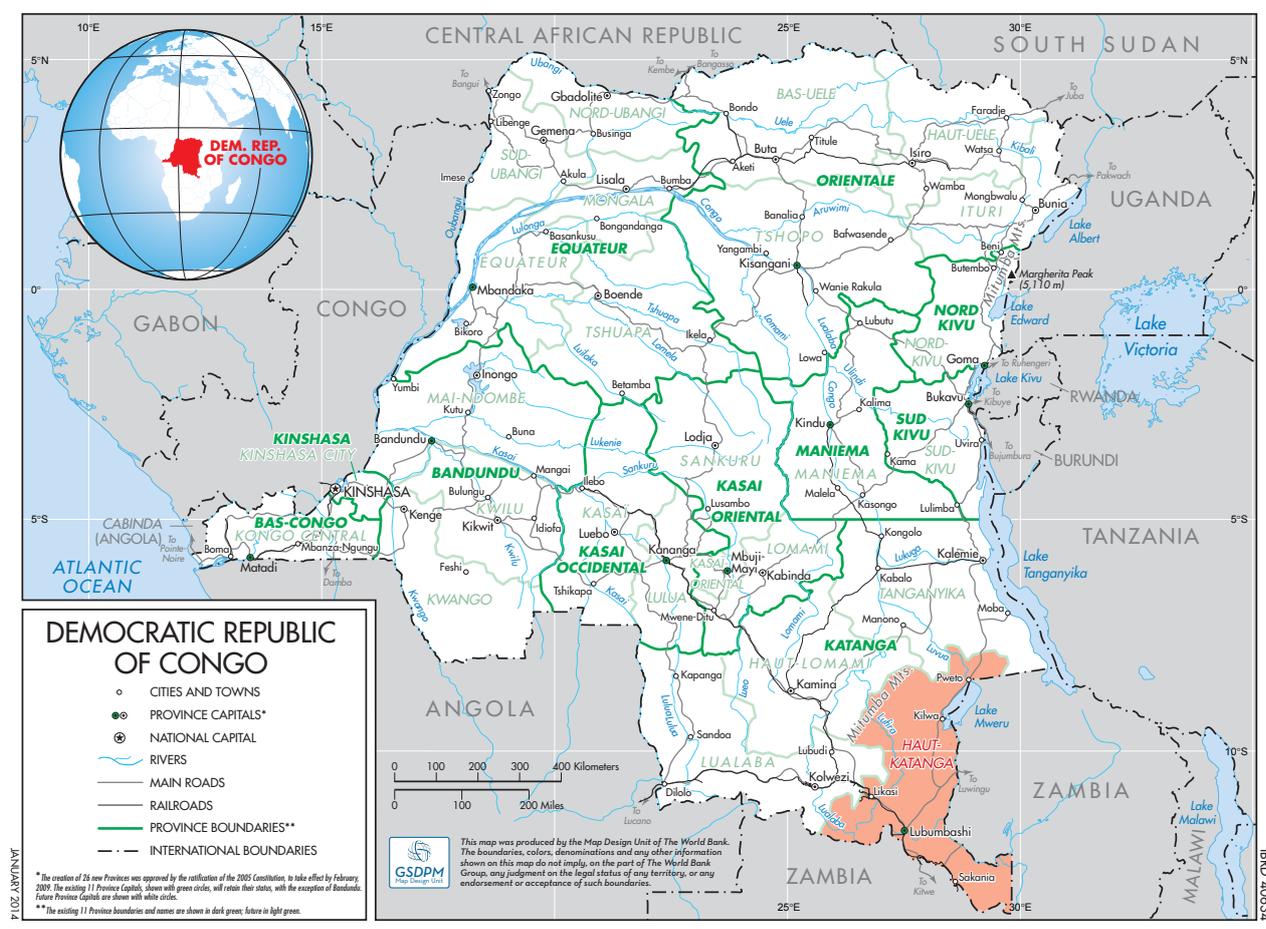
Indicator	Haut-Katanga District (IE baseline, 2009)	National (DHS, 2007)
General characteristics		
Household size	6	5
% female	49	51
% married	53	58
Household goods (%)		
Radio	35	45
Television	7	14
Refrigerator	1	4
Bicycle	65	24
Motorcycle	2	1
Housing has electricity	5	15
Maternal health (%)		
Any antenatal care	76	88
At least 4 antenatal care visits	40	46
Delivery in health facility	69	71
Delivery by caesarian-section	10	4

mechanisms on outcomes such as the quantity of services provided by contracted health facilities. The primary objective of the Haut-Katanga impact evaluation was to compare the effect of two financing mechanisms (one conditional on the quantity of services delivered and the other a lump sum payment based on the number and cadre of facility staff) on health service delivery outcomes. More precisely, the impact evaluation aims at analyzing the effect of the financial incentive mechanism linked to service volume on (i) the availability of services, (ii) the price of health services and cost to patients, (iii) health workers' satisfaction, work-related stress and motivation, (iv) service utilization, (v) patient satisfaction and (v) the population health status. The Haut-Katanga impact evaluation design involved randomly assigning health areas (the unit of randomization) to intervention and comparison groups. The average causal effect of the performance-related subsidies on the outcomes of interest could thus be captured by differences in outcomes across the intervention and comparison groups.

The design of the intervention and impact evaluation

Different to the other nine districts included in the project (which did not have any PBF-type intervention), in Haut-Katanga district a pilot was launched which involved the introduction of performance-based payments to health centers and referral centers using a "point system" linked to the volume of targeted services delivered by contracted facilities. To measure the effect of this financial incentive payment mechanism (treatment group), the impact evaluation introduced a predictable monthly transfer for the comparison group. In order to ensure neutrality in the level of financing between the two groups and allow for comparison of the two payment mechanisms independent of the level of money put into the system, the impact evaluation design was such that the total budget allocated to performance bonuses across health facilities in the intervention group was to be the same as the total budget for lump sum payments allocated across all health facilities in the comparison group (thus eliminating the risk of varying

FIGURE 1: District of Haut-Katanga



levels of financing between the two groups affecting outcomes of interest). Payments were to be made on a monthly basis after verification of declared results by facilities in the treatment group. **The impact evaluation, thus, compares the results of a lump-sum monthly transfer to a transfer that is tied to the quantity of services provided through use of the point system.**

The design of the performance-based strategy was kept simple so that it could be feasibly implemented in the post-conflict setting of Haut-Katanga. Consequently, the strategy lacked some of the more technically complex attributes that are often seen in PBF interventions, such as (i) stringent quality measures tied to performance payments, (ii) business plans, (iii) quality investment units, (iv) individual health worker performance evaluations, (v) indice

tool, (vi) intense coaching for the introduction of advanced strategies, (vii) fee-exemptions and waiver schemes financed through PBF targeting the poor, and (viii) rural hardship fee-adjustments, which while desirable, take a longer time and are more difficult to implement successfully. Successful PBF schemes¹ typically also include the referral hospital, and the health administration in performance contracting (Fritsche et al., 2014 (forthcoming)). Furthermore, the budget used in the Haut-Katanga scheme was estimated at \$0.43 per capita per year (average monthly facility payments were \$550 and the average catchment area population was 12,900) was much lower than is typical in similar contexts where output budgets range between \$2 and \$3 per capita per year).

¹ For example in Rwanda, Burundi, and Cameroon

TABLE 2: Summary of design elements of Haut-Katanga pilot

PBF design element	Katanga Pilot, DRC
Benefit package of targeted services	7 services covered at primary health center level and 3 at referral health centers ^a
Results verification	Monthly verification of consistency between reported volumes and health facility registers; Only few community verifications to verify if the information in the registers was true; sanctions for discrepancies were light
Community client satisfaction surveys	No client satisfaction element included
Quality assessment	No use of the quantified quality checklist; no quality measures tied to payment
Fee-for-service provider payment mechanism	Not real fee-for-service; more of a weighting for relative value, and then apportioning available budget based on value obtained
Payment schedule and levels	Supposed to be monthly but irregular payments; total earnings remained the same during the intervention at \$550 per month per health facility on average

^a The incentivized services included seven services at the primary care level (outpatient first curative consultations, prenatal consultations, deliveries, obstetric referral, children completely vaccinated, tetanus toxoid vaccination, and family planning consultations) and three at the secondary care level (C-section, blood transfusion, and obstetric referrals to hospitals).

Table 2 summarizes the design elements of the Haut-Katanga pilot.

The “point system” was such that a particular number of points was given for each service that was provided (for services specified in the contract). After adding the number of points a score was produced for each facility. Each facility’s score was then divided by the sum of all scores in the treatment group to produce the percentage of the total budget that should be allocated to each facility. *Facilities’ payment was therefore determined by the quantity of services provided relative to the other health facilities. In addition, payments in the treatment group were not linked to the quality of care provided by these facilities; the incentive scheme was linked to the relative volumes produced by each facility.*

In contrast, the amount allocated to each facility in the comparison group was calculated based on the staff in the facility. A list of eligible workers was established at the beginning of the pilot conjointly by the PARSS project and Ministry of Health and each worker received a given payment depending on his/her grade and experience. Health facilities were then randomly assigned to either the treatment or comparison group.

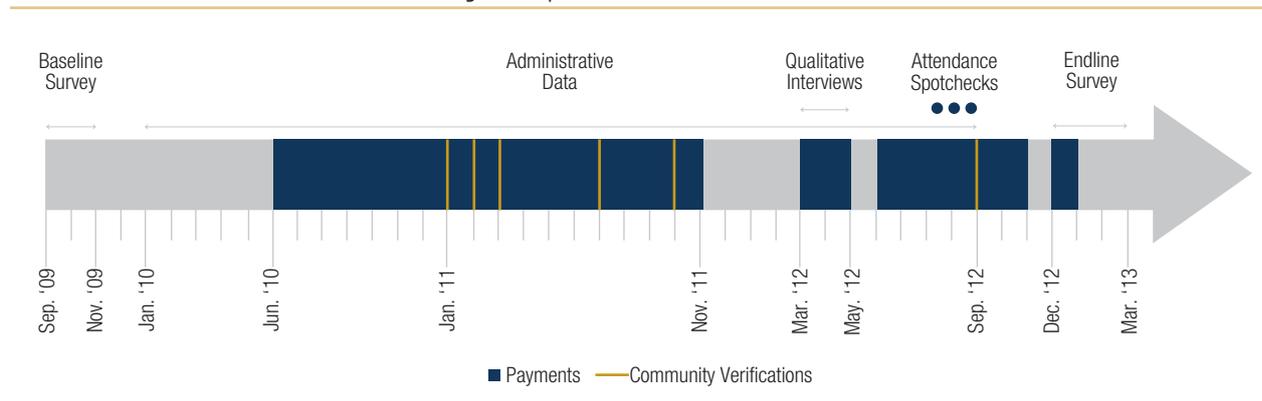
According to the design of the pilot, the volume of services was measured first by use of monthly Health Management Information System (HMIS) reports

submitted by facilities in which the number of services delivered each month were reported. Subsequently verification of declared service volumes was conducted by verification agents through (i) comparing reported volumes with those found in health facility registers, and (ii) verifying that the information noted in the registers was true by conducting community verifications.² In principle, verification of facility registers was to be conducted monthly and community verification quarterly. In reality, community verification occurred less frequently than expected and only with a sample of health facilities during each round of verification.

It is important to note that the comparison of the two payment mechanisms was conducted within the overall context of the PARSS intervention, where both groups were benefitting similarly from construction/rehabilitation, provision of equipment and medicine, staff training, and user fee reduction policy. However, the treatment and the comparison groups varied in two aspects. First, payments in the treatment group were supposed to be contingent on verification of declared results while payments in the comparison groups were solely based

² Community verifications were conducted by selecting a random sample of patients from registers and visiting them to ensure that they actually visited the facility and received the services noted in the facility registers.

FIGURE 2: Timeline of the Haut-Katanga PBF pilot



on the number of health staff officially employed at the health facility. Second, in the treatment group health facility managers were granted increased financial autonomy to distribute the payment subsidies among facility staff, whereas in the comparison group facility managers were expected to distribute bonus payments according to grade and classification. In the treatment group, health facilities' autonomy was limited to the distribution of these performance bonuses while use of resources from other categories (functioning, investment) remained rigid. Facilities had to maintain separate budget lines for bonuses, functioning and investment. Equally important, in Haut-Katanga health facilities transferred a set percentage of their revenue from drug sales to the Provincial Inspection.

Implementation of the PBF pilot started in 2009 and was first managed by an international NGO, International Rescue Committee, who was then replaced by the PARSS project implementation unit (UCP) within the Ministry of Health. Activities for IRC/PARSS included verification of service volumes declared by facilities, calculation and processing of payments, and contracting local organizations for conducting community verifications. The first payments were made in June 2010 and ended in September 2012 with several interruptions in the payment schedule over the two year period. Due to several challenges encountered during the implementation of the pilot, the actual design of the Haut-Katanga PBF scheme differed substantially from what was originally original designed. Table 3 summarizes the

differences between the planned and actual design of the PBF scheme.

Five sources of data were used for the impact evaluation. First, a baseline household and health facility survey was administered in Sept–Nov 2009. Second, administrative data from IRC/PARSS was collected from January 2010 to December 2012. Third, a qualitative survey was administered to a sub-sample of 30 health facilities in April–May 2012. Fourth, unannounced visits (spot-checks) were performed in July, August and September 2012 in order to collect objective data on worker attendance in the health facilities. Finally, an endline survey was administered in December 2012–February 2013. Data collection for the baseline and endline surveys was done by the Lubumbashi School of Public Health.

Due to implementation challenges during the start-up phase the sample of facilities participating in the impact evaluation changed between 2009 and 2010. As a result, only about 70% of health facilities that were included in the baseline survey were included in the pilot and endline survey. As a result only the endline survey was used for estimating the impacts of the program. The endline sample used for the main analysis included 96 health areas (123 health facilities) and 1,708 households. The average treatment effect (ATE) was estimated for each outcome of interest. The ATE measures the difference in mean outcomes between health facilities in the treatment group and health facilities in the comparison group.



TABLE 3: Implementation of the PBF pilot

	Planned Design		Actual Design	
	Treatment Group	Comparison Group	Treatment Group	Comparison Group
Volume of Payment		Same	No difference in volume of payment overall—although higher payments in comparison group during first six months, budget neutrality was respected for remainder of pilot	
Payment Mechanism	Payment based on the quantity of health services provided by the facility = Fee-for-service mechanism	Payment based on the size of the facility (the number of health workers on governmental payroll)	Payment based on the quantity of health services provided by the facility with a fixed total budget = Point system	As planned
Technical Verification	One visit every month by IRC/PARSS supervisors and the health zone administration—for verification of consistency between facility registers and declared data		As planned (although no visits were made when there was no payment)	
Community Verification	One verification every month of 30 patients per health area in a subsample of health areas. Sanctions on a case by case basis	No verification planned	Although 28 verifications were expected, only 6 were conducted. No adjustment proportional to discrepancies	2 community verifications for comparison purpose (no consequence on payments)
Autonomy	Facility manager could decide on payment allocation within the facility	Facility manager followed the governmental payroll which indicates who should receive what		As planned
Equipment investment support	Some equipment provided by PARSS			As planned
Drugs	Drugs for pharmacy stocking provided by PARSS		As planned	
Training	Health worker clinical training (by IRC)		As planned	
User fees	All health facilities in the pilot committed to reduce user fees		Overall reduction in user fee schedule despite substantial variation in level of reduction across facilities	
Baseline survey		Yes	Due to delay in project start-up and changes in health facility included in the pilot, The baseline sample did not match perfectly the pilot sample. Baseline survey used for balance checks but not for impact estimates	
Endline survey		Yes		As planned

RESULTS

Despite (i) several interruptions in the payment schedule during the two years and (ii) the comparison group receiving substantially higher payment levels during the first six months of the intervention (June–November 2010), over the two year period the two groups of health facilities received the same level of bonus payments (on average \$550 per month per facility). All other PARSS-related investments were also equally delivered across the two groups. In the end, the main differences between the two groups were (i) the financial incentives generated by the two different payment mechanisms, (ii) the level of community verification of results, and (iii) the level of financial autonomy facility managers had in using the payments to distribute financial bonuses to facility staff.

As payments were linked to the volume of services delivered, there was a natural risk of over-reporting by facilities in the treatment group. In order to curb this, payments were supposed to be contingent on rigorous and routine verification at the facility and community levels. In addition to a strong verification system, financial sanctions was supposed to be integrated into the process in order to reduce providers' incentives to submit fraudulent reports and over-reporting of phantom patients. In reality, the system suffered from several important weaknesses. First, community verifications were substantially fewer than planned—on average facilities only received 3 community verifications throughout the pilot (versus the 28 that were planned), with large variations across facilities in the number of verifications conducted (some did not receive any, others up to eight). Second, there was actually no 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 verified through community identification. For example, if 18% of patients were not found through community verification, the facility would only receive a reduction of 18% in their total payment and no additional sanctions were enforced.

The evaluation found that the propensity for facilities to over-report on service volumes decreased with the intensity of verification. In essence there were fewer discrepancies between declared and validated services as supervisors strengthened the attention paid to patient records. Moreover, no significant difference was found between the two groups in the proportion of patients that could not be found during the community verifications. However, the evaluation found that comparison group facilities under-reported services volumes and in certain cases treatment group facilities had high levels of over-reporting. Relative to the comparison group, health facilities in the treatment group had a strong financial incentive to keep complete records of their patients, resulting in improved accountability and availability of data as invalid patients were discounted from future payments. In contrast, the comparison group had no such direct incentive to maintain comprehensive patient records. Results from the evaluation show that a substantial proportion of health workers in the comparison group did not fill out consultation forms for each patient visiting the facility. As for the purpose of the impact evaluation completed consultation forms were used to validate patient volumes in both treatment and comparison groups, it is hence probable that service volumes were most likely underestimated in the comparison group. These opposite tendencies to over-report in the treatment group and under-report in the comparison group should be acknowledged when routine administrative data are used to compare the two

groups. These results highlight that when payments are linked to verified performance, health facilities strengthen mechanisms for data management.

In terms of financial consequences of the two payment mechanisms for the health facilities, it was found that the payment mechanism in the treatment group led to higher payment volatility at the facility level than the fixed payment mechanism in the comparison group. As payments based on the point system were adjusted relative to the performance of other facilities in the treatment group, the ambiguity of expected revenues may have led to feelings of insecurity and difficulties in financial planning among staff in the treatment group facilities. However, it was found that the flexibility provided to health facility managers in the treatment group with respect to bonus allocations among facility staff led to a more egalitarian distribution of payments, with a higher proportion of non-technical staff receiving bonuses in the treatment group than in the comparison group. Such findings underline the result that increasing management autonomy for health facilities can lead to greater transparency and equity in resource allocation among staff, where distribution criteria are based on individual performance and not administrative rules, hierarchy or grade.

The introduction of the financial incentives in the treatment group led to concrete changes in health workers behaviors and the service delivery models they adopted. Facility staff were found to be present at facilities more often in the treatment group than in the control group (reduction in absenteeism) and introduced different strategies to increase demand and utilization of health services by the population. One strategy adopted to increase the number of patients receiving services was to decrease the price of services, as it was found that fee schedules for targeted services were significantly lower in the treatment group than in the control group. A second strategy that facility staff in the treatment group used to increase service volumes was to offer more organized preventive health sessions at facilities in order to sensitize the population about the services offered by the facility and the health benefits associated with these

services. Third, facilities in the treatment group were found to conduct more community-based outreach activities, such as the provision of preventative and curative services at the village level through monthly visits to communities. Overall, the financial incentives of the payment mechanism in the treatment group led to an intensification of activities aimed at increasing demand and utilization of health services. Equally important, the evaluation found that introducing a financial incentive linked to service volumes did not have a negative effect on the quality of services, both for targeted and non-targeted health services (but quality did not increase either).

Surprisingly, the evaluation found that despite the introduction of different strategies to increase demand and access to health services by the local population, the efforts of the health workers in the treatment group did not lead to a significant increase in utilization or coverage of these services. It is crucial to think about the other barriers that impede the increase of the utilization of health services by the population. Given that the overall PARSS intervention included a package of interventions that included the reduction of user fees, the provision of drugs and equipment to facilities, infrastructure investment and rehabilitation, which were provided to both the treatment and comparison groups, it is possible that the additional incremental effect of either of the two payment methods had minimal to no effect on the population's health-seeking behavior. As all other components of the PARSS intervention were equal across the two groups and these interventions were introduced prior to the start of the evaluation and its baseline survey, the impact evaluation was not able to measure the effect of these other project interventions on health-seeking behavior or health status. Alternatively, the lack of response of the population might reflect other obstacles that were not addressed in the strategies implemented by the incentivized health facilities, such as geographic barriers, preference for alternative treatment options (traditional care, use of the informal sector, etc.), or general dissatisfaction with the quality of care provided by health facilities (results from the household survey found that quality of care was the most important factor driving demand and utilization of facility-based health services).

The evaluation identified several population-level health outcomes affected by the intervention. First, weight-for-age z-scores and height-for-age z-scores for children under five were found to be lower in the treatment group than the control group (0.17 and 0.10 standard deviations, respectively), suggesting that the intervention may have had a negative effect on health status of children in the targeted areas. Second, a positive effect was also found, as the number of work days missed was found to be 19% less in the treatment group than the control group.

As facilities in the treatment group reduced the prices of their services more than the comparison facilities without attracting more patients, the evaluation found lower user fee revenues in the treatment group than the comparison group. In fact, the treatment group had 42% less total resources for the facilities than comparison group, and health workers total income was 34% less in the treatment group than in the comparison group. This could have reduced the financial power of facilities to restock essential supplies and drugs, as suggested by the significantly lower quantity and quality of equipment in the treatment group as compared to the comparison group.

Finally, the financial incentive mechanism had a significant negative impact on job satisfaction of facility staff. Job satisfaction was 14% lower in treatment group than in control group, and the treatment group had 72% more facility heads concerned by the volatility of health facility revenues. The lower levels of job satisfaction found in the treatment group might be an effect of the limited response of the population to the newly introduced efforts to increase patient demand, which resulted in a reduction in facility and staff revenue (i.e. an increase in work led to a reduction in income). It is therefore not clear to what extent the negative effect of the financial incentive on job satisfaction is due to the exposure to the payment mechanism itself or to the difficulties induced by the overall PARSS pilot in the specific context of Haut-Katanga. It is also important to note that staff attendance was found lower in the treatment health facilities than in the comparison group facilities several months after the pilot ended, which suggests that

the incentive-based payments deterred some of staff intrinsic motivation.

The impact evaluation examined the effect of financial incentives to health care providers linked to performance on health service delivery results in the district of Haut-Katanga, DRC. The results from the impact evaluation can be summarized across four main conclusions:

1. Using financial incentives linked to performance outcomes can be an efficient and effective approach to improve the supply of priority health services;
2. Introduction of the financial incentive mechanism did not lead to a deterioration of the quality or availability of non-targeted health services;
3. The financial incentive payment mechanism led to an overall increase in the level of health worker motivation but may have contributed to reduction in intrinsic motivation after the intervention ended; and
4. The financial incentives linked to quantity did not lead to significant changes in the coverage or quality of services provided³.

These results suggest that certain strategies could be integrated into the intervention in order to increase demand and utilization of health services. Most importantly, payments should not only be linked to the quantity of services delivery, but also quality. As close to 80% of household survey respondents identified quality of care as the primary determinant of health-seeking behavior, incentives linked to quality outcomes may have led to improved quality and in turn improved demand and utilization. Additional strategies that could be applied to boost utilization include community-level awareness campaigns to inform target populations of the services offered by facilities and their benefits.

³ When compared with a lump sum payment mechanism that was not performance-related

TABLE 4: Summary of key results from the impact evaluation

Implementation of the payment mechanisms	
Understanding of the financial incentive payment mechanism	Facility staff understood that the payment mechanism was linked to the quantity of services provided, but only limited understanding of how payments were calculated
Implementation of the payment mechanisms	On average same level of payment in both groups Payment irregularity in both groups (5 months without payment out of 28) The treatment group had 26% higher payment volatility than the comparison group Greater variations in payment levels in the treatment group The autonomy associated with financial incentive system in the treatment group led to a more egalitarian distribution of the bonus payments among facility staff
Implementation of performance verification	Same frequency of technical verifications Same procedure for patient validation but different incentives to report accurate data that led to biased declarations Very few community verifications in the treatment group (3 verifications per health facility) No financial sanctions except payment adjustment for discrepancies
Strategies to increase the number of patients	The financial incentive mechanism induced more effort from the health workers to increase the number of patients for targeted services
Cost of health services	User fees for targeted services and the prices of drugs were 20% to 60% lower in the treatment group than in the comparison group The payment mechanism in the treatment group had no effect on the prices of non-targeted services
Availability of service delivery	The financial incentive mechanism led to no change in the number and type of health services offered There were 43% more preventive sessions for targeted services in the treatment group than in the comparison group No change in service delivery strategies for non-targeted services
Outreach activities	Treatment group facilities offered 30% to 50% more outreach activities for targeted services than comparison group facilities, without any change for non-targeted services
Impact on service utilization	The strategies used to increase the number of patients in the treatment group proved unsuccessful
Number of patients	There was no difference between the two groups in the number of patients visiting health facilities
Demand for health services	There was no difference in the demand for health services between the two groups
Impact on the quality of health services	The financial incentives in the treatment group did not affect the quality of health services
Patient perceived quality of care	No differences between the two groups in patients' perceived quality of care
Patient satisfaction	No differences in patient satisfaction found between the two groups
Impact on population health status	Mixed effects on population health outcomes Weight-for-age and height-for-age scores were significantly lower in the intervention group than the control group Number of days missed at work in the last month because of health issues lower in treatment group than the comparison group
Impact on health facilities' resources	The payment mechanism in the treatment group led to fewer financial resources for the facility and lower workers' payment
Total resources at the facility level	Health facilities total resources were 42% less in the treatment group than the comparison group
Health worker income	Health workers total income was 34% less in the treatment group than in the comparison group
Quality of facility infrastructure and equipment	The quality and the quantity of equipment was poorer in the treatment group than in the control group (of about 0.6 standard deviation)

(continued on next page)

TABLE 4: Summary of key results from the impact evaluation *(continued)*

Implementation of the payment mechanisms	
Impact on staff satisfaction and motivation	Perceived workload was found to be lower in the treatment group compared to the comparison group. Attachment to remuneration was found to be higher among health workers in the treatment group relative to those in the comparison group.
Staff satisfaction, stress and anxiety	Job satisfaction was 14% lower in treatment group than in control group The treatment group had 28% fewer health workers feeling they have too much work than the comparison group The treatment group had 72% more facility heads concerned by the volatility of health facility revenues
Staff motivation	The financial incentives had a positive impact on staff attendance during the program (14% more during spot checks) but a negative one at the time of the endline survey (25% less 4 months after payments stopped) The treatment group had 34% more health workers who attach importance to remuneration than the comparison group

LINKS BETWEEN PAYMENT SCHEME DESIGN CHARACTERISTICS AND IMPLEMENTATION OUTCOMES

Appropriate program design and implementation are vital for successful PBF pilots. Based on trial and error, PBF programs have evolved to include certain design and implementation characteristics for enhanced performance. These characteristics and their expected outcomes on health system performance are summarized in the forthcoming PBF Toolkit (Fritsche et al., 2014 (forthcoming)) and described in Table 5 (a detailed description of each element is found in Annex 1).

TABLE 5: PBF design and implementation characteristics linked to improved results

No.	Characteristic
1	Existence of a well-balanced benefit package (15–25 services) at health center and referral hospital
2	Rigorous ex ante and ex post results verification occurs ^a
3	Separation of functions among regulator, provider, and purchaser serves to improve accountability and credibility of results
4	Use of community client satisfaction surveys to gather information from clients on use and satisfaction
5	Use of a quantified quality checklist (balanced score card) with the result tied to payments
6	Use of a fee-for-service provider payment mechanism (makes measuring outputs easier and links efforts directly to rewards)
7	Strategic purchasing with a focus on underprovided and underutilized preventive services
8	Individual fees and total earnings that are significant and paid regularly
9	Most money to the most cost-effective services: two-thirds of performance payments go to the community or health center level and one-third to the first-level referral hospital
10	Inclusion of equity measures that orient payments to destitute geographical areas, facilities and the poor
11	Enhanced financial and managerial autonomy for health facility resource mobilization and use
12	Strengthened role of health facility management committee in decision-making processes
13	Quarterly or monthly payment cycles with facilities using indice tools for improved financial management
14	Performance frameworks for the regulator (health administration at the district and provincial levels)
15	Use of quality improvement and investment units to provide means for health facilities to upgrade quality
16	Health facility management instruments such as the business plan, indice tool, and individual monthly performance evaluation
17	Coaching and technical assistance by the purchasing agent
18	Existence of a district PBF steering committee for improved governance at the decentralized level
19	Use of a web-enabled application with public front end for improved data management, strategic purchasing and public accountability
20	Continuous coordination between technical assistance and the government
21	Capacity building at health facility, district, and national levels

^a Ex ante verification refers to verification that occurs prior to payment, while ex post verification refers to verification that occurs after payment.

TABLE 6: Potential links between PBF design elements and implementation outcomes

Advised design and implementation	Haut-Katanga Pilot, DRC	Possible effects of weak design and implementation	Observed Effects in the Haut-Katanga Pilot
Well-balanced benefit package at all levels	Only partial coverage at primary care centers (7 services) and referral health centers (3 services).	Potential for providers to focus on certain services to the detriment of others; lesser linkage and referrals between health center and hospital levels.	Effort concentrated on targeted services but no detrimental effects on non-targeted services. No difference in volume of referrals between health centers and referral health centers between the two groups.
Rigorous results verification	Regular register verifications but very limited community verification; instead of 28 (once per month); only 6. Sanctions for discrepancies were light. Counter-verifications implemented by the purchaser.	Opportunities for increases in phantom patients; incomplete or weak verification mechanisms can lead to lack of trust in results.	No significant difference in the number of phantom patients between the two groups. Most discrepancies correspond to patients who could not be found due to seasonal or permanent migration. The propensity to over-report patients remained concentrated in few health facilities.
Use of client satisfaction surveys	No client satisfaction surveys.	No feedback on perception of clients on services rendered.	Client satisfaction was not accounted for in payments.
Use of a quantified quality checklist with the result tied to payments	No use of the quantified quality checklist; no quality measures tied to payment.	Could potentially lead to an increase of quantity combined with a lesser increase of quality, no increase in quality, or even a decrease in quality.	No change in service quality (health workers' attitude, waiting time, consultation time, technical quality). Slight reduction in availability of drugs and equipment.
Use of a fee-for-service provider payment mechanism	Not real FFS; more of a weighting for relative value, and then apportioning available budget based on value obtained.	Reward for effort unclear, increasing productivity may lead to a devaluation of the value of each service delivered. May be interpreted by providers as being penalized for high achievement.	Even though the fact that more patients would lead to higher payment was very clear, the exact calculation of the reward was unclear.
Individual fees and total earnings that are significant and paid regularly	Irregular payments; total earnings remained the same during the intervention at \$550 per month per health facility on average.	Lower rewards provide fewer opportunities to invest in improving the quantity and quality of services.	Despite small reward, health workers increased efforts to attract more patients.

Most well-designed and well-implemented PBF programs from other contexts (such as Burundi and Rwanda) exhibit a mix of the characteristics listed in Table 5. The following section provides an assessment of the link between several key design characteristics of the Haut-Katanga pilot and potential effects of these elements on results observed. While not all outcomes were observed in the pilot or confirmed by its impact evaluation, the assessment builds on experiential knowledge from PBF

programs in various settings over the past decade. The PBF Toolkit (Chapter 17) was used to structure the assessment (Fritsche et al., 2014 (forthcoming)). It is important to note that PBF programs are continuously evolving on the basis of lessons learned which is why design and implementation characteristics are expected to evolve too. Table 6 provides examples of potential effects that may result when design and implementation characteristics are changed. A more detailed assessment is found in Annex 4.

Well-balanced benefit package at all levels

With only seven services purchased at the primary care level and three at the referral hospital level, one of the main weaknesses of the Haut-Katanga PBF design was the unbalanced benefit package. With less than 15 services in a benefit package and limited coverage at the referral center level, possible effects of the design constraint include a focus on certain services to the detriment of others and weak linkages between health center and hospital levels. While the impact evaluation found that there were no negative effects on non-targeted services (no change in the number and type of health services offered and no change in service delivery strategies for non-targeted services), it was found that health facilities in the treatment group conducted more awareness-raising and information activities for targeted services than facilities in the control group. One might assume that if the PBF intervention covered a broader array of services than the seven that were purchased, facilities might have conducted promotion activities for these services as well, leading to more comprehensive strategies to inform the population of essential services offered.

Rigorous results verification

While ex ante verification was well executed in the Haut-Katanga pilot, implementation of ex post verification remained limited⁴. Only 6 community verifications were implemented while 28 were expected in the intervention design (one per month during the time service provision was collected to compute payments, i.e. from June 2010 to September 2012). Moreover, performance verifications should have been done in all health areas, while only a sub-sample was effectively audited. The first verification took place in January 2011 (six months after the intervention started) in 12 health facilities. The second and third verifications occurred as expected in February and March 2011 (resp. in 24 and 45 health facilities). Then, a fourth verification took place in July 2011

in 50 health facilities. Those four verifications were conducted only in PBF health facilities. Only two other verifications occurred during the remaining duration of the pilot, in October 2011 (83 health facilities) and September 2012 (61 health facilities). On average 3 community verifications per facilities took place in the PBF group between June 2010 and December 2012. Low capacity of community-based organizations and difficult geography of Katanga Province, coupled with limited resources allocated to community verifications, may have contributed to the limited number of community verifications conducted during the pilot period.

Equally important, the sample of patients taken for community verifications was limited, with only 30 patients were to be checked per health area. The number is quite low relative to the number of patients facilities served, which was more than 800 per month on average. A supervision mission conducted in late 2010 found that for certain facilities in the PBF group over 30% of the patients found in facility registers were not found during community verifications⁵. For certain services like family planning, declared volumes of services delivered increased rapidly during the first year, despite the fact that there were clear contraceptive stock-outs at the provincial level. It was also identified that children were being vaccinated without consideration of their age or vaccination schedule.

Another weakness in the implementation of performance verifications was the lack of concrete penalties for fraudulent reporting. There was a threshold under which no penalty was applied (the threshold varied from 5% to 20% across the six verifications) to take into account the fact that some patient attrition was inevitable. Above the threshold, the payment to the

⁴ Two types of mechanisms exist for quantity verification in PBF schemes: those that are made before any PBF payment is made (ex-ante verification) and those that are undertaken after payments are made (ex-post verification). The latter includes community client satisfaction surveys and other types of counter-verification.

⁵ Evaluation of the Result Based Financing pilot project, The International Rescue Committee, Haut Katanga District, Katanga Province, DR Congo, December 20th, 2010

health facility was reduced by the same percentage as the percentage of inconsistencies. For instance, if the threshold was 10%, then a health facility with an inconsistency rate of 8% was not sanctioned, and a health facility with an inconsistency rate of 18% had an 18% reduction in its payment. All in all, there was no sanction at all because, if any penalty, the payment was just adjusted to what it would have been in the absence of over-reporting.

However, there was no significant difference between the two groups in the proportion of patients that could not be found during the community verifications (17% in the comparison group and 21% in the PBF group). This suggests that the propensity to make up phantom patients in order to inflate service volumes was limited in Haut-Katanga.

Use of client satisfaction surveys

In PBF designs, community client satisfaction surveys are used to gather information from clients on use and to gather their opinions on perceived quality of care and satisfaction with the services they received. Satisfaction scores are often reflected in the final payment to health facilities, with payments being reduced when satisfaction is low. In the Haut-Katanga pilot, there was no use of client satisfaction surveys. In addition, payments were not adjusted for satisfaction scores. As a result, no feedback was provided on perception of clients on services rendered.

Use of a quantified quality checklist with the result tied to payments

One of the key objectives of PBF is an improvement in the quality of care provided to clients. To incentivize health facilities to improve quality, quantified quality checklists (or balanced score cards) are used to independently assess quality, with payments tied to quality and adjusted for quality assessment scores. The impact evaluation of the Rwanda PBF scheme, which tied payments to quality assessment scores,

found significant increases in both the quantity and quality of services delivered (Basinga et al., 2011). In the Haut-Katanga pilot, a quantified quality checklist was not used and no quality measures were tied to payment. Potential consequences of not linking payments to quality include an increase in the quantity of services delivered without positive changes in the quality of services. As can be seen from the evaluation results, the PBF intervention did not lead to any significant changes in quality outcomes, including consultation time, clinical actions performed during consultations, and prescription patterns to patients⁶.

Use of a fee-for-service provider payment mechanism

Using a fee-for-service mechanism in PBF makes measuring outputs easier and links efforts directly to rewards. In the Haut-Katanga model, a true fee-for-service payment mechanism was not applied. In essence, the payment mechanism in the treatment group was a weighting for relative value, and then apportioning available budget based on value obtained. In essence, this means that an increase in productivity leads to a devaluation of the value of each service delivered. Potential consequences of this model include lack of clarity among health workers on the level of payment they would receive based on the level of their performance. However, the evaluation found that it was clear for the health workers in the PBF group that more patients would imply more money.

Individual fees and total earnings that are significant and paid regularly

As the evaluation aimed to ensure budget neutrality between the treatment and comparison groups, on

⁶ Note that the 2007 and 2011 surveys found that both quality and utilization did improve in the PARSS zones. The impact evaluation results conclude that the incentive-based payment mechanism had an insignificant contribution to these improvements.

average treatment group health facilities received the same payment level from PARSS. The Haut-Katanga impact evaluation found that one strategy that facilities in the PBF group used to attract more clients was a reduction in user fee prices. Although prices did decline more in the PBF group than in the comparison group, service volumes did not increase (relative to trends in the control group). Yet due to the reduction in user fee revenue, overall revenue was found to be less in the PBF group than the control group. In fact, the treatment group had 42% less total resources for the facilities than comparison group, and health workers total income was 34% less in the treatment group than in the comparison group.

The lower levels of total revenue in the treatment group could have affected both provider motivation and the financial power of facilities to restock essential supplies and drugs. For example, job satisfaction was 14% lower in treatment group than in control group, and the treatment group had 72% more facility heads concerned by the volatility of health facility revenues. As mentioned during in-depth interviews and focus group discussions, facilities in the treatment group felt that unit prices were too low to generate any real incentives. These negative outcomes highlight the importance of having payment mechanisms (whether they be user fees or PBF payments) that are significant and paid regularly, generating sufficient

income to (a) pay staff a significant monthly bonus income and to hire additional staff if necessary and (b) pay for non-salary recurrent cost items. In Haut-Katanga, the envelope allocated to both the control and the treatment groups was small (USD0.42 per capita) compared to PBF programs in other countries where output budgets average around \$3 per capita per year. Theoretically, this might have created weaker incentives for the staff to increase efforts as the payments were expected to be minimal. However, the results indicate that even with the relatively small incentives, health care workers did change their behaviors in an attempt to attract more patients.

Payments were interrupted in both groups from December 2011 to March 2012 due to (i) the political insecurity created by national election in November 2011 and (ii) internal administrative bottlenecks in the payment cycle. Payments due within this period of time were disbursed from April 2012 on, grouped with later payments. This problem of payment discontinuity was raised a lot in all of the qualitative interviews with facility heads and health workers, who noted that payment delays from PARSS led to demotivation of facility staff and district health teams. However, payment irregularity was the same in the treatment and in the comparison group and thus cannot explain any difference in staff motivation or anxiety between the two groups.



RECOMMENDATIONS

While the evidence base and learning agenda for Results-Based Financing continues to grow, over the past ten years, implementation experiences from several countries (including DRC, Rwanda, Burundi, Cameroon, Nigeria, Benin, Chad, Zambia, and Zimbabwe to name a few) underline several important characteristics of successful Performance-based financing schemes that should be taken into consideration when designing future PBF programs in the Democratic Republic of the Congo and abroad. While not every scheme has incorporated all of the recommendations provided below (for example Rwanda did not include geographical equity bonuses or investment units during the national scale-up⁷), experiences from each country do provide insight into characteristics that may contribute to improved results. These include:

1. Increase output budgets for operations in similar contexts to around \$3 per capita per year⁸;
2. Introduce the quality checklist at the health center and hospital level and make performance payments contingent on the quality checklist;
3. Introduce health facility management tools (business plans, indice tools, individual performance evaluations);
4. Introduce performance contracting for the health administration at all levels;
5. Strengthen community verification and counter-verification;
6. Strengthen the ex-ante verification and coaching;
7. Introduce a comprehensive web-enabled data management system (data analysis, reporting, payment);
8. Introduce geographic equity bonuses (hardship allowances); and
9. Introduce investment units to allow facilities to upgrade their quality quicker.
10. Strengthen governance through the introduction of decentralized PBF Steering Committees at the district level.

⁷ In fact, the geographical equity bonus was part and parcel of the Cordaid Cyangugu Province pilot since 2002 but was dropped in 2006 during the scaling-up as it did not pass into the national model. The innovation was then scaled-up in Burundi in 2010.

⁸ Estimating exact output budgets should depend on a variety of circumstantial factors, including the amount of available funding over the medium term. To assess sustainability, among other things, the per-capita cost of the funds for PBF should be compared against the amount of fiscal space for such a scheme over the longer run.

REFERENCES

- Basinga, P., Gertler, P.J., Binagwaho, A., Soucat, A.L., Sturdy, J., & Vermeersch, C.M. (2011). Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation. *Lancet*, 377, 1421–1428.
- Fritsche, G., Soeters, R., Meessen, B., Ndizeye, C., Bredenkamp, C., & Heteren, G. (2014 (forthcoming)). Performance-based Financing Toolkit. Washington, DC: The World Bank.
- Gertler, P., Giovagnoli, P., & Martinez, S. (2013). Rewarding Performance To Enable a Healthy Start: The Impact of Plan Nacer on Birth Outcomes of Babies Born into Poverty. Under Review for Publication: The World Bank.
- World Bank. (2013). World Development Indicators Online database.

ANNEX 1: PBF design and implementation characteristics linked to improved results

Characteristic	Detailed information
Well-balanced benefit package at all levels	A minimum of 15–25 services exist at each level: health center/community level and first-level referral hospital.
Rigorous results verification	A mix of ex ante verification and ex post verification occurs.
Separation of functions	Separation of functions among regulator, provider, and purchaser serves to improve accountability and credibility of results.
Use of community client satisfaction surveys to gather information from clients on use and to gather their opinions	Feedback is gained on use of services and opinion of the population
Use of a quantified quality checklist (balanced score card) with the result tied to payments	A comprehensive mix of measures on structure and process gives a balanced view on quality. The quality checklist is applied by the district or provincial health administration (regulatory function). Other results include observational and supervisory effects and improvement of technical efficiency.
Use of a fee-for-service provider payment mechanism	Using a fee-for-service mechanism is evidence based. It makes measuring outputs easier and links efforts directly to rewards.
Strategic purchasing with a focus on underprovided and underutilized preventive services	Fees are open at the micro-level (health facility), which leads to money following the effort, and budgets are closed at the macro-level, which leads to cost containment. Fees are adapted as a function of results (what is desired) and available budget (use of lever services—high-volume services such as curative services—to stay within budget at the macro-level). ICT solutions allow individual health facility fees to be managed on a quarterly basis.
Individual fees and total earnings that are significant and paid regularly	Income from PBF and other sources needs to be sufficient to (a) pay staff a significant monthly bonus income and to hire additional staff if necessary and (b) pay for non-salary recurrent cost items.
Most money to the most cost-effective services	Two-thirds of the money goes to the community or health center level and one-third to the first-level referral hospital. Improvement of allocation efficiency (reprogramming existing money to the frontlines) occurs.

(continued on next page)

ANNEX 1: PBF design and implementation characteristics linked to improved results *(continued)*

Characteristic	Detailed information
Equity	Various equity instruments exist: (a) delivering more money to destitute areas (ring-fenced global budget); (b) delivering more budget to destitute health facilities (higher fees); and (c) providing higher fees for services consumed by indigents.
Autonomy	Health facilities' decision rights include procuring their drugs and other inputs, having their own bank accounts, and deciding on their income. Hiring and firing of staff would be ideal.
Health facility management committee	The committee enhances local decision rights of health facilities combined with making the local population part of the oversight and governance mechanisms.
Payments and financial management	A quarterly payment cycle can still be combined with a monthly bonus payment to staff. The indice tool aids in managing all cash income in a holistic fashion and managing bonus payments.
Performance frameworks for the regulator	Health administration at the district and provincial levels and sometimes at the national level is made responsible for tasks that are under its control.
Quality improvement units and investment units	Negotiated through the business plan, the quality improvement and investment units provide means for a health facility to upgrade its quality.
Health facility management instruments	Instruments include the business plan, indice tool, and individual monthly performance evaluation.
Coaching and technical assistance	Usually occurring with the purchasing agent, coaching and technical assistance are vital.
District PBF steering committee	The committee furnishes governance at the decentralized level, links health system performance to the health administration, and provides a platform for government and the local community to discuss health system performance.
Web-enabled application with public front end	The application provides access to data at all levels, enables strategic purchasing, and enhances public accountability for performance.
Coordination	Coordination occurs between technical assistance and the government to support and enhance system performance.
Capacity building	System strengthening occurs at health facility, district, and national levels.

(Fritsche et al., 2014 (forthcoming))

Annex 2: In-depth impact evaluation results

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 treatment (financial incentive) group. As the treatment was randomly assigned across health areas, treatment effects are 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 first column of our result tables. The p-value for the test that this coefficient is equal to zero is presented in the second column of the result tables.

The unit of observation *i* varies depending on the study instrument used: health areas, health facilities, health workers, patients after consultations at health facilities, or households living in the study area. As randomization was done at the health area level and not the individual level and perfect balance was not achieved between the treatment and comparison groups, estimations control for a set of characteristics *X* that varies according to the unit of observation:

- **At the health area level:** dummy variables include the health zone, whether the majority of the health facilities in the area are private, whether the majority of the health facilities in the area are faith-based, whether the majority of the health facilities in the area are urban and whether the majority of the health facilities in the area are

health post (versus bigger structure as health centers and reference health centers).

- **At the health facility level:** dummy variables include the health zone, whether the health facility is private, whether the health facility is faith-based, whether the health facility is urban, whether the health facility is a health post, the catchment population size and distance to the closest supervision unit.
- **At the health worker level:** dummy variables include the health facility dummy variables, plus 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:** dummy variables include the health facility control variables, plus a dummy indicating that the patient is a female as well as the age of the patient and the reason for the visit.
- **At the household level:** dummy variables the health area control variables, plus a dummy indicating that the household was displaced due to the conflicts in the North of Haut-Katanga, as well as the sex and age of the household member. For women, it also includes a dummy indicating that the woman is literate.

Finally, error terms were clustered at the health area level to take into account potential correlation between units in the same assignment unit.

TABLE 1: Implementation of the PBF pilot

	Average Treatment Effect (ATE)	p-value (ATE=0)	Mean of Dep. Var. (Control)	St.dev. of Dep. Var. (Control)	Observations
1.1. Implementation of the Payments					
Average PARSS payment (Congolese Franc (FC)), Jun. 2010–Dec. 2012	-17489.36	0.768	500646.9	422386.6	149
Sd. Dev. of PARSS payment, Jun. 2010–Dec. 2012	82162.46	0.053*	321404.8	261933.6	149
<i>Distribution Scheme within the Facility</i>					
An apportionment basis is used to distribute the payment (from the Facility Head)	0.574	0***	0.031746	0.1767314	123
An apportionment basis is used to distribute the payment (from the Health Worker)	0.103	0.111	0.0955056	0.2947412	331
<i>Who gets the Payment</i>					
% workers who receive PARSS payment (source: Facility head)	0.163	0***	0.7718557	0.2236079	123
The health worker receives PARSS payment (source: Health workers)	0.014	0.833	0.7865168	0.4109218	331
<i>Distribution of the last PARSS Payment</i>					
Last PARSS payment per health worker (FC)	-10807.59	0.432	95849.72	73842.64	327
St. Dev. of the last PARSS Payment per worker within the facility	-16323.46	0.045**	44842	41468.93	85
1.2. Implementation of Performance Verification					
a. Technical verifications					
Number of technical verifications in the last 12 months	0.564	0.481	7.253968	4.931689	123
The health worker uses a consultation report for each patient	0.161	0.001***	0.7808989	0.4148041	331
The facility is sometimes visited by supervisors	0	0	1	0	123
Total number of visits by (any) supervisors in the last 12 months	-4.155	0.047**	19.07936	13.88466	123
b. Community verifications					
Number of community verifications (Jun. 2010–Sept. 2012)	1.811	0***	0.948718	0.2220001	154
Av. % missing patients	3.543	0.251	16.92081	20.02851	149

TABLE 2: Strategies to increase the number of patients

	Average Treatment Effect (ATE)	p-value (ATE=0)	Mean of Dep. Var. (Control)	St.dev. of Dep. Var. (Control)	Observations
2.1. Cost of Health Services					
a. User fees and cost of drugs					
<i>The Facility Head reports:</i>					
User fee for the first curative consultation	-692.45	0.281	1263.492	4557.316	123
User fee for the second curative consultation	-178.082	0.18	459.4828	799.0377	112
User fee for delivery	-224.185	0.655	2747.414	2423.25	113
User fee for the first prenatal visit	-407.873	0.095*	850	1741.42	118
User fee for the second prenatal visit	-80.801	0.053*	132.2034	264.8622	115
User fee for postnatal visit	-57.43	0.386	105.3571	430.8215	111
User fee for preschool consultation	-6.718	0.838	86.66666	154.8281	112
Fee Summary Index* at the facility level	-1.077	0.141	0.166473	4.212105	93
Fee Summary Index** at the facility level, targeted services	-0.807	0.061*	0.0366889	2.866472	109
Fee Summary Index*** at the facility level, non-targeted services	-0.398	0.346	0.1007338	2.064238	95
<i>Patients and Community Members report:</i>					
Fee paid for the delivery	301.24	0.762	9532.258	11570.85	773
Fee paid for the last postnatal visit	-71.637	0.35	400.8342	712.8497	392
Fee paid for the last prenatal visit	-112.969	0.125	665.5804	976.022	929
Fee paid for the last immunization shot	-22.096	0.237	87.71028	316.9161	2039
Cost of drugs purchased by the patient at the health facility (FC)	-1106.16	0.005***	2252.593	5166.591	980

*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.

** Idem but only with targeted services: first curative consultation, delivery, and prenatal visits

*** Idem but only with non-targeted services: second curative consultation, postnatal visit, and preschool consultation

(continued on next page)

TABLE 2: Strategies to increase the number of patients *(continued)*

	Average Treatment Effect (ATE)	p-value (ATE=0)	Mean of Dep. Var. (Control)	St.dev. of Dep. Var. (Control)	Observations
2.2. Availability of Service Delivery					
a. Health facility management					
<i>Opening</i>					
The facility was open the day of the survey	-0.053	0.465	0.9206349	0.2724789	123
The patient could consult each time s/he visited	-0.019	0.322	0.9375	0.2422843	993
The household member could consult each time s/he visited	0.016	0.351	0.857081	0.3500661	4323
Number of opening hours per week (as reported by the facility head)	-6.522	0.524	138.9262	47.86586	116
Number of opening days in the last month (as reported by the facility head)	-0.139	0.816	29.73016	1.885482	119
<i>Internal Management</i>					
Number of workers in the facility	-0.923	0.309	7.047619	5.692181	123
% health workers in the facility	0.027	0.425	0.683401	0.1826084	123
% doctors in the facility	-0.001	0.933	0.0271569	0.0575394	123
Number of workers who left the facility in the last 12 months	-0.009	0.972	0.7619048	1.011455	123
Number of staff meetings in the last 12 months (as reported by the facility head)	1.034	0.729	12.01587	19.59468	123
Number of staff trainings in the last 12 months (as reported by the facility head)	-1.175	0.159	3.904762	5.366305	123
The health worker received a training in the last 12 month	-0.03	0.653	0.409396	0.4933808	274
b. Scheduling of available services					
Number of services offered by the facility (between 0 and 23)	-0.492	0.35	13.55556	3.644606	123
Number of targeted services offered by the facility (between 0 and 10)	-0.141	0.606	7.730159	1.715267	123
Number of non-targeted health services offered by the facility (between 0 and 13)	0.143	0.723	5.825397	2.393133	123
Number of preventive services* scheduled by the facility (between 0 and 5)	-0.083	0.725	3.634921	1.440224	123
Number of targeted preventive services** scheduled by the facility (between 0 and 3)	0.146	0.248	2.523809	0.7799326	123
Number of non-targeted preventive services*** scheduled by the facility (between 0 and 2)	0.038	0.794	1.111111	0.8252294	123

(continued on next page)

TABLE 2: Strategies to increase the number of patients *(continued)*

	Average Treatment Effect (ATE)	p-value (ATE=0)	Mean of Dep. Var. (Control)	St.dev. of Dep. Var. (Control)	Observations
Number of scheduled preventive services* provided in the last 12 months	20.084	0.291	100.4426	82.87933	118
Number of scheduled targeted preventive services** provided in the last 12 months	31.542	0.044**	73.91803	57.09679	119
Number of scheduled non-targeted preventive services*** provided in the last 12 months	10.808	0.107	26.87097	31.89197	120
c. Number of outreach activities in the last 12 months					
<i>Performed at the Facility Level (source: Facility Head)</i>					
Total	13.121	0.453	53.53968	69.05206	120
Targeted Services*	12.65	0.361	39.92064	46.35219	121
Non-Targeted Services**	3.924	0.521	13.61905	26.10118	122
<i>Performed by the Health Worker (source: the Health Worker)</i>					
Total	7.184	0.171	15.23295	44.47532	326
Targeted Services*	5.976	0.096*	9.829545	26.42281	326
Non-Targeted Services**	1.208	0.523	5.403409	19.53698	326
<i>Performed in Patient' Community (source: Patients straight out of consultation)</i>					
Total	8.825	0.162	18.83629	35.8214	898
Targeted Services*	8.294	0.194	16.04734	30.52676	906
Non-Targeted Services**	0.171	0.851	2.890359	11.66756	942
d. Awareness among women in the community					
A health worker made the mother aware of postnatal visits	-0.028	0.577	0.5093167	0.5004315	953
The mother knows the schedule of postnatal care sessions	-0.009	0.821	0.4171779	0.4935977	960
A health worker made the pregnant woman aware of prenatal visits	-0.094	0.004***	0.8154122	0.3883109	1118
The pregnant woman knows the schedule of prenatal care sessions	-0.052	0.319	0.7053571	0.4562896	1121
A health worker made the pregnant woman aware of immunization	-0.019	0.688	0.6465201	0.4784884	1090
The pregnant woman knows the schedule of immunization sessions	-0.015	0.654	0.7403315	0.438857	1083
A health worker made the woman aware of family planning	-0.032	0.401	0.4351554	0.4960433	1842
The woman knows the schedule of family planning sessions	-0.02	0.6	0.351153	0.4775811	1874

*Preventive services include: immunization, pre- and postnatal care, family planning and HIV prevention

**Targeted preventive services include: immunization, prenatal care and family planning

TABLE 3: Service utilization

	Average Treatment Effect (ATE)	p-value (ATE=0)	Mean of Dep. Var. (Control)	St.dev. of Dep. Var. (Control)	Observations
3.1. Number of Patients					
<i>At the Facility Level, Last Month</i>					
Number of patients for targeted services	-61.714	0.628	605.6102	1194.306	112
Number of patients for all services	-49.916	0.732	832	1378.686	109
<i>At the Health Worker Level, Last Month</i>					
Number of patients for targeted services	-21.383	0.468	156.8494	176.6688	316
Number of patients for all services	-29.925	0.387	239.3313	245.167	309
3.2. Demand for Health Services					
General Demand					
The household member visited a health facility in the last 12 months*	-0.028	0.072*	0.4961274	0.5000388	9113
Have been sick in the last 12 month but did not visit a health facility	0.012	0.483	0.2500537	0.4330902	9124
Number of days the household member waited before visiting a health facility	0.014	0.957	3.643269	7.464664	3553
Child Immunization					
Ever had an immunization shot	-0.002	0.94	0.8486739	0.3585063	2448
Number of immunization shots based on the immunization card	-0.023	0.961	2.706977	3.186173	833
Has a scar from tuberculosis immunization	0.016	0.677	0.6	0.4900902	2441
Perinatal Care					
The mother delivered in a health facility	-0.015	0.684	0.8241309	0.3810987	961
The mother had a C-section	0.018	0.121	0.0173697	0.130807	773
Number of prenatal visits	-0.292	0.13	3.482944	2.243731	1117
Number of prenatal visits with a health worker at a health facility	-0.281	0.14	3.357782	2.122774	1120
Number of postnatal visits	0.058	0.655	1.10041	1.778309	957
Number of postnatal visits with a health worker at a health facility	0.055	0.622	0.8650306	1.426543	959
The mother is supplemented in iron	0.005	0.888	0.3875	0.487615	1121
The mother takes drugs to avoid malaria	-0.037	0.369	0.5392857	0.4988999	1121
Number of months the mother breast-fed her new-born	0.3	0.335	5.494845	3.787549	955
Family Planning					
The women is pro family planning	-0.044	0.132	0.4632353	0.4989086	1874
The partner is pro family planning	-0.022	0.443	0.316894	0.46551	1871
The women uses family planning	-0.068	0.01**	0.2044025	0.4034759	1878
The women uses a modern contraceptive method	0.005	0.69	0.0505263	0.2191437	1873

*We control for an additional variable: the household member was sick in the last 12 months



TABLE 4: Quality of services

	Average Treatment Effect (ATE)	p-value (ATE=0)	Mean of Dep. Var. (Control)	St.dev. of Dep. Var. (Control)	Observations
4.1. Technical Quality of Health Services					
<i>By the patient</i>					
Consultation time (minutes)	1.028	0.422	16.09263	15.51822	974
Compliance rate with medical procedure, any care service	-0.015	0.695	0.3538175	0.3248204	984
Drugs were prescribed to the patient and the patient was not examined	0.02	0.66	0.4077491	0.49187	991
<i>By the household member</i>					
Consultation time (minutes)	2.581	0***	16.98827	15.74057	4309
Number of days in the health facility after the delivery	-0.077	0.689	2.313283	1.702673	767
Compliance rate with medical procedure, prenatal care service	0.004	0.818	0.6657578	.1680248	923
Compliance rate with medical procedure, postnatal care service	0.048	0.123	0.6166667	0.258334	389
4.2. Patient's Understanding of Health Services					
The patient understands diagnosis and next steps	0.007	0.813	0.8268877	0.3786932	992
The patient knows what drugs to be taken	-0.072	0.039**	0.9042357	0.294539	991
The household member understands diagnosis	0.017	0.241	0.9372237	0.2426138	4258
4.3. Patients' Satisfaction					
<i>The Patient reports that s/he was...</i>					
Satisfied	0.013	0.359	0.9430147	0.2320279	994
Satisfied thanks to user fees	0.012	0.48	0.0277778	0.1644879	990
Satisfied thanks to care quality	0.003	0.937	0.5722222	0.4952152	990
Satisfied thanks to welcome quality	-0.027	0.442	0.2796296	0.4492334	990
Satisfied thanks to equipment quality	0	0.997	0.0333333	0.1796719	990
Dissatisfied thanks to user fees	0	0.	0	0	993
Dissatisfied thanks to care quality	-0.005	0.671	0.0349265	0.1837626	993
Dissatisfied thanks to welcome quality	0	0.946	0.0073529	0.0855121	993
Dissatisfied thanks to equipment quality	-0.006	0.359	0.0110294	0.1045364	993

TABLE 5: Population health status

	Average Treatment Effect (ATE)	p-value (ATE=0)	Mean of Dep. Var. (Control)	St.dev. of Dep. Var. (Control)	Observations
Weight-for-age z-score	-0.181	0.048**	-1.061262	1.741033	2428
Height-for-age z-score	-0.184	0.044**	-1.654884	1.864021	2402
Weight-for-height z-score	-0.347	0.306	0.4450215	5.516395	2403
Number of days missed at school in the last month because of health issues	-0.122	0.617	1.581796	3.318446	1561
Number of days missed at work in the last month because of health issues	-0.662	0.025**	3.547844	6.645518	2835
Number of household member that died in the last 12 months	0.007	0.732	0.1366313	0.4006933	1708
Number of women that died during or because of a pregnancy in the last 12 months	-0.004	0.427	0.009434	0.0967264	1707
Number of children under 5 that died in the last 12 month	0.012	0.55	0.0896226	0.3171387	1707
The child born in the last 12 month is still alive	-0.01	0.093*	0.9897751	0.1007032	961

TABLE 6: Health facility resources

	Average Treatment Effect (ATE)	p-value (ATE=0)	Mean of Dep. Var. (Control)	St.dev. of Dep. Var. (Control)	Observations
6.1. Total Resources at the Facility Level					
<i>Revenue of the Facility the last month as reported by the Facility Head</i>					
Revenue from user fees	-156138.6	0.148	310434.5	770580.8	120
Revenue from drugs and medical lab	-136695.9	0.083*	252311.2	494647.2	118
Revenue from PARSS payment	29812.86	0.669	116388.9	251311	123
Revenue from PARSS to cover the running costs	-13589.66	0.358	23044.76	112156.2	123
Revenue from other sources (NGOs and government)	-34445.51	0.249	31396.83	176646.7	123
Total	-306889.1	0.04**	738938.4	1267279	118
6.2. Health Worker Payments					
<i>Payment to health workers (reported by the Facility Head)</i>					
Average total payment per worker in the last month (Franc Congolese – FC)	-19252.79	0.079*	56168.16	71476.75	118
Average wage per worker in the last month (FC)	-1103.906	0.853	9439.635	49938.38	120
Average other payment per worker in the last month (FC)	-13211.64	0.049**	29590.41	39748.77	121
Average PARSS payment per worker in the last month (FC)	-1553.621	0.83	15444.8	31795.58	123
<i>Payment to health workers (reported by health workers)</i>					
Total payment in the last month (FC)	-35885.75	0.031**	127139.5	174494.9	282
Wage received in the last month (FC)	-4999.407	0.5	23654.04	88004.44	326
Other payments received in the last month (FC)	-28682.54	0.061*	102552.8	153866.8	285
6.3. Quality of the Facility Infrastructure and Equipment					
Quality Index based on interviewers' observation (Principal Component Analysis)	-0.525	0.014**	0.1990995	1.511479	116
Infrastructure index (Principal Component Analysis)	0.184	0.372	-0.171534	1.425423	110
Equipment index (Principal Component Analysis)	-0.639	0.026**	0.052816	2.226755	116
Number of types of vaccine currently available (between 0 and 5)	-0.744	0.034**	4.16129	1.738603	118
Number of types of vaccine that have been unavailable at some point in the last 12 months (between 0 and 5)	0.036	0.929	1.52381	1.740014	118
Number of types of drug currently available (between 0 and 9)	0.236	0.646	6.7	3.185241	117
Number of types of drug that have been missing once in the last 12 months (between 0 and 9)	-0.276	0.589	5.333333	3.445148	111

TABLE 7: Staff satisfaction and motivation

	Average Treatment Effect (ATE)	p-value (ATE=0)	Mean of Dep. Var. (Control)	St.dev. of Dep. Var. (Control)	Observations
7.1. Staff Satisfaction, Stress and Anxiety*					
a. Job satisfaction					
Level of satisfaction of the facility staff for his job (from 0 to 10)	-0.769	0.045**	5.705394	2.783944	455
The facility staff would go for a position in another facility	-0.031	0.564	0.7095436	0.4549178	455
The facility staff would go for a position in another facility for financial reasons	-0.089	0.155	0.6224067	0.485794	455
b. Subjective workload					
The facility staff finds his workload heavy	-0.086	0.093*	0.5291666	0.5001917	454
The facility staff reports too much work in the last 7 days	-0.169	0.002***	0.6092437	0.4889482	444
The facility staff felt tired due to the job in the last 7 days	-0.092	0.079*	0.5606695	0.4973471	445
c. Conflicts, stress and anxiety					
The facility staff worries about insecure / volatile remuneration	0.095	0.117	0.3886256	0.4885971	388
The facility staff worries about low remuneration	-0.057	0.256	0.4691943	0.5002369	388
The facility staff reports that the facility is in competition with other facilities	-0.007	0.898	0.3583333	0.4805129	454
Level of conflicts among workers perceived by the facility staff (from 0 to 10)	-0.155	0.521	1.717842	2.203041	453
The health worker reports that PARSS payment allocation is a source of conflict in the facility	0.129	0.172	0.1413043	0.3502439	165
7.2. Staff Motivation					
a. Staff effort					
<i>Attendance</i>					
The facility head supervises worker attendance	-0.007	0.918	0.8253968	0.3826766	122
The facility head ever used financial sanction to monitor worker attendance	0.147	0.031**	0.1111111	0.316794	123
The health worker would be financially sanctioned if s/he does not show up	0.072	0.245	0.3502825	0.4784117	330
Av. % workers present in the facility on unannounced visits 1, 2 and 3	0.074	0.067*	0.5807223	0.2924829	138
Number of workers in the facility the day of the survey	-1.354	0.032**	3.84127	3.418198	123

(continued on next page)

TABLE 7: Staff satisfaction and motivation (continued)

	Average Treatment Effect (ATE)	p-value (ATE=0)	Mean of Dep. Var. (Control)	St.dev. of Dep. Var. (Control)	Observations
Av. attendance rate of workers in the facility in the last 7 days (as reported by the facility head)	-0.09	0.155	0.7752835	0.1929815	123
Attendance rate in the facility in the last 7 days (as reported by the Health Worker)	-0.067	0.042**	0.7799358	0.1429585	331
<i>On-the-job effort</i>					
Number of health workers who were working when the interviewer arrived	-0.779	0.034**	2.15873	2.001663	123
<i>The Patients report:</i>					
Waiting time (minutes)	-0.083	0.975	14.6011	33.88586	994
Waited because of inactive staff	0.066	0.581	0.3150685	0.467758	137
Waited because of excessive demand	-0.066	0.581	0.6849315	0.467758	137
<i>The Household Members Who Visited in the Last 12 Months report:</i>					
Waiting time (minutes)	1.247	0.287	15.07084	19.60033	4317
Waited because of inactive staff	-0.041	0.471	0.2225806	0.4166515	595
Waited because of excessive demand	0.053	0.382	0.7677419	0.4229555	595
b. Importance attached to job remuneration					
The facility staff places much importance on job remuneration	0.044	0.399	0.3485477	0.4775021	455
The facility staff emphasizes financial benefits as the main advantage or disadvantage of his position	0.131	0.021**	0.3833333	0.4872145	454

* Staff includes facility heads and health workers.

ANNEX 3: Possible effects of weak PBF design and implementation

Advised design and implementation	Actual design and implementation	Possible effects of weak design and implementation
Well-balanced benefit package at all levels	Less than 15 services in a benefit package at both primary care centers and referral health centers.	Focus on certain services to the detriment of others; lesser linkage between health center and hospital levels.
Rigorous results verification	Ex ante verification not well executed and no ex post verification.	Increase in phantom patients; lack of trust in results. Very limited ex-ante verification; instead of 28* (once per month); only 6*. Sanctions for discrepancies were light. No counter verification. Only a sub-sample was audited and not the entire sample as planned with on a very limited sample of patients (30 while on average more than 1,200 per health area per month). 21% over-reporting found in PBF group. Most importantly sanctions for discrepancies were light with cut-off points between 5–20%, and only an adjustments for discrepancies. A non-significant deterrence while verifications (ex-ante and ex-post) very light and incomplete.
Separation of functions	Separation of functions not well executed; regulator and purchaser too close to provider.	Decrease of trust in reported results; decrease in sustainability because of lesser funding (both internal and donor funding).
Use of community client satisfaction surveys to gather information from clients on use and to gather their opinions	Community client satisfaction surveys not done.	Lack of trust in results; increase in phantom patients; no feedback on perception of clients on services rendered.
Use of a quantified quality checklist (balanced score card) with the result tied to payments	Simple quality measures consisting of single indicators or no quality measure used instead of a comprehensive quantified quality checklist.	Increase of quantity combined with a lesser increase of quality, no increase in quality, or even a decrease in quality.
Use of a fee-for-service provider payment mechanism	Percentage point coverage increase of select services purchased instead of a fee-for-service.	Narrow focus on certain services; problems with catchment population (denominator); unreliable baselines; penalties for high achievers; conflicts in assessing performance; long payment cycles.

^a Summing the 'fee-for-service' scheme was such that each unit of service that was provided was remunerated by a particular fee. Summing the number of each service delivered times the fee for this service produced a score for this facility. Each facility's score was then divided by the sum of all scores to produce the percentage of the total budget that should be allocated to each facility.

^b Rusa, L., W. Janssen, et al. (2009). "Performance-based financing for better quality of services in Rwandan health centers: 3-year experience." TMH 14: 830-837.

(continued on next page)

ANNEX 3: Possible effects of weak PBF design and implementation (continued)

Advised design and implementation	Actual design and implementation	Possible effects of weak design and implementation	Haut-Katanga Pilot, DRC
Strategic purchasing with a focus on underprovided and underutilized preventive services	Fees fixed for a prolonged period of time; no ability to analyze expenditures because of lack of appropriate ICT tools; focus on reimbursements for curative care.	No ability to renegotiate fees in case forecasts were mistaken; no ability to follow budget expenditure; focus on reimbursing curative care that leads to the under-provision of preventive services.	Rigid system; no strategic purchasing, no appropriate ICT tools, focus on reimbursement of curative care in combination with the introduction of a user fee reduction policy.
Individual fees and total earnings that are significant and paid regularly	Income from PBF and other sources insufficient to (a) pay staff a significant monthly bonus income and hire additional staff if necessary and (b) pay for non-salary recurrent cost items.	Small bonus payments insufficient to remedy staff coping mechanisms; insufficient funds for the purchase of drugs, medical consumables, equipment, and minor repairs, leading to lesser quantity and quality production.	Irregular payments; total earnings remained the same during the intervention at \$550 per month per health facility on average. A reduction in overall earnings and no incentives to work harder as bonus payments per individual were low. Total health facility income decreased significantly, and total salary decreased by 28%.
Most money to the most cost-effective services	Most money to hospital services.	Financing of less cost-effective services (hospital) to the detriment of more cost-effective services.	Reasonably well targeted services, however only a limited package and absence of comprehensive package at first level referral hospital.
Equity	Equity instruments not used.	Facilities in hard-to-reach areas will struggle to attract qualified staff and therefore to offer quality services; in case of user charges, higher barriers to access, to services for indigents than for the less poor.	No equity adjustments; no rural hardship adjustments; no fee exemptions for the poorest. Fees decreased across the board which is regressive (as not all are equally poor).
Autonomy	Very limited or no autonomy or money managed by higher levels of administration (none own bank account); no gain share (no bonuses paid).	Drugs frequently out of stock; staff less motivated; fewer innovations.	Autonomy limited to distribution of performance bonuses while use of resources from other categories (functioning, investment) remained rigid. Facilities obligated to continue separate budget lines for bonuses, functioning and investment. Facilities had to reduce user fees significantly which limited autonomy; no influence on hiring and firing of staff; no money to purchase drugs; drug supply through central push system.
Health facility management committee	No specific health facility management committee or no involvement in local governance of the health facility.	Lesser sense of ownership of community; fewer checks and balances.	Significant resources spent by IRC during the first year on community oversight mechanisms.
Payments and financial management	A six-month or annual payment cycle used; no indice tool used.	Lesser link between individual performance and overall achievement results; conflicts related to bonus payments; fragmented management of income.	Payments irregular; no use of indice tool; no use of individual performance evaluation framework.

(continued on next page)

ANNEX 3: Possible effects of weak PBF design and implementation (continued)

Advised design and implementation	Actual design and implementation	Possible effects of weak design and implementation	Haut-Kotanga Pilot, DRC
Performance frameworks for the regulator	No performance frameworks for the health administration.	Quantified quality checklist not timely carried out by health administration; data not complete, leading to difficulties in paying for performance of the health facilities; less supervision and training or coaching from the district and provincial health administration.	Weak performance framework for the regulator by PARSS.
Quality improvement units and investment units	No quality improvement units and investment units used.	No improving of aspects of structural quality such as lack of equipment; certain minor infrastructural repairs to be slower or not done due to financial constraints.	No investment unit/quality improvement units used.
Health facility management instruments	No business plan, no indice tool, and no individual monthly performance evaluation used.	No ability for the purchaser to negotiate certain targets; more difficult to intercept moral hazard of the provider; difficulties managing cash income in a holistic manner; difficulties in distributing performance bonuses; staff conflicts	No use of business plans or indice tools, no individual performance evaluation used.
Coaching and technical assistance	No coaching of health facility management provided; no or very limited technical assistance provided to the health facilities and district health administration.	At the health facility level, less performance because of less advanced strategies; at the district level, less capacity development related to analyzing performance and less ability to support enhancing performance of health facilities.	Coaching was part of the PARSS supervisor/verifier terms of reference, however: no or very limited technical assistance provided to the health facilities and district health administration.
District PBF steering committee	No district PBF steering committee.	Less ownership of government of the PBF system; no leveraging of health administrative capacity; less input from the local community in governance of public health system.	No district PBF steering committee.
Web-enabled application with public front end	Fixed database or Microsoft Excel-based management tool.	No public access to data or financial information; much less availability of data for action.	No web-enabled application used; data drawn from separate forms—separate from HMIS; unreliable information and over-reporting.
Coordination	Poor coordination or no coordination between government and technical assistance agencies.	Less availability of technical assistance; more fragmentation of health system than could be the case; less support of development partners than could be the case.	Poor coordination or no coordination between government and technical assistance agencies. From July 2010 there was no technical agency any more. The PARSS set up a separate project management unit with limited technical capacity.
Capacity building	Very little or no capacity building.	Less quality and quantity performance results than could be the case.	Very little or no capacity building.

