2

Impact of performance-based financing on health-care quality and utilization in urban areas of Cameroon

Omer Zang,[i] Sebastien Djienouassi,[i] Gaston Sorgho,[i] Jean Claude Taptue[ii]
Corresponding author: Omer Zang, e-mail: zang_omer@yahoo.fr

Performance-based financing has attracted considerable interest from governments and aid agencies in low-income countries as a means to increase productivity and quality of health-care providers. Supply-side PBF is an instrument that links financing to pre-determined results, with payment made upon verification that the agreed results have actually been delivered by the health facility. In Africa alone, more than 35 countries, including Cameroon, are implementing or are in the process of introducing payment methods that reward performance.[1] Many impact studies, with varying degrees of rigour, have been or are being carried out in various settings on PBF and other similar financial incentives aimed at health workers. Randomized experiments were carried out in order to monitor health worker attendance in India and incentivized service quality by physicians in the Philippines. Results showed that in India, the monitoring system was initially extremely effective but became ineffective after 18 months due to administration laxity. In the Philippines, service quality-based incentives had significant effects.[2,3]

In Africa, to date, only two experimental studies of the impact of PBF on health service provision and utilization have been completed, in Rwanda and the

SUMMARY—This article looks at a pilot project designed to estimate the impact of performance-based financing (PBF) on the quality and utilization of health care in a predominantly urban setting – the Littoral region of Cameroon. It uses three quasi-experimental impact evaluation methods involving matching and difference-in-difference. Results show that the PBF pilot had a positive and significant impact on most essential aspects of quality of care. Meanwhile, there was no impact on any of the indicators of health service utilization with the exception (limited) of modern contraceptive methods. These findings suggest that the setting and indicators chosen are important in achieving maximum impact. However, it should also be noted that improvements in utilization might be limited as a result of high baseline figures. Finally, the findings show that the quality of care seems to be the most promising aspect in terms of improvements related to PBF in urban settings.

Voir page 69 pour le résumé en version française.
Ver a página 69 para o sumário em versão portuguese.

i World Bank
ii Ministry of Health, Cameroon

Democratic Republic of the Congo. In Rwanda, PBF proved an efficient way to increase health service quality and utilization, resulting in improved child health outcomes.[4,5] In the Democratic Republic of the Congo, Elise Huillery et al (2013) found that financial incentives improved effort from health workers to increase targeted service provision, but demand for health services was not responsive to these incentives.[6] Most other studies using non-credible comparison groups or comparing simple before and after situations advocate PBF as a way to increase accountability, efficiency, quality and quantity of service delivery.[7–14] Loevinsohn and Harding (2005) reviewed ten studies on the effect of contracting with non-state entities, including non-governmental organizations (NGOs), as a way to improve health-care delivery, and concluded that contracting for the delivery of primary care can be very effective and that improvements can be rapid.[15]

Given the rising popularity of this financing strategy, robust evidence about its effects is still needed.[16,17] One way of improving the robustness of the evidence might be, when an experimental design was not prepared, the use of multiple quasi-experimental methods to assess the impact.[18]

Cameroon has made little progress towards achieving the Millennium Development Goals (MDGs). In fact, with a few exceptions such as immunization, most key indicators of maternal and child health and nutrition have stagnated or worsened since 1990. The mortality rate in under-five-year-olds rose in Cameroon in the 1990s and has stagnated in the 2000s. Maternal mortality has remained high and even worsened since 1998.[19] Analysis of the health system of Cameroon indicates that linking performance to results could indeed make a difference. The PBF programme in Cameroon also includes enhanced supervision as well as financial and hiring autonomy on PBF subsidies.[20] It is clear in Cameroon, suboptimal allocation of resources and resource use inefficiencies are key underlying determinants of the limited improvements achieved in the health sector. An important part of the problem is that the operational level receives a small fraction of the health budget while the lion's share of resources is allocated for administration.[20] Inefficiencies are also created by the inadequate alignment between the burden of disease in Cameroon and health expenditures. Governance problems are at the root of the second key constraint to district health system functioning in Cameroon.[20] Furthermore, non-transparent human resource management practices combined with low salary levels drive health workers to abuse public funds by charging informal payments or over-billing patients for services and, ultimately, deter use by the poor.[20,21] Finally, cost recovery mechanisms are extensively used in Cameroon, and are the greatest source of revenue for health facilities.[7,21,22]

The objective of the study reported in this article is to estimate the impact of a PBF pilot project on quality and utilization of health care in an urban setting and cross-check the results with various quasi-experimental impact evaluation methods.

Method

Within the framework of the quasi-experimental impact evaluation study of Cameroon's PBF project in the Littoral region, a baseline survey was conducted in January 2011 and follow-up survey in February 2013. Three quasi-experimental impact evaluation methods were applied for this study.

Study area

Cameroon is a central African country with a population of almost 22 million as estimated in 2014. Cameroon's health system is organized into 10 regions, 189 health districts and around 1 700 health areas. As assessed in 2012 by its Ministry of Health, the country has 1 888 public integrated health centres, 760 private health centres, 155 sub-divisional health centres and 164 district hospitals. Overall, the health system of Cameroon is staffed with 1 842 medical doctors, 18 954 nurses and 1 340 pharmaceutical personnel.[20] In 2011, per capita total health expenditure was US$ 61. Private spending (out-of-pocket) accounted for 51.4% of total expenditure on health, while public spending accounted for 33% and external resources for 14.4%. The share of private pre-payment is very low at 1.2% of total health expenditure. Government expenditure on health is low (US$ 328.2 million), amounting to 6.2% of total government expenditure in 2011.[20]

The 2011 Demographic and Health Survey – Multiple Indicator Cluster Survey (DHS-MICS) survey estimated that 53% of children aged 12–23 months were fully immunized. Despite a high level of awareness, 23% of women aged 15 to 49 years who are in a relationship use some form of contraception. In the same year, 85% of pregnant women consulted a skilled health care provider during their pregnancy. Moreover, close to 40% of women were still giving birth at home in Cameroon. For all these indicators, urban/rural and interregional disparities are very large.[19,20]

Littoral is the most developed region of Cameroon and hosts the country's economic capital. With a population of 2.8 million it is the third most peopled region of the country. The region accounts for 3 701 health personnel; almost 18% of national health human resources for 12% of the country's population. The health map of the region, made up of 19 health districts, 163 health areas and around 310 primary and secondary level health facilities, suggests that, overall, the setting of the PBF pilot project is mostly urban. The intervention targets four treatment health districts, home to 636 000 inhabitants.[22]

The same sampling design was applied for both the baseline and endline surveys, hereafter designated as PBF-LT surveys. The four health districts of the treated area are: Cité des Palmiers, Edea, Loum and Yabassi. The six control health districts are: Nylon, Mbanga, Melong, Logbaba, Nkongsamba and Manjo. For the purpose of delimiting comparable health zones for the project and the evaluation, these health districts were disaggregated into 152 health zones among which 40 were randomly sampled for the baseline survey and then revisited in the endline survey. In the rest of this article, they will be designated as delimited health zones. A total of 52 private and public health facilities were numbered in those delimited health zones and the household survey randomly targeted 1 000 households; 25 per delimited health zone.

Power calculations were conducted assuming antenatal care and full immunization coverage as results of interest to assess the validity of the overall sample size. They concluded that for a level of power of 80–90%, a sample size of 1 000 households was enough to allow any impact on utilization to be detected. Power calculations were not conducted for health facilities sampling.

Table 1 shows the sample sizes of baseline and endline surveys. Among 1 000 households surveyed, 62.5% were located in urban areas. Out of the 40 sampled delimited health zones investigated, 25 were urban.

Three questionnaires were designed for the PBF-LT surveys and served as guidelines for households and main health facilities of delimited health zones data collection. The household questionnaire collected data on the household composition; some characteristics of under-five-year-old children; some household characteristics – including their assets, income and expenses, sickness episodes; under-one-year-old children immunization; characteristics of the latest pregnancy of the household; and attitudes of women of reproductive age to contraception. The health facility questionnaire focused on health facility identification; catchment area population size; facility expenses and income; and personnel payment. A health facility quality checklist served to collect quality scores on the following components: structural quality; outpatient care; maternity; family planning; vaccination and antenatal care protocols; laboratory; and drug and supply availability (safety stock measured by the monthly average consumption). Each component of the quality checklist had a maximum score ranging from 3 to 12 points and included a series of items that "quality verifiers" should observe during the survey. The overall quality score summed up all component scores included in the checklist for a given facility. The maximum value was 68 points (see Table 3).

Cameroon's National Committee on Ethics and Human Health provided ethical clearance for all the surveys related to the process of impact evaluation of the PFB on health in Cameroon.

Impact evaluation quasi-experimental methods

The study used the propensity score matching method technique (PSM), the double difference or difference-in-difference method (DD) and a mix of these two methods as a third impact evaluation approach.

The basic idea behind propensity score matching is to match each participant with an identical nonparticipant and then measure the average difference in the outcome variable between participants and nonparticipants.[18,23] The balancing property test is captured by the area of common support. It represents the propensity scores within the range of the lowest and highest estimated values in the treatment group. With the propensity scores generated, the outcomes of interest between the treatment group and the matched control group are then compared to see whether the intervention affects the outcome of interest. This is possible by estimating the average treatment effect on the treated (ATT) of the programme participation, using kernel-based matching – identified as the most robust method. The method is usually accompanied with bootstrapping of standard errors.[24,25] The matching method is meant to reduce bias by choosing the treatment and comparison groups on the basis of observable characteristics.[26]

The double difference method uses panel data, collected from a baseline survey before the programme was implemented and after the programme has been operating for some time. The DD method can be implemented using a regression on panel or pseudo-panel data:[28]

$$Y_{it} = a + bTt + \beta T_i + \delta_{it} + \varepsilon_{it}$$

where T is the treatment variable, t is the time dummy, and b is the coefficient of the interaction of T and t. b gives the estimate of the impact of the treatment on the outcome Y.

The health outcomes the intervention targets considered in the impact modelling are: overall quality score and its main components; human resources availability (physicians, nurses, nurses' aides and unqualified staff); and health service coverage (outpatient consultation, under-one-year-old children vaccinated on time, unwanted pregnancy, modern contraceptive method utilization, institutional delivery, two or more antenatal care visits, and pregnant women antitetanic vaccination).

All coefficients of the model are expected to be positive as the assumption behind the PBF intervention is that it should increase either the health-care quality score, staff availability or health service coverage.

For this study, in the case of health facilities, panel data were used; and in the case of households, pseudo-panel data were used as only the same delimited health zones were considered for the baseline and endline surveys – but not the same households. The DD method is popular in non-experimental evaluations. A basic assumption behind the simple implementation of the DD method is that other covariates do not change over a few years.[28,29]

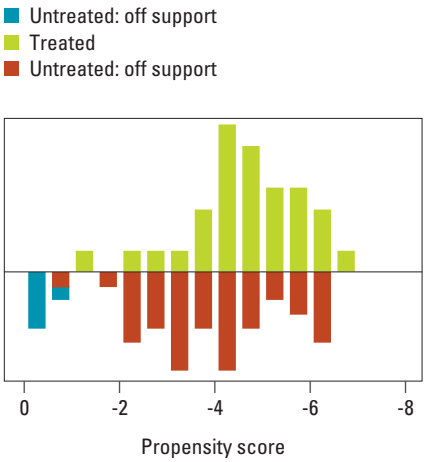
The DD method can be refined in a number of ways. One way is by using PSM with the baseline data to make sure that the comparison group is similar to the treatment group and then applying double differences to the matched sample (DD-PSM). This way, the observable heterogeneity in the initial conditions can be dealt with.[30]

Table 1. Sample sizes of baseline and endline surveys

	Households			Health facilities		
	Rural	Urban	Total	Rural	Urban	Total
Treated	150	250	400	6	10	16
Control	225	375	600	9	15	24
Total	375	625	1000	15	25	40

Source: PBF-LT surveys

Figure 1. Common support derived from double difference



Source: PBF-LT surveys

Using the min-max method, the distribution of the PSM density in the two groups was graphed in order to portray the common support area (see Figure 1). As a result, the bulk of the delimited health zones in the project area could find a matching unit in the control group under the matching criteria used: catchment area population size, square of catchment area population size, number of qualified health personnel, square of number of qualified health personnel and number of qualified health personnel to catchment area population size.

Another way to check the quality of the matching is to test the equality of some observables between the treated and the control groups under common support. Table 2 shows no significant difference for the selected observables between the treated and the untreated groups.[30,31]

Limitations of the study

Some methodological limits can be observed in the sampling strategy used for the data collected for this study. Power calculations were not made to figure out beforehand whether the sample size was large enough to be able to capture any impact. Only post-sampling power calculations were made with the constraint of 1 000 household sampling size and concluded favourably. Nevertheless, the data used still remain potentially more robust than many studies of the same type using routine data and smaller sample sizes.

Results

The findings relate to the quality of care at the health facility, including human resource aspects, and curative, maternal and children health service utilization. All results are presented in Table 3.

Quality of care

- The three methods have very similar differences on overall average quality score impact (DD: 20.2/68, $p<0.01$; ATT: 22.4, $p<0.01$; and DD-PSM: 19.2, $p<0.01$).
- There is a positive impact score of about 5/12 points on the institutional quality of care (DD: 5.3, $p<0.01$; ATT: 4.9, $p<0.01$; and DD-PSM: 5.1, $p<0.01$).
- The impact score on outpatient care quality is positively significant and close to 3/11 points (DD: 2.9, $p<0.01$; ATT: 2.7, $p<0.01$; and DD-PSM: 2.6, $p<0.01$).
- The maternity quality score could also register a significant and positive impact score ranging from 3.6/12 to 5.3/12 points across the methods (DD: 3.6, $p<0.01$; ATT: 5.3, $p<0.01$; and DD-PSM: 3.7, $p<0.01$).
- On family planning quality score, a positive and significant impact is recorded with a magnitude of about 2.5/8 points (DD: 2.5, $p<0.01$; ATT: 2.6, $p<0.05$; and DD-PSM: 2.5, $p<0.01$).
- Vaccination and antennal care quality are measured under the same component in the quality checklist and record a 2/7-point score impact (DD: 2.0, $p<0.05$; ATT: 2.0, $p<0.05$; and DD-PSM: 1.9, $p<0.05$).
- Among the three methods, laboratory quality impact score is positive and significant only with difference-in-difference with a 1/5-point impact difference at 0.1 margin of error (DD: 1.021, $p<0.1$).
- On drug availability, only simple difference method could be positive and significant with a 3/10-point impact difference at 0.1 margin of error (ATT: 3.0, $p<0.1$).
- Impact on pharmaceutical supply availability score (1/3 point score) is positive and significant in all three methods with various levels of significance (DD: 1.0, $p<0.05$; ATT: 1.0, $p<0.1$; and DD-PSM: 1.0, $p<0.1$).
- Human resource is also a key aspect of the quality of care at the health facility. All three methods show a non-significant coefficient on physician presence at the health centre. No significant impact is either found on other human resource availability: nurses, nurses' aides and unqualified staff.

Health service utilization

All health service utilization indicators' coefficients, measured by coverage percentages, are not significant with the exception of modern contraceptive methods (0.085 coverage rate-difference) for which only the kernel PSM method yields a significant coefficient (ATT: 0.085, $p<0.1$).

Discussion

It is considered that PBF could impact on the quality of health-care service provided by contracted facilities through a positive influence on quality of most of the components as assessed in the study. This may be driven by improved staff motivation through financial bonuses as shown by an experimental evaluation of financial incentives on staff in the Philippines.[3] Meanwhile, laboratory and drug availability showed mitigated levels of impact significance. No significant impact either was found on health personnel availability. Moreover, all health

Table 2. Post matching mean comparison tests

Variable	Treated	Control	% bias	T	p>t
Square catchment area population size	2.1e+08	2.0e+08	2.9	0.14	0.888
Square number of qualified staff	131.34	129.81	0.2	-1.18	0.244
Population x qualified staff	1.5e+05	1.5e+05	2.2	-0.86	0.395
Catchment area population size	13318	12912	5.6	-0.01	0.989
Number of qualified staff	10.219	10.259	-0.4	-1.57	0.120

Source: PBF-LT surveys

Table 3. Impact on quality, human resource and service utilization

	DD	Kernel PSM (ATT)			DD with kernel PSM (DD-PSM)
	Diff-in-diff	Treated	Control	Difference	Diff-in-diff
Impact on the quality of care assessed at the health facility (score)					
Overall quality score (/68)	*20.229****	*53.440*	*30.950*	*22.490****	*19.252****
Institutional quality (/12)	5.354***	8.438	3.555	4.882***	5.082***
Outpatient care (/11)	2.958***	9.938	7.221	2.716***	2.575***
Maternity (/12)	3.625***	9.688	4.374	5.314***	3.723***
Family planning (/8)	2.500***	4.375	1.759	2.616**	2.509***
Vaccination and antenatal care (/7)	1.979**	5.313	3.352	1.961**	1.914**
Laboratory (/5)	1.021*	4.688	3.769	0.921	0.688
Drug availability (/10)	1.792	8.188	5.136	3.052*	1.743
Supply availability (/3)	1.000**	2.813	1.784	1.029*	1.017*
Impact on human resources in the health facility (number of staff per facility)					
Physicians	0.771	1.688	0.640	1.047	0.756
Nurses	1.896	5.063	2.105	2.957	1.926
Nurses' aides	2.625	10.500	6.441	4.058	2.794
Unqualified staff	1.500	6.063	3.206	2.856	1.657
Impact health services utilization (coverage rates)					
Outpatient consultation	-0.020	0.726	0.759	-0.033	-0.034
Under-one-year-old children vaccinated on time	-0.036	0.803	0.833	-0.030	-0.036
Unwanted pregnancy	0.103	0.111	0.058	0.053	0.097
Modern contraceptive method	0.036	0.373	0.288	0.085*	0.036
Institutional delivery	-0.011	0.981	1.000	-0.019	-0.004
Two or more antenatal care	-0.021	0.953	0.968	-0.015	-0.011
Pregnant women antitetanic vaccination	-0.022	0.867	0.877	-0.098	-0.025
Drug availability (/10)	1.792	8.188	5.136	3.052*	1.743

Source: PBF-LT surveys
Note: *P<0.1; **P<0.05; ***P<0.01

service utilization, measured by coverage rates, was not significantly impacted by the intervention, with the exception of modern contraceptive methods. Actually, it should be noted that, with the exception of contraception, coverage at baseline of all other health services targeted in this study were already very high at the outset, leaving little room for improvement. Comparable results on service utilization were found in the Democratic Republic of the Congo and, to some extent, in Rwanda.[4,6]

Conclusion

The pilot study indicates that the implementation of PBF in an urban area of Cameroon could significantly and positively impact on key aspects of clinical care quality without really leveraging more utilization of health services. These findings demonstrate that within the framework of health PBF, the context (urban/rural) and the list of indicators matter as underpinning factors of future impact. In urban settings, the quality of care seems to be the most likely area for improvement as there may be little room to improve health service utilization in many settings.

Acknowledgements

We wish to acknowledge the support of the World Bank. Meanwhile, all errors remain ours, and the opinions expressed in this paper are ours alone and should not be attributed to the institutions with which we are associated for this work. We are grateful to André Arsène Bita Fouda who played a significant role as the regional delegate for public health for the project's implementation. We thank Celestin Kimanuka for coordinating the data collection process along with Simon Mbunya. And special thanks also to Enandjoum Bwanga, the national coordinator of the Health Sector Support Investment Project; and to Paul Jacob Robyn, the coordinator of the main experimental impact evaluation of the project into the three other regions financed by the World Bank. Finally, we gratefully acknowledge the staff and patients of the health facilities as well as the households for the time and information they provided.

References

1. Bonfrer I et al. The effects of performance-based financing on the use and quality of health care in Burundi: an impact evaluation. *The Lancet* 2013; 381:S19.
2. Banerjee A et al. Putting a band aid on a corpse: Incentives for nurses in the Indian public health care system. *J European Econ Assoc* 2008; 6(2–3):487 –500.
3. Peabody JW et al. Financial Incentives and Measurement Improved Physicians' Quality of Care in the Philippines. *Health Affairs* 2011; 10(4):773–81.
4. Basinga P et al. Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation. *The Lancet* 2011; 377:1421–28.
5. Gertler PJ et al. *L'évaluation d'impact en pratique*. World Bank 2011.
6. Huillery E, Seban J. *Performance-Based Financing for Health: Experimental Evidence from the Democratic Republic of Congo*. ENS-France Working paper 2013.
7. Soeters R et al. Performance Based Financing Experiment Improved Health Care in the Democratic Republic of Congo. *Health Affairs* 2011; 30(:8):1518–1527.
8. Rusa L et al. Performance-based financing for better quality of services in Rwandan health centres: 3-year experience. *Trop Med & Int Health* 2009; 14(7):830–837.
9. Soeters R et al. *Comparison of two output based schemes in Butare and Cyangugu provinces with two control provinces in Rwanda*. Rapport pour le GPOBA, WB-Ministry of Health, Rwanda 2005.
10. Eichler R et al. *Performance-Based Incentives for Health: Six Years of Results from supply-side programs in Haiti*. Working paper 121. Washington, DC: Center for Global Development 2007.
11. Soeters R, Griffths F. Improving government health services through contract management: a case from Cambodia. *Health Policy Plan* 2003; 18(1):74–83.
12. Forsberg E et al. Financial incentives in health care. The impact of performance-based reimbursement. *Health Policy* 2001; 58(30:243–262.
13. Sondorp E et al. Afghanistan: *Paying NGOs for performance*. Washington, DC: Center for Global Development 2008.
14. Meessen B et al. Output-based payment to boost staff productivity in public health centres: Contracting in Kabutare district, Rwanda. *Bull World Health Organ* 2007; 85(2):108–15.
15. Loevinsohn B, Harding A. Buying results? Contracting for health service delivery in developing countries. *The Lancet* 2005; 366:676–681.
16. Eldridge C, Palmer N. Performance-based payment: some reflections on the discourse evidence and unanswered questions. *Health Policy Plan* 2009; 24:160–166.
17. Oxman AD, Fretheim A. Can paying for results help to achieve the Millennium Development Goals? A critical review of selected evaluations of results-based financing. *J Evid Based Med* 2009.
18. 18. Khandker SR. *Handbook on Impact Evaluation*. World Bank 2010; 262.
19. Institut National de la Statistique – Cameroun. *EDS-MICS 2011*. INS 2012.
20. World Bank. *Country Status Report on Health*. World Bank 2013.
21. Renaud A, Ndema EA. *Redynamisation des Soins de Santé à l'Est du Cameroun – Phase II : Rapport de la mission d'évaluation finale*. Cordaid 2010.
22. World Bank. *Concept note for the impact evaluation of a PBF pilot in Cameroon*. World Bank 2012.
23. Ho et al. Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis* 2007; 15:199–236.
24. Rosenbaum P, Rubin D. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 1983; 70:41–50.
25. Rubin D. Using multivariate matched sampling and regression adjustment to control bias in observational studies. *J Am Stat Assoc* 1979; 74:318–28.
26. Heckman JH et al. Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *Rev Econ Studies* 1997; 64:605–654.
27. Heckman JH et al. Characterizing Selection Bias Using Experimental Data. *Econometrica* 1998; 66:1017–1098.
28. Imbens G, Wooldridge J. Recent Developments in the Econometrics of Program Evaluation. *J Econometric Lit* 2009; 47(1):5–86.
29. Imbens G, Wooldridge J. *What's New in Econmetrics?* Lecture notes 10. NBER Summer 2010.
30. Ravallion M. *Evaluating Anti-Poverty Programmes*. Development Research Group, World Bank 2008.
31. Villa JM. *Diff: simplifying the causal inference analysis with difference-in-differences*. University of Manchester 2012.