Social Science & Medicine 123 (2014) 96–104

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

The effects of performance incentives on the utilization and quality of maternal and child care in Burundi

Igna Bonfrer [a, *], Ellen Van de Poel [a], Eddy Van Doorslaer [a, b]

[a] *Institute of Health Policy & Management, Erasmus University Rotterdam, The Netherlands*
[b] *Erasmus School of Economics, Erasmus University Rotterdam, The Netherlands*

A R T I C L E I N F O

Article history:
Received 29 April 2014
Received in revised form
28 October 2014
Accepted 4 November 2014
Available online 6 November 2014

Keywords:
Burundi
Performance based financing
Incentives
Maternal and child care
Africa

A B S T R A C T

Africa's progress towards the health related Millennium Development Goals remains limited. This can be partly explained by inadequate performance of health care providers. It is therefore critical to incentivize this performance. Payment methods that reward performance related to quantity and quality, called performance based financing (PBF), have recently been introduced in over 30 African countries. While PBF meets considerable enthusiasm from governments and donors, the evidence on its effects is still limited. In this study we aim to estimate the effects of PBF on the utilization and quality of maternal and child care in Burundi. We use the 2010 Burundi Demographic and Health Survey (August 2010–January 2011, $n = 4916$ women) and exploit the staggered rollout of PBF between 2006 and 2010, to implement a difference-in-differences approach. The quality of care provided during antenatal care (ANC) visits improved significantly, especially among the better off, although timeliness and number of ANC visits did not change. The probability of an institutional delivery increased significantly with 4 percentage points among the better off but no effects were found among the poor. PBF does significantly increase this probability (with 5 percentage points) for women where PBF was in place from the start of their pregnancy, suggesting that women are encouraged during ANC visits to deliver in the facility. PBF also led to a significant increase of 4 percentage points in the probability of a child being fully vaccinated, with effects more pronounced among the poor. PBF improved the utilization and quality of most maternal and child care, mainly among the better off, but did not improve targeting of unmet needs for ANC. Especially types of care which require a behavioral change of health care workers when the patient is already in the clinic show improvements. Improvements are smaller for services which require effort from the provider to change patients' utilization choices.

1. Introduction

Africa's progress towards the health related Millennium Development Goals (MDGs) remains limited. The goal of reducing child mortality from 178 to 59 deaths per 1000 live births by 2015 is unlikely to be met given the current rate of 109 deaths per 1000 live births (United Nations, 2013). One of the reasons for this stagnation in health improvements is the inadequate performance of health care providers in low income countries (LICs) (Miller and Babiarz, 2013; Rowe et al., 2005). A study across a set of LICs found, through unannounced visits, a staggering 35 percent of absenteeism among health care providers. Since many of the providers actually present in the facility were not working, this percentage may still paint a too favorable picture (Chaudhury et al., 2006). Even if providers are delivering health care services, these are often of insufficient quality – referred to as the know-do gap (Leonard and Masatu, 2010; Peabody et al., 2006). Das and Gertler (2007), for example, compared doctor knowledge in Tanzania through a clinical vignettes study to their performance in actual daily practice. Results showed that doctors completed only 24 percent of the elements they knew how to do (as apparent from a vignettes study) when presented with a patient with malaria and 38 percent for a child with diarrhea (Das and Gertler, 2007). A similar result was found for Rwanda where providers knew on average 63 percent of appropriate procedures but delivered only 45 percent (Gertler and Vermeersch, 2012).

Given these examples of inadequate performance, it is critical to incentivize health care providers to behave in line with the best interest of their patients. A large number of African governments is currently piloting payment methods that reward performance in

* Corresponding author.
E-mail address: bonfrer@bmg.eur.nl (I. Bonfrer).

http://dx.doi.org/10.1016/j.socscimed.2014.11.004

the health care sector; Burundi and Rwanda have been the first countries to implement these methods nationwide (The World Bank Health Results Innovation Trust Fund, 2013). Through these performance based financing (PBF) schemes, health care facilities are paid retrospectively based on the quantity and quality of services provided. This is different from traditional health care financing mechanisms where budget flows are linked to for example number of beds or estimated drug needs. The PBF schemes typically affect health care provision in two ways: first, through incentives for providers to expend more effort in specific activities and second, through an increase in the amount of financial resources (Gertler and Vermeersch, 2012).

Over the last decade PBF has gained popularity among practitioners and governments (Magrath and Nichter, 2012; Meessen et al., 2011; Meessen et al., 2006; Soeters and Vroeg, 2011). More than 30 Sub Saharan African (SSA) countries are now in the process of introducing payment methods that reward performance or have already done so (Fritsche et al., 2014; Meessen, 2013). This enthusiasm is likely to be sparked further by the World Bank's recent pledge of 700 million dollar to be spent on women and children's health through performance based financing by 2015 (Kim, 2013). While the World Bank has initiated several PBF pilots across Africa with associated impact evaluations, the current knowledge base about the effects of PBF in LMIC is still quite limited (Eldridge and Palmer, 2009; Ireland et al., 2011; Kalk et al., 2010). A recent study by Miller and Babiarz confirmed that no formal evaluations are available for eighteen African countries where PBF has been piloted, including Burundi (Miller and Babiarz, 2013). A systematic review by Witter et al. (2012) on pay for performance in low and middle income countries, identified only one study (Peabody et al., 2011) – on the effects of bonuses for doctors in the Philippines – meeting high quality impact evaluation standards, with low risk of any bias. It found PBF to improve children's general self-assessed health and to reduce wasting but showed no effect on patient volumes. However, in this experiment similar effects were observed in another intervention group for which health insurance reimbursements to the hospitals were increased, suggesting that the effect mainly derived from increased resources. Though not considered low risk of bias by Witter et al., rigorous evidence has also been generated on the effects of PBF in Rwanda. Basinga et al. (2011) and Gertler & Vermeersch (2012) use a difference-in-differences analysis to show that PBF increased the use and quality of maternal and child services, and child nutritional outcomes. Sherry et al. (2013) use the same experimental design, but a different dataset and find a significant increase in the proportion of women delivering in facilities but no impact of PBF on antenatal care utilization, child vaccinations and contraceptive use.

More recently Bonfrer et al. (2014) have examined the effects of a pilot PBF program in Burundi with implementation support from the Dutch Non-Governmental Organization (NGO) Cordaid in 9 out of 17 provinces (Bubanza, Bururi, Cankuzo, Gitega, Karuzi, Makamba, Muramvya, Rutana and Ruyigi). Using three waves of data, they study the effects of PBF on antenatal care, institutional deliveries, vaccinations, modern family planning, reported patient satisfaction and a quality score based on a checklist for health care facilities. They find positive effects of PBF on institutional deliveries, antenatal care utilization, modern family planning and the quality score, though this latter finding is not reflected in an increased reported patient satisfaction. The present study builds on this earlier work and extends the analysis in several respects. First, and perhaps most importantly, we evaluate the *nationwide* effects of PBF, i.e. including provinces where NGOs other than Cordaid provided technical support to the Ministry of Health (MoH) for implementation. Furthermore, the use of the Burundi Demographic and Health Survey (BDHS) offers important advantages over the data collected by Cordaid. Not only is the sample size about 9 times larger (7742 births), the BDHS also provides a broader range of outcome measures related to (the timing of) vaccinations and the content of ANC provided. In addition to the number of ANC visits, it registers whether the mother's blood pressure was taken, whether she received an anti-tetanus vaccination and whether the first visit was in the first trimester of pregnancy. Finally, it is important to note that the BDHS data were collected independently from the PBF program, while the earlier study used data collected by the implementing agency Cordaid, which might potentially have affected reporting and induced bias. Independent data collection is especially important in a context of PBF, where random visits and interviews are conducted to verify quantity and quality of care which are used as parameters to determine payments to facilities.

Given the limited evidence on the effectiveness of PBF in SSA and the considerable expansion of this financing mechanism across the continent, there is an urgent need for evidence on impact. This study contributes evidence on the effects of PBF in Burundi, a country where PBF has gradually become a nationwide policy in the period from 2006 to 2010. Burundi is a post-conflict country, among the lowest income countries in the world with a GDP of barely 251 current US$ per capita (World Bank, 2012) and located in central Africa. The case of Burundi is especially interesting as health systems in post-conflict states are often forced to innovate which can generate useful lessons for other settings (Witter, 2012). The health status of the population is poor as reflected in an infant mortality rate of 67 per 1000 live births compared to 35 per 1000 worldwide (World Bank, 2012).

The primary aim of the PBF scheme in Burundi was to improve maternal and child health (Busogoro and Beith, 2010). We therefore study its effects on the quantity of child and maternal care use and its quality based on the reported services provided during antenatal care (ANC) visits. We use the BDHS and exploit the staggered rollout of PBF across provinces between 2006 and 2010, to implement a difference-in-differences approach.

In the following sections we first discuss the details of the PBF scheme introduced in Burundi, followed by a description of the data and the statistical analyses. Then we discuss the common tend assumption, followed by the estimated effects. These effects are discussed and we end with some concluding remarks.

2. Performance based financing in Burundi

Starting from the end of 2006, PBF was implemented in almost 700 health care facilities in Burundi (The World Bank Health Results Innovation Trust Fund, 2013). Based on quantity and quality of services provided, facilities receive performance related funding (Bertone and Meessen, 2012) which on average makes up 40 percent of the total facility budget (The World Bank Health Results Innovation Trust Fund, 2013). Quantity is measured using various output indicators including ANC, vaccinations, family planning, and HIV care (Ministère de la santé publique Republique du Burundi, 2010). Different levels of PBF payments are associated with these output indicators as shown in Table 1.

Health care facilities report monthly to the MoH about the quantity of incentivized services delivered. A provincial committee verifies and validates the reported quantities through unannounced visits to facilities. On top of the quantity based payments, facilities receive a quality bonus ranging from 0 to 25 percent. Local regulatory authorities assess the quality every three months on a randomly chosen day using a standardized checklist (Appendix 1) procedure for availability of medical supplies, equipment, administrative procedures, prescription behavior, lab services and hygiene (Busogoro and Beith, 2010; Kamana, 2012; Soeters, 2013; The World Bank, 2010). Based on information about quantity and

Table 1
Average payments for output indicators.

Health care service j	Average payment P_j in US dollar
Children 6–59 months receiving Vit A	0.05
Outpatient consultancy – new case	0.25
Antenatal care: new and standard visits	0.40
Diagnosis and treatment of STD	0.50
In patient bed day	0.50
Pregnant woman fully immunized	0.50
Small surgery intervention	0.50
Latrine newly constructed	0.70
Child treated after birth HIV mother	1.00
Family planning: referral of tubal ligation and vasectomy	1.00
HIV mother treated	1.00
Patient referred to hospital and feedback obtained	1.00
Pregnant woman counseled and tested for HIV	1.00
Person voluntary counseled and tested for HIV	1.00
Bed net distributed	1.50
Child under 1 completely immunized	1.50
HIV case diagnosed and referred	1.50
Family planning: new and re-attendants, oral & injectable	2.00
HIV mother referred to hospital	2.00
Institutional delivery by qualified staff	2.00
Family planning: implant or IUD	5.00
Patient diagnosed with TB (3 sputum checks)	10.00
TB patient correctly treated during 6 months	20.00

Source: Ministère de la santé publique (2011)

quality, the formula used to calculate the total subsidy to facility i in period t is

$$PBF\ \text{subsidy}_{it} = \left(\sum_{j=1}^{J} P_j N_{ijt}\right) \cdot Q_{it} \ \text{ with } \ 1 \le Q_{it} \le 1.25 \qquad (1)$$

where P_j is the subsidy received by the health care facility per health care service j. N_{ijt} is the number of services j delivered in facility i over period t. Q_{it} is the quality bonus which health care facility i receives in period t, ranging from 0 to 25 percent, depending on the score obtained from the checklist (Bonfrer et al., 2014). Health care facilities have some autonomy to decide on the allocation of PBF revenues across two broad categories: up to 50 percent of payments can be used for staff remuneration and the remaining share must be invested in service quality improvements (Busogoro and Beith, 2010).

Table 2 details the rollout dates of the program: PBF was first implemented by the MoH with help from non-governmental organizations (NGOs) in three provinces in December 2006 and over time all other provinces were added (Bonfrer et al., 2014; Busogoro and Beith, 2010; Ministère de la santé publique Republique du Burundi, 2011). The selection of provinces in the early (pilot) phase of the PBF roll-out was done by the MoH, with the aim to ensure comparability across intervention and control provinces. When presenting descriptive statistics, we group the provinces into early adopters starting December 2006 (Bubanza, Cankuzo and Gitega), middle adopters starting October 2008 (Bururi, Karuzi, Makamba, Rutana, Ruyigi) and later adopters starting mostly April 2010 (Ngozi, Bujumbura-rural, Cibitoke, Kayanza, Kirundo, Muramvya, Muyinga, Mwaro and Bujumbura-mairie).

Before the introduction of PBF, in May 2006, user fees for deliveries and care for under-fives were removed throughout Burundi in governmental, non-faith based, facilities (Nimpagaritse and Bertone, 2011). These facilities received payments from the government for the services provided for free. Following the problematic implementation of this policy (Kamana, 2012; Nimpagaritse and Bertone, 2011), the MoH decided to make the reimbursement for these maternal and child care services also performance based by incorporating it into the PBF scheme in April 2010. Given that this was a nationwide program implemented before the start of PBF, it does not affect our estimates which are based on differences in the changes in outcomes over time, isolating the part of change attributable to PBF.

Table 2
Timing of PBF introduction.

Bubanza	1-12-2006
Bujumbura-mairie	1-4-2010
Bujumbura-rural	1-4-2010
Bururi	1-10-2008
Cankuzo	1-12-2006
Cibitoke	1-4-2010
Gitega	1-12-2006
Karuzi	1-10-2008
Kayanza	1-4-2010
Kirundo	1-4-2010
Makamba	1-10-2008
Muramvya	1-4-2010
Muyinga	1-4-2010
Mwaro	1-4-2010
Ngozi	1-12-2009
Rutana	1-10-2008
Ruyigi	1-10-2008

Source: Bonfrer et al., 2014 and Busogoro and Beith, 2010

3. Methods

3.1. Data

We use the Burundi Demographic and Health Survey (BDHS) data collected in 2010 (August 2010–January 2011, after nationwide rollout of PBF as shown in Table 2), in which a nationally representative sample of 4916 women was asked about maternal and child care use for all of their pregnancies in the past five years. It thus provides information on births occurring in the period 2005–2010, during which PBF was rolled out in Burundi, and for every reported birth we identify whether PBF was being implemented in the mother's province of residence in the month that the delivery took place. In the case of antenatal care we identify whether PBF was in place nine months earlier, at the start of the pregnancy.

As outcome measures we use all (11) pregnancy and birth related health care services registered in the BDHS that are incentivized in the PBF scheme. The first four outcomes relate to ANC and were collected for the most recent birth among women who gave birth in the last five years. Since virtually all women in the sample used ANC at least once (99%), we create a variable indicating whether the mother used ANC from a doctor, nurse or midwife more than once before she gave birth (>1 *ANC*). *First trimester ANC visit* indicates whether the mother used ANC from a doctor, nurse or midwife at least once before the end of the first trimester of the pregnancy. Although timeliness of the first visit was not directly incentivized through PBF, it is an important measure to monitor quality. *Blood pressure* (*BP*) *measurement* and ≥1 *anti-tetanus vaccination* indicate whether the mother had her BP measured and whether she received at least one anti-tetanus vaccination in the prenatal period. Both indicators are internationally recognized as essential elements of good quality ANC (Lincetto et al., 2007; World Health Organization, 2006). Anti-tetanus vaccinations in pregnant women were incentivized through the PBF subsidy 'pregnant woman fully immunized' while there was no specific subsidy for the BP measurement of pregnant women. *Institutional delivery* indicates whether the mother gave birth in a public health care facility while a doctor, nurse or midwife was present.

Regarding child outcomes, which are collected for all children born in the last five years, we include indicators of whether children were fully vaccinated -as shown on their vaccination cards-by the age of one year (excluding those younger than one at the time of survey) (*child fully vaccinated at* 1 *year*) and more specific indicators for each of the vaccinations received (BCG vaccination at 1 year, polio vaccination at birth, three additional polio vaccinations at 1 year, three DTP vaccinations at 1 year and *measles vaccination at* 1 *year*). Although information about child mortality is also available in the BDHS, we have insufficient power to detect a plausible effect (for $\alpha = 0.05$ and power $= 0.80$ a sample size of at least 32,358 births would be necessary to detect a neonatal mortality reduction of 12% as we currently observe across births without and with PBF). Table 3 shows sample sizes for all outcome measures for children with and without PBF at birth or from the start of the pregnancy. All models control for household characteristics (size, socioeconomic status, age and sex of the household head and access to water and electricity) and for mother's demographics (age at birth, education, first pregnancy) as shown in Table 4. Socioeconomic status is measured by a wealth index, included in the BDHS, and estimated from principal component analysis on a large set of assets and dwelling characteristics (Filmer and Pritchett, 2001).

No ethical approval was sought because no primary data collection took place for this study.

3.2. Statistical analysis

We identify the effects of PBF by comparing changes in the outcome measures in provinces with PBF (treated) to changes in provinces without PBF (controls). Subject to the common trend assumption – which requires that the trend among the controls is a valid counterfactual of what would have happened to the treated in the absence of PBF – this difference-in-differences strategy isolates the part of the change that is causally attributable to the impact of PBF (Imbens and Wooldridge, 2009). We assess the plausibility of this assumption in the Results section.

We implement the difference-in-differences approach by estimating the following probit model for each of the k dichotomous outcome measures y^k_{itp} (>1 *ANC*, *first trimester ANC visit*, *BP measurement*, *anti-tetanus vaccination*, *child vaccinations* and *institutional delivery*) for pregnancy/child i at time t in province p as follows:

$$y^{k*}_{itp} = \alpha_k + \beta_k PBF_{tp} + X_{itp}\Psi + \delta_t + \phi_p + \varepsilon_{itp} \quad (2)$$

Table 4
Sample means for all control variables.

	Early adopters	Middle adopters	Later adopters	Total
Household size	5.85	6.02	5.90	5.92
Lowest wealth quintile	0.25	0.17	0.19	0.19
Lower wealth quintile	0.23	0.23	0.17	0.20
Higher wealth quintile	0.18	0.20	0.17	0.19
Highest wealth quintile	0.13	0.17	0.31	0.24
Child is mother's firstborn	0.21	0.20	0.21	0.21
Mothers' age at birth <21 years	0.54	0.56	0.52	0.54
Mothers' age at birth >35 years	0.46	0.44	0.48	0.46
Mother no primary education	0.52	0.54	0.46	0.49
Age of household head in years	36.39	36.35	36.81	36.61
Male household head	0.87	0.85	0.83	0.84
Safe drinking water	0.62	0.69	0.65	0.65
Household has electricity	0.06	0.04	0.12	0.09

where y^{k*}_{itp} is a latent index and the error term (ε_{itp}) is drawn from a normal distribution. Our main interest lies in the effect (β_k) of the PBF indicator (PBF_{tp}) which is switched on if province p had PBF at time t when the child is born (or when the mother was pregnant). The model includes 63 birth period indicators (δ_t) for the month and year in which the child was born to capture the time trend in outcomes common to the intervention and control areas, and a full set of province effects (Φ_p) to capture time invariant differences between provinces. We use 9 birth half year indicators instead of the 63 month-year indicators for >1 *ANC* because this model had problems converging with the large number of fixed effects and the relatively small sample size. Controlling for time varying individual variables (X_{itp}) accounts for the differences in observable characteristics between treated and controls, and adds precision. Standard errors are adjusted for clustering at the province level (Angrist and Pischke, 2008; Bertrand et al., 2004). We have confirmed robustness of our results to using ordinary least squares regression models (significance is lost for the effect on two of the specific vaccination variables, see Appendix 2).

For each outcome measure, the result is presented as the average partial effect of the PBF indicator among the births (or pregnancies) which had taken place at a time and in a province where PBF was implemented, i.e. the average treatment effect on the treated (ATET). In extended probit models, the effect is allowed to differ by poverty status by interacting the PBF indicator with a variable indicating whether the household is in the bottom two

Table 3
Sample sizes and baseline means of outcome measures.

	Sample sizes			Baseline means				
	No PBF	PBF	Total	Early adopters	Middle adopters	Later adopters	Joint test of significance, *p*-value	Total
>1 antenatal care	3603	1299[a]	3603	0.98	0.94	0.96	0.26	0.96
First trimester antenatal care	3576	1289[a]	3576	0.18	0.14	0.28	0.00	0.23
Blood pressure measurement	3576	1289[a]	3576	0.36	0.50	0.59	0.00	0.53
≥1 Anti-tetanus vaccination	3614	1302[a]	3614	0.59	0.60	0.68	0.08	0.64
Institutional delivery	5253	2489	7742	0.50	0.46	0.45	0.26	0.46
Child fully vaccinated at 1 year	4595	1177	5772	0.31	0.27	0.29	0.48	0.29
BCG vaccination at 1 year	4595	1177	5772	0.34	0.32	0.33	0.79	0.33
Polio vaccination at birth	5253	2489	7742	0.31	0.29	0.30	0.83	0.30
Three additional polio vaccinations at 1 year	4595	1177	5772	0.34	0.30	0.32	0.58	0.32
Three DTP vaccinations at 1 year	4595	1177	5772	0.34	0.31	0.33	0.64	0.32
Measles vaccination at 1 year	4595	1177	5772	0.33	0.29	0.32	0.61	0.31

[a] PBF in place when the mother was pregnant.

wealth quintiles (poor). All statistical analyses were done in Stata 12.

4. Results

4.1. Summary statistics and the common trend assumption

Table 4 compares subgroup means for all control variables. While there are some significant differences, this is not necessarily a problem for the difference-in-differences analysis. The more important assumption — that the trends in outcomes for treatment and control groups are parallel in the absence of treatment — cannot formally be tested. However, to assess its credibility we compare pre-intervention trends in maternal and child health care across the three groups of provinces using data from the Multiple Indicator Cluster Surveys (MICS) collected in Burundi in 2000 and 2005 that provide information on births in 1996–2004, i.e. prior to the introduction of PBF. The MICS does not collect exactly the same outcome variables as the BDHS, but it has information on ownership of a child's vaccination card, one BCG vaccination at birth or closely after that, three or more doses of polio vaccination, three doses of DTP vaccination and at least one dose of measles vaccination. Fig. 1 displays the trends in these five indicators for mother and child care and confirms that trends were very similar across the three groups of provinces in the period prior to the start of the PBF program. Estimating a probit model with indicators for birth years, indicators for early, middle and later adopters and the interactions between these two shows for all outcome variables that interactions are not jointly significant (p-values range from 0.321 for Measles to 0.698 for DTP), implying that the trends were parallel before the intervention. We do see that the improvement in vaccination coverage rates occurred somewhat later for the early adopters (2002) compared to middle and late adopters (2001). Overall fluctuations in coverage rates are likely to be driven by socio-political instability arising from the ethnic conflict between Hutus and Tutsis from 1993 till 2005. Similar fluctuations were not only reflected in vaccination rates but they are apparent in the Burundi economic indicators in general, as reflected in for example trade levels (see Appendix 3).

Furthermore, in the remaining columns in Table 3 we present baseline (2005–2006) means for selected outcome variables from the BDHS for the early, middle and later adopters. For each outcome, we test for area differences by estimating a simple model with only three covariates: indicators for the early, middle and later adopters. The p-values from the joint test of significance (in 7th column of Table 2) show that there are no significant between-area differences in terms of baseline levels for antenatal care, anti-tetanus vaccinations, institutional delivery and child vaccinations. This suggests that no structural differences were present across the groups of provinces nominated for the different stages of roll-out.

4.2. Estimated effects

Table 5 presents estimated average marginal effects of PBF for all outcome measures. The estimates suggest that PBF in Burundi did *not* affect the probability of a child's mother receiving more than

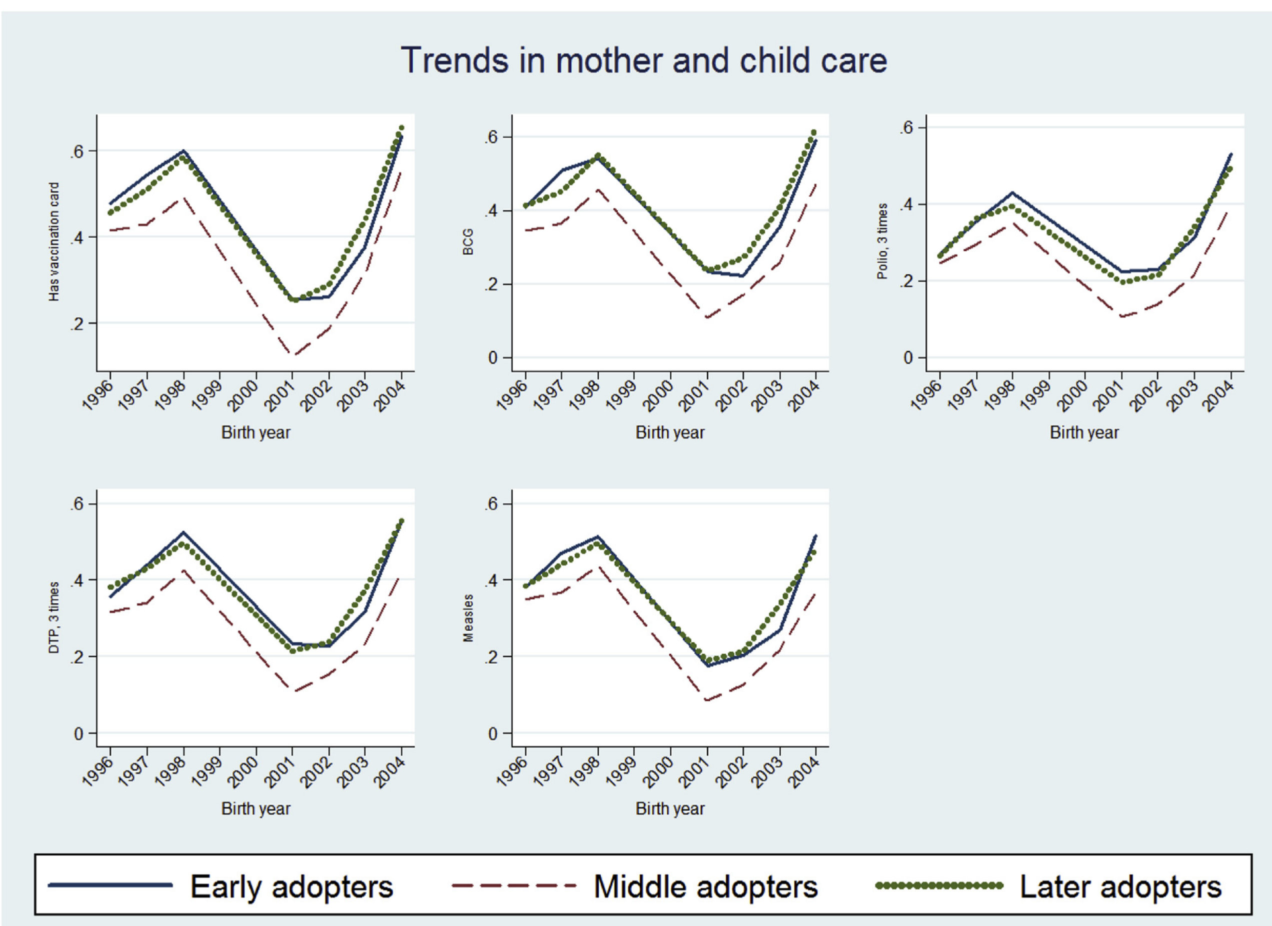

Fig. 1. Pre-intervention trends in mother and child care.

Table 5
Probit model showing average marginal effects of PBF on treated, using interaction effects.

	Full model		Full model	Full model		Subset model	
	Sample size full model		Alternative treatment definition	Poor	Non-poor	Sample size subset model	
>1 antenatal care visit	4902	−0.004	n/a	−0.001	−0.007	2471	−0.006
First trimester antenatal care	4865	0.014	n/a	0.025	0.006	2448	0.029
BP measured during pregnancy	4865	0.061*	n/a	0.055	0.064**	2448	0.063*
≥1 anti-tetanus vaccination	4916	0.100***	n/a	0.106***	0.107***	2477	0.114**
Institutional delivery	7742	0.027	0.051*	0.003	0.043**	3961	0.051***
Child fully vaccinated at 1 year	5772	0.044*	−0.008	0.066**	0.027	2974	0.064**
BCG vaccination at 1 year	5772	0.037***	0.003	0.049*	0.027*	2974	0.060***
Polio vaccination at birth	7742	0.069***	0.011	0.063***	0.076***	3961	0.071***
Three additional polio vaccinations at 1 year	5772	0.032**	−0.029	0.048*	0.019	2974	0.048***
Three DTP vaccinations at 1 year	5772	0.031**	−0.024	0.047*	0.019	2974	0.050***
Measles vaccination at 1 year	5772	0.042**	−0.002	0.058*	0.030	2974	0.061**

*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.
[d] All models contain control variables as shown in Table 4 and region and time controls.
[e] The full model with alternative treatment definition estimates the effects of PBF when in place from the start of pregnancy (as opposed to being in place at birth) for the outcomes related to delivery and vaccinations. For the other outcome variables, relating to antenatal care, the treatment was already defined in this manner, so no alternative treatment definition is defined (indicated as n/a).
[f] The subset model is based on data from the subset of provinces used in the earlier study by Bonfrer et al. (2014) with data from the NGO Cordaid: Bubanza, Bururi, Cankuzo, Gitega, Karuzi, Makamba, Muramvya, Rutana and Ruyigi.

one ANC visit during pregnancy or that the ANC visit took place within the first trimester. However, we do find a marginally significant ($p = 0.065$) increase of 6 percentage points (pp) in the probability that a mother reported to have had her BP measured at least once during her pregnancy. The likelihood of receiving one or more anti-tetanus vaccinations (as part of the ANC) also increased significantly ($p < 0.000$) and substantially (10 pp). This suggests that while PBF did not lead to a further increase in the quantity or timeliness of ANC provided – most likely because the utilization rate for >1 antenatal care was already so high at baseline (96%) – it did improve the quality of ANC provided as measured by whether certain important components were delivered. The positive coefficient of PBF is not significantly different from zero for the probability of delivering in an institution. We further examined this finding in a number of ways. The finding was not different between urban or rural residence of the mother. We also did not find any effect on deliveries when the outcome measure was relaxed to no longer require the presence of a doctor, nurse or midwife. Neither did we find any effect when the outcome measure was expanded to not only include governmental but also private facilities (detailed results available upon request). We did, however, find that PBF significantly raises the probability of a child being fully vaccinated (4pp; $p = 0.060$), an effect that was driven by an increase in all components of the vaccination package (BCG, polio, DTP and measles). We found no significant effect of PBF on neonatal mortality, likely caused by the lack of power necessary to pick up an effect of credible size, as mentioned in the Data section.

Columns 5 and 6 in Table 4 allow for heterogeneity of the effects across poor and non-poor respondents. The effects of vaccinations are generally stronger for children in poor households, except for the polio vaccine, which is the only vaccination provided immediately at birth. For antenatal care, we find little poor versus non-poor differences in the probability of having a tetanus vaccination, but an increased probability of BP measurement during pregnancy only occurs among the non-poor. Institutional deliveries increased significantly (by 4pp; $p = 0.028$) among the non-poor, while there is no effect among the poor.

5. Discussion

In this study we use the nationwide rollout of PBF in Burundi to estimate its effects on the utilization and quality of maternal and child care services. We do not find any evidence that PBF affects the likelihood of the mother receiving more than one ANC visit during pregnancy, or of an ANC visit to occur in the critical first trimester. However, we do find a significant rise in the likelihood of BP measurement and anti-tetanus vaccination as part of the ANC. We also found the increase in BP measurement during pregnancy to be fully driven by the (larger) effect among the non-poor. This implies that PBF improved the quality of care during ANC visits, especially among the non-poor, but not the targeting of unmet ANC needs. The fact that institutional deliveries were found to increase significantly among the non-poor, but not among the poor, fuels concerns about lower effectiveness of PBF where it is needed most. Greater effects on institutional deliveries among the better off could indicate that PBF might not improve equity in outcomes as it does not overcome demand side barriers. While in principle fees for delivery care are waived, it is likely that other costs, like transportation, might constrain poor women more to deliver in a facility. This argument seems to hold less for vaccinations, for which demand side constraints may be less binding as its administration is less urgent, costly and time consuming. PBF was also found to significantly increase the administration of a full vaccination package, including all of its components (BCG, polio, DTP and measles). The vaccinations effects are stronger for children in poor households, except for the polio vaccine provided at birth. The latter may relate to the higher rates of institutional deliveries within the non-poor group, which gives them better access to this service.

Are these large or small effects and how do they compare to earlier work on PBF in Burundi? Our findings show a smaller impact of PBF overall compared to the PBF pilot program which was evaluated in a subset of 9 out of 17 provinces (Bonfrer et al., 2014). The main differences in findings are the absence of an effect on institutional deliveries and on having more than one ANC visit during pregnancy. In part, this is due to the selection of provinces for the PBF pilot evaluation: we also find a significant 5pp increase in the probability of delivering in an institution if we restrict our analyses to the 9 provinces included in the pilot evaluation, but even then the effect is considerably smaller than the reported pilot effect (22pp). Another source of differences in the effect on institutional deliveries is the timing of the data collection. Our evaluation includes births which took place directly after the introduction of PBF while the evaluation of the pilot was based on information from births which in almost all cases occurred at least 9 months after the introduction of PBF. To test whether and to which extent this difference in timing of data collection drives results, we changed the definition of the PBF indicator to reflect PBF presence at the start of the pregnancy as opposed to only just before the delivery (see column 3 in Table 5). We then do find a significant effect of PBF on institutional deliveries of 5 pp ($p = 0.058$), which suggests that PBF needs to be in place at the start of a pregnancy to have an effect on institutional deliveries. To test whether ANC utilization is the pathway explaining this finding, we include >1 *ANC* as an explanatory variable in the probit model and find that the effect of PBF on institutional deliveries ($p = 0.182$) indeed vanishes and that ANC as explanatory variable is positive and significant (17pp; $p = 0.000$). This strongly suggests that health care providers encourage women during subsequent ANC visits to deliver in the facility, as also suggested by Gertler & Vermeersch (2012) for Rwanda. Overall, however, our estimated effect of PBF on institutional delivery in the nationwide program remains smaller than in the earlier pilot evaluation.

The effect on the probability of more than one ANC visit during pregnancy found in the pilot study (10pp) is not reproduced with the BDHS when restricting the analysis to the subset of provinces, though Bonfrer et al. (2014) also did not find this in all cases. The estimated size of the effect on anti-tetanus vaccination for pregnant women is the same in both studies (11pp) but was not significant in the pilot study, possibly due to its smaller sample size. Estimated effects on child vaccinations are difficult to compare, since the pilot study only collected information on whether the child received at least one vaccination for each of the diseases (and no information on its timing). Moreover, the pilot study was based on mothers' reports of their child's vaccination while the BDHS data are derived from the child's vaccination card. The reported baseline vaccination rates in the pilot study were therefore much higher than those reported in the BDHS (95% versus 46% for BCG vaccination), and this probably explains why the pilot evaluation did not find an effect while this study did (3.7pp increase, $p = 0.000$). The more detailed BDHS information also reveals substantial effects of PBF on the probability of children receiving the full course of vaccination. Finally, our findings on blood pressure measurement and anti-tetanus vaccination suggest that quality of care has improved because of PBF. Bonfrer et al. (2014) also reported that pilot provinces included in the first phase of roll-out started off at a lower baseline and had therefore more room for improvement. They found effects indeed to be smaller (institutional delivery) or no longer significant (antenatal care) for births in provinces where PBF was rolled out in the second phase. This may also partly explain the differences between the effects of the pilot program and our results based on all provinces, where a larger share of the sample did not obtain PBF in the first phase.

Our results can also be compared to the estimates effects of PBF in neighboring Rwanda, the only other country with nationwide PBF (see Appendix 4 for details), even though we need to be cautious in comparing PBF program effects across countries given the heterogeneity in organizational, social, and institutional environments (Miller and Babiarz, 2013). First we therefore highlight the main differences in health care financing, content of the PBF scheme itself, and set-up of the impact studies in both countries.

Sherry et al. (2013) use a similar approach as ours with the Rwanda *Demographic and Health Surveys* of 2005 and 2007/08. They find a significant increase in the proportion of women delivering in facilities (9.8pp) but no impact of PBF on antenatal care utilization, child vaccinations and contraceptive use. They do find an impact on non-rewarded services related to antenatal care specifically urinalysis (5.1 pp) and iron supplementation (9.3 pp). No significant effect is found on other aspects of antenatal care like vitamin A supplementation. They also do not find any effect on health outcomes. Basinga et al. (2011) confirm considerable effects on institutional deliveries (23pp) and the absence of significant impacts of PBF in Rwanda o vaccinations and antenatal care.

The set-up of PBF was very similar in Burundi and Rwanda, but subsidies were slightly lower in Rwanda. One clear difference lies in the fact that in Rwanda the control provinces received additional funding, enabling differentiation between the incentive and resource effect, while this was not the case in Burundi (Basinga et al., 2011). The financing system for the demand side also differed between the two countries: Rwanda introduced PBF within a system of community based health insurance while in Burundi user fees for deliveries and care for under-fives were removed. The differential relative importance given to certain services is apparent from the subsidy levels and shows a greater focus on HIV/AIDS care and institutional deliveries in Rwanda and a focus on family planning and TB care in Burundi. Generally, subsidies for vaccinations were higher in Burundi while those for institutional delivery were higher in Rwanda. We indeed see effects of PBF on vaccinations in Burundi (4pp increase; $p = 0.060$), while this is not the case for Rwanda (Basinga et al., 2011; Gertler and Vermeersch, 2012; Sherry et al., 2013) and no or smaller effects on institutional delivery in Burundi compared to Rwanda (9.8 pp increase) (Sherry et al., 2013). The latter difference might also be explained by the relatively low subsidy for institutional deliveries in Burundi where it is not among the highest subsidized services, as it is in Rwanda. Further research is necessary to provide insights into the effectiveness of PBF in the context of demand side interventions like community based health insurance (Robyn et al., 2014) or user fee removal. All in all, it seems that the differences in the relative importance given to specific aspects of the PBF design also produced differences in incentivized outcomes.

6. Conclusion

It seems clear that PBF has had some positive impacts in Burundi. Especially for types of care which require a behavioral change of health care workers when the patient is already in the clinic we see improvements. While the likelihood of measuring a pregnant woman's BP and giving an anti-tetanus injection during an ANC visit increased, mothers did not respond by going for more than one ANC visit or to initiate these visits during the critical first trimester. This is consistent with the argument of Gertler & Vermeersch (2012) that it is more difficult for PBF to increase utilization of services that depend on patient choices than services that are under the provider's control. They also argue that initiation of care takes more effort than its continuation, suggesting that policy makers might want to increase the PBF unit payments for first time antenatal care use. Like Gertler and Vermeersch, we believe that conditional cash transfers to poor women for the timely utilization of ANC may be more effective in influencing decisions of care utilization over which patients have direct control. Provider incentives like PBF seem more appropriate to incentivize the quality or content of ANC visits which is more controlled by health care providers. Cost-effectiveness comparisons could shed light on the relative efficiency of both health care financing options.

Some qualifications are in order. While many of the outcomes used are accepted measures of clinical performance (e.g. prenatal care use in first trimester of pregnancy), they are still indicators of processes, not outcomes. The assumption is that these result in improved patient or population health outcomes, but this may not always hold. It would obviously be preferable if performance incentives rewarded health improvement directly rather than the use of health services or other health inputs.

There are some limitations to the analyses presented in this paper. First, because of the non-random rollout of PBF, it is possible that unobservables affect both the placement of the PBF and our outcomes of interest which creates bias on our impact estimates. While it is not possible to test for this problem, the common trends in health care use in the period prior to PBF does give credence to the parallel trends assumption. Given that this intervention was implemented at the province level ($n = 17$), a randomized set-up is unlikely to guarantee comparability across intervention and control provinces. Generally, randomization at a lower level is preferable from a design point of view, but in practice often difficult to achieve. Second, the implementation of PBF in Burundi, as in many other contexts, involved a change in provider incentives coupled with a substantial increase in their budget. It was in this study not possible to separate these effects, while this would obviously be very valuable information for policy makers.

Notwithstanding these limitations, our study is among the first to show evidence of PBF in an African country affecting the utilization and quality of maternal and child care. It is also the first to compare results from a PBF pilot to those of a nationwide program. This is timely given the many PBF pilots set up across SSA and likely to transit to nationwide programs in the near future. Although the World Bank has invested considerably in evaluating PBF, not many of these impact evaluations have yet been finalized. The rapid growth in the use of performance pay (Miller and Babiarz, 2013) will provide ample scope for evaluation and evidence generation on the question whether and how PBF can live up to the expectation of improving access to good quality care in SSA.

Acknowledgments

We thank Robert Soeters, Frank van de Looij, Olivier Basenya and Gashubije Longin for sharing their in depth knowledge about the practicalities of performance based financing in Burundi. We thank Michael Grimm and Geert Frederix for helpful comments on the manuscript. We are also grateful to MeasureDHS for the public availability of the Burundi Demographic and Health Survey 2010 and to UNICEF for the public availability of the Multiple Indicator Cluster Surveys 2000 and 2005. Ellen Van de Poel acknowledges the Netherlands Organization for Scientific Research (Veni project 451-11-031).

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.socscimed.2014.11.004.

References

Angrist, J., Pischke, J., 2008. Mostly Harmless Econometrics. Princeton University Press, Princeton.
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 (9775), 1421–1428.
Bertone, M.P., Meessen, B., 2012. Studying the link between institutions and health system performance: a framework and an illustration with the analysis of two performance-based financing schemes in burundi. Health Policy Plan..
Bertrand, M., Duflo, E., Mullainathan, S., 2004. How much should we trust difference-in-differences estimates? Q. J. Econ. 119, 249–275.
Bonfrer, I., Soeters, R., Van de Poel, E., Basenya, O., Longin, G., van de Looij, F., et al., 2014. Introduction of performance based financing in Burundi associated with improvements in care and quality. Health Aff. (forthcoming).
Busogoro, J., Beith, A., 2010. Pay for Performance for Improved Health in Burundi. Retrieved 04/03, 2013 from. http://www.healthsystems2020.org/content/resource/detail/2575/.
Chaudhury, N., Hammer, J., Kremer, M., Muralidharan, K., Rogers, F., 2006. Missing in action: teacher and health worker absence in developing countries. J. Econ. Perspect. 20 (1), 91–116.
Das, J., Gertler, P.J., 2007. Variations in practice quality in five low-income countries: a conceptual overview. Health Aff. 26 (3), W296–W309.
Eldridge, C., Palmer, N., 2009. Performance-based payment: some reflections on the discourse, evidence and unanswered questions. Health Policy Plan. 24 (3), 160–166.
Filmer, D., Pritchett, L., 2001. Estimating wealth effects without expenditure data - or tears: an application to educational enrollments in states of india. Demography 38 (1), 115–132.
Fritsche, G., Soeters, R., Meessen, B., van Heteren, G., Ndizeye, C., Bredenkamp, C., 2014. Performance Based Financing Toolkit. Retrieved 27 March 2014, 2014 from. http://go.worldbank.org/L6CDLHSSW0.
Gertler, P., Vermeersch, C.M., 2012. Using Performance Incentives to Improve Health Outcomes. In: Policy Research Working Paper, 6100, pp. 1–39.
Imbens, G., Wooldridge, J.M., 2009. Recent developments in the econometrics of program evaluation. J. Econ. Literature 47, 5–86.
Ireland, M., Paul, E., Dujardin, B., 2011. Can performance-based financing be used to reform health systems in developing countries? Bull. World Health Organ. 89 (9), 695–698.
Kalk, A., Paul, F.A., Grabosch, E., 2010. Paying for performance in rwanda: does it pay off? Trop. Med. Int. Health 15 (2), 182–190.
Kamana, J., 2012. Selective Free Health Care & Performance Based Financing in Burundi: An Opportunity to Improve the Provision of Health Services, Especially for Reproductive Health. Ministry of Public Health and Fight Against HIV/AIDS (Performance Based Financing National Technical Cell).
Kim, J., 2013. World Bank Group President Jim Yong Kim's Speech at World Health Assembly: Poverty, Health and the Human Future. Retrieved 01/17, 2014 from. http://www.worldbank.org/en/news/speech/2013/05/21/world-bank-group-president-jim-yong-kim-speech-at-world-health-assembly.
Leonard, K.L., Masatu, M.C., 2010. Professionalism and the know-do gap: exploring intrinsic motivation among health workers in tanzania. Health Econ. 19 (12), 1461–1477.
Lincetto, O., Mothebesoane-Atoh, S., Gomez, P., Munjana, S., 2007. Standards for Maternal and Neonatal Care, Provision of Effective Antenatal Care. World Health Organization, Geneva.
Magrath, P., Nichter, M., 2012. Paying for performance and the social relations of health care provision: an anthropological perspective. Soc. Sci. Med. 75 (10), 1778–1785.
Meessen, B., 2013. Going Beyond Impact Evaluation Performance Based Financing of Health Care Provision in Low Income Countries (Unpublished manuscript).
Meessen, B., Soucat, A., Sekabaraga, C., 2011. Performance-based financing: just a donor fad or a catalyst towards comprehensive health-care reform? Bull. World Health Organ. 89 (2), 153–156.
Meessen, B., Musango, L., Kashala, J.I., Lemlin, J., 2006. Reviewing institutions of rural health centres: the performance initiative in butare, rwanda (passage en revue des institutions rurales de centre de santé: L'initiative de performance a butaré au rwanda; revisando instituciones de centros sanitarios rurales: La iniciativa de rendimiento de butare, ruanda). Trop. Med. Int. Health 11 (8), 1303–1317.
Miller, G., Babiarz, K., 2013. Pay-for-performance Incentives in Low- and Middle-income Country Health Programs (Unpublished manuscript).
Ministère de la santé publique Republique du Burundi, 2011. Rapport de la mise en oeuvre du financement base sur la performance au burundi.
Ministère de la santé publique Republique du Burundi, 2010. Manuel des procedures pour la mise en oeuvre du financement base sur la performance au burundi.
Nimpagaritse, M., Bertone, M., 2011. The sudden removal of user fees: the perspective of a frontline manager in burundi. Health Policy Plan. 26 (Suppl. 2) ii63-ii71.
Peabody, J., Taguiwalo, M., Robalino, D., Frenk, J., 2006. Improving the quality of care in developing countries. In: Jamison, D., Breman, J., Measham, A. (Eds.), Disease Control Priorities in Developing Countries, second ed. World Bank, Washington D.C, pp. 1293–1307.
Peabody, J., Shimkhada, R., Quimbo, S., Florentino, J., Bacate, M., McCulloch, C.E., et al., 2011. Financial incentives and measurement improved physicians' quality of care in the philippines. Health Aff. 30 (4), 773–781.
Robyn, P.J., Bärnighausen, T., Souares, A., Traoré, A., Bicaba, B., Sié, A., et al., 2014. Provider payment methods and health worker motivation in community-based health insurance: a mixed-methods study. Soc. Sci. Med. 108 (0), 223–236.
Rowe, A.K., de Savigny, D., Lanata, C.F., Victora, C.G., 2005. How can we achieve and maintain high-quality performance of health workers in low-resource settings? Lancet 366 (9490), 1026–1035.
Sherry, T., Bauhoff, S., Mohanan, M., 2013. Paying for Performance in Health Care: Results from the Randomized Roll-out of Rwanda's National Program. In: Economic Research Initiatives at Duke (ERID) Working Paper, vol. 136, pp. 1–45.
Soeters, R., 2013. PBF in Action Theory and Instruments No. 251113. Cordaid-SINA Health, The Hague.
Soeters, R., Vroeg, P., 2011. Why there is so much enthusiasm for performance-based financing, particularly in developing countries. Bull. World Health Organ. 89 (9), 700.
The World Bank, 2010. Brand New Day: Newly Launched Nationwide Performance-based Financing (PBF) Scheme in Burundi Reflects the Hopes of a Nation.
The World Bank Health Results Innovation Trust Fund, 2013. Results-based Financing for Health. Retrieved 01/09, 2014 from. http://www.rbfhealth.org/project/our-projects.

United Nations, 2013. The Millenium Development Goals Report 2013. United Nations, New York.
Witter, S., Fretheim, A., Kessy, F., Lindahl, A., 2012. Paying for Performance to Improve the Delivery of Health Interventions in Low- and Middle-income Countries (Review), vol. 2. The Cochrance Review, pp. 1–81.
Witter, S., 2012. Health financing in fragile and post-conflict states: what do we know and what are the gaps? Soc. Sci. Med. 75 (12), 2370–2377.
World Bank, 2012. World Development Indicators. The World Bank.
World Health Organization, 2006. Pregnancy, Childbirth, Postpartum and Newborn Care: A Guide for Essential Practice. World Health Organization, Geneva.