ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Mathonnat, Jacky; Pelissier, Aurore

Working Paper

How a Results-Based Financing approach can contribute to the health Sustainable Development Goals - Policy-oriented lessons: what we know, what we need to know and don't yet know

FERDI Working Paper, No. P204

Provided in Cooperation with:
Fondation pour les études et recherches sur le développement international (FERDI), Clermont-Ferrand

Suggested Citation: Mathonnat, Jacky; Pelissier, Aurore (2017) : How a Results-Based Financing approach can contribute to the health Sustainable Development Goals - Policy-oriented lessons: what we know, what we need to know and don't yet know, FERDI Working Paper, No. P204, Fondation pour les études et recherches sur le développement international (FERDI), Clermont-Ferrand

This Version is available at:
https://hdl.handle.net/10419/269485

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

FONDATION POUR LES ÉTUDES ET RECHERCHES SUR LE DÉVELOPPEMENT INTERNATIONAL

Working Paper 204 October 2017 Development Policies

How a Results-Based Financing approach can contribute to the health Sustainable Development Goals

Policy-oriented lessons: what we know, what we need to know and don't yet know

Jacky Mathonnat, Aurore Pélissier

Jacky Mathonnat, CERDI-CNRS, Université Clermont Auvergne and FERDI

Aurore Pélissier, Équipe Économie de la Santé – Laboratoire d'Economie de Dijon, Université de Bourgogne Franche-Comté.

Abstract

The 'results-based financing' (RBF) approach, an umbrella term, occupies a particularly important place in the debates regarding meeting the challenge of 'better health for all at all ages' as stated in the Sustainable Development Goals 3. Our analysis of the RBF schemes results published in the literature underlines five major points. 1. The RBF approach is a very promising one, but highly context dependent. 2. There is no silver bullet. It should be based on a sound realistic theory of change with incentives relevant to the specificity of every context. 3. RBF includes not only financial incentives and can contribute to the reform of the health care system. 4. It appears very important to combine RBF on both the supply side and demand side (as conditional cash transfers, vouchers, etc.) to create positive synergies. 5. There are many gray areas on issues important to decision-makers. More rigorous research (including qualitative) is needed to create valuable public goods on this crucial issue.

Acknowledgements

A preliminary version of this study was presented at the Workshop IDDRI-FERDI 'Transformative investment for sustainable development', Paris, 27 March 2017. The authors thank the participants for their comments. They also thank Jaime de Melo for his suggestions. The authors remain solely responsible for the analysis made in this document.

LA FERDI EST UNE FONDATION RECONNUE D'UTILITÉ PUBLIQUE.
ELLE MET EN ŒUVRE AVEC L'IDDRI L'INITIATIVE POUR LE DÉVELOPPEMENT ET LA GOUVERNANCE MONDIALE (IDGM).
ELLE COORDONNE LE LABEX IDGM+ QUI L'ASSOCIE AU CERDI ET À L'IDDRI.
CETTE PUBLICATION A BÉNÉFICIÉ D'UNE AIDE DE L'ÉTAT FRANÇAIS GÉRÉE PAR L'ANR AU TITRE DU PROGRAMME «INVESTISSEMENTS D'AVENIR» PORTANT LA RÉFÉRENCE «ANR-10-LABX-14-01».

INTRODUCTION 2

1. WHAT IS RBF AND WHY USE IT? 3

1.1. The concept 3

1.2. From results to performance 5

1.3. What types of indicators? 6

2. THE MAIN EFFECTS: POSITIVE RESULTS COEXIST WITH MIXED OR NO RESULTS 7

2.1. A brief overview 7

2.2. Cambodia: the complex effects of a complex architecture during a decade of pbf 9

2.3. Burundi: pbf on a national scale 11

2.4. Afghanistan: p4p in a context of conflict and rebuilding the state 14

2.5. Pbf and the poor in rwanda: 'pro-rich' effects that also benefit the poor 16

2.6. Pbf and improving prescribing practices in rural china 18

2.7. Argentina: rbf in an upper-middle-income country 20

3. GETTING THE BEST FROM RBF AND AVOIDING PITFALLS 21

3.1. One size does not fit everyone: context is crucial 21

3.2. When reported results depend on evaluation methods, the quality of data and information collected 23

3.3. Short and long-term effects, and the sustainability of rbf mechanisms 26

3.4. Externalities, side and net effects 27

3.5. Cost-effectiveness 28

3.6. Community rbf for pushing for better health care: an important but (almost) missing vector in the literature. 29

4. CONCLUSION 30

ANNEX. A selection of studies used for the case-studies, with methodologies and main findings 32

REFERENCES 35

Introduction

The ambition of the Sustainable Development Goals (SDGs) in the field of health is a major challenge for developing countries and their external partners. This is not just a matter of stepping up the resources available for health, but also introducing or developing the use of managerial and organisational approaches as an integral part of reforms likely to further the achievement of the SDGs targeting the health sector. So-called 'results-based financing' (RBF), 'performance-based financing' (PBF), or Paying for Performance (P4P) - in this paper, we use the generic term RBF, but also keep the terms PBF and P4P when these are used in the studies we cite - occupy a particularly important place in the debates on these issues.

While interest in the RBF approach is undoubtedly now shored up by the SDGs' ambition in the field of health, it had already gained ground with the drive towards the Millennium Development Goals, when it became clear that insufficient headway would be made in the field of health if the 'business-as-usual' approach was maintained, with only an increase in available resources. Various studies had sounded the alarm regarding both the poor quality of care frequently observed – which did not forcibly stem from a lack of resources – and the fact that the systemic dynamics did not encourage better use of the scarce available resources. As a result, health indicators were lastingly doomed to fall short of what they could have been. The RBF approach provides a response – but only one among others; it is not a panacea, but is being implemented with a growing interest in view of meeting these enduring challenges and helping speed up the advance towards 'healthy lives…for all at all ages', as stated in Goal 3 of the SDGs.

Moreover, the RBF approach is explicitly integrated in the principles of the 2005 Paris Declaration on Aid Effectiveness and the Accra Agenda for Action (one of the goals being 'Managing for Results: Managing resources and improving decision-making for results'[1]), and in line with the 2015 Addis Ababa Action Agenda on financing for development.[2]

The aim of this study is to summarise the main effects of the RBF approach in developing countries as seen in the literature and to draw lessons that are directly useful for the effectiveness and efficiency of health policies. *Overall, this approach turns out to be very promising*. It shows many positive effects even though, at the end of the day, the results seem mixed and leave many different gray areas, as some programmes/projects have not produced the expected results or have generated negative effects. As promising as it may be, the RBF approach thus should be considered with a reasoned pragmatism, in order to better contribute to achieving the SDGs in the health field. It should be viewed as one tool in the set of health financing tools now available.

[1] See notably Art. 14 'Partner countries commit to…exercise leadership…in developing and implementing their national development strategies through broad consultative processes; Translate these national development strategies into prioritised results-oriented operational programmes…'; cf. also articles 43 to 46.

[2] Art. 50 : 'We recognize that we share common goals and common ambitions to strengthen international development cooperation and maximize its effectiveness, transparency, impact and results'. United Nations, Outcome document of the Third International Conference on Financing for Development: the Addis-Ababa Action Agenda, 2015.

This paper is organised as follows. The first section describes what the RBF approach involves (1). The main effects observed, as they are documented in the literature, are then summarised and followed by examples of RBF implemented in six very different countries (Afghanistan, Argentina, Cambodia, China, Burundi and Rwanda), but each one has features pointing to a specific interest for setting up strategies that include RBF approaches (2). While the rationale behind RBF approaches is in fact quite simple, its operationalisation is complex and there are pitfalls that must be avoided. This is what section 3 aims to show, along with highlighting grey areas that remain regarding important issues for decision-makers (3). The conclusion (4) briefly highlights the main orientations that emerge from the literature in a 'policy-oriented' perspective.

1. What is RBF and why use it?

1.1. The concept

Both economic theory and common sense support the idea that payment for healthcare, regardless of the source of funding (budget, insurance, public or private foreign aid, user-fees) should not be disconnected from meaningful indicators of quantity, quality or value, as well as without losing sight of the fundamental issue of efficiency. It therefore makes sense to tie a *fraction* of providers' payments (public or private staff and public or private health facilities) to their results or 'performance'. Improving results or 'performance' in healthcare delivery requires, on the one hand, changes in providers' behaviour and, on the other hand, that relevant incentives be put in place, given that many current based payment methods do not explicitly stimulate good performance.

In the broadest sense of the term, RBF can be defined as a payment made to a national or sub-national government body, to a health facility or any other healthcare provider, or even to a consumer of health services,[3] once predefined results have been attained and verified. This concept encompasses diverse approaches and modalities in view of its operationalisation, but all of them share three common features. First, they differ from the conventional approach to financing in terms of inputs, whether this be through a budgetary line or one-off payment.[4] This approach is heavily criticised for its lack of flexibility, effectiveness and efficiency. Secondly, the payments are conditional and, lastly, they contain an incentive-based aspect designed to influence the behaviour of the beneficiaries (cf. Figure 1 below).

That said, RBF should not be viewed as a financing tool limited to the supply and demand of healthcare (Ireland et al., 2011). The RBF approach can be embedded and used *as a strategic tool* at multiple levels in the architectures designed to reform the healthcare system. It raises questions on the grounds for and the modalities of decision-making within the concerned institutions and, on top of that, within the organisational structure of the health pyramid (management autonomy,

[3] Using voucher or conditional cash transfer schemes, for example. But this paper is focussed on supply-side schemes.

[4] Although in some cases the set objectives require an increase in inputs.

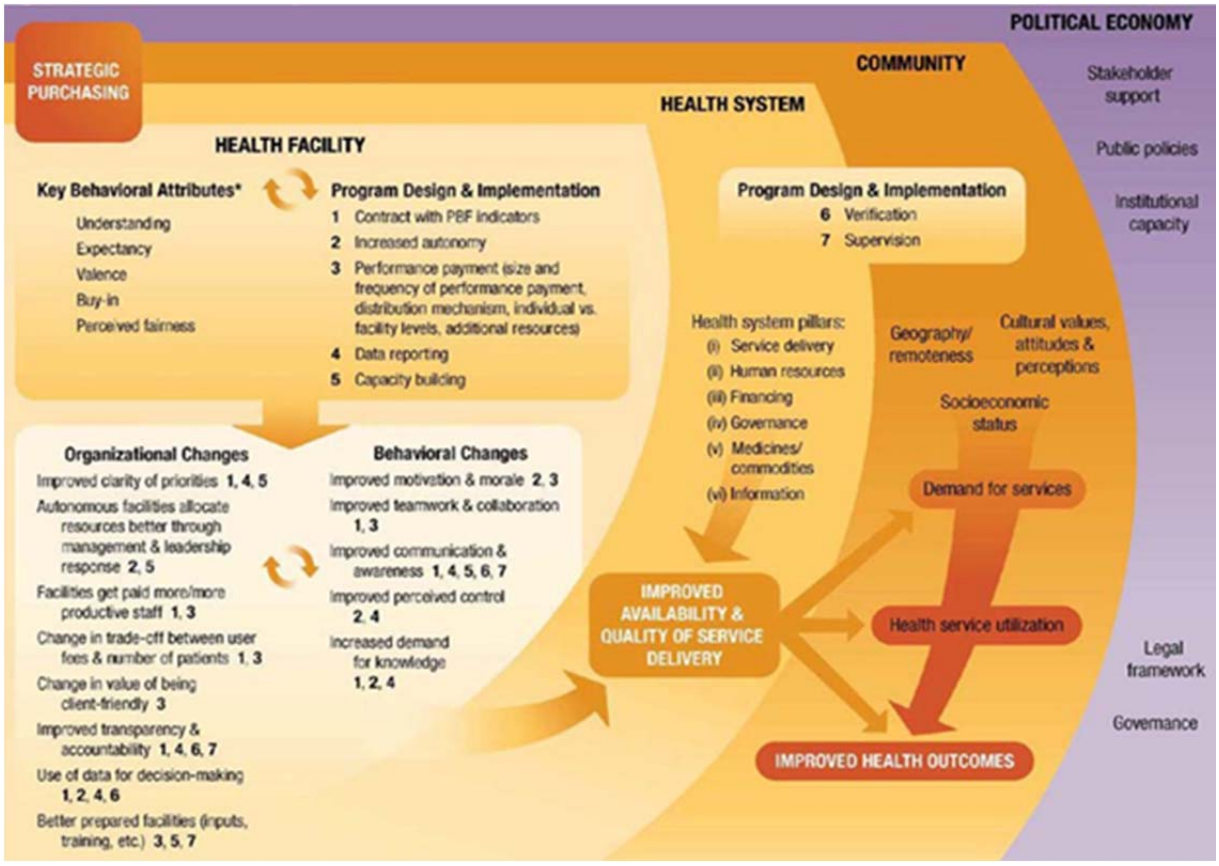
separation of the functions between buyers of health services, providers, verification, participation of the different stakeholders in the decision-making process, production and structuring of information, accountability, etc.).

The RBF concept is close to other approaches such as Cash On Delivery (COD) promoted as a new approach to aid by the Center for Global Development. It is also used in aid for trade (Cadot and de Melo, 2014). The main difference with the RBF mechanisms for healthcare is that they give complete autonomy to agents regarding the decisions and methods used to achieve the contractual results, contrary to what is generally found in RBF projects. For example, in the RBF project in Zimbabwe supported by Cordaid and the World Bank, the use of health facility Operational Plans (OPs), also known as Business Plans, to guide the use of RBF payments is a cornerstone of RBF implementation at the front-line service delivery level. The supporting supervision included in the project is a morale booster for health workers and promotes relationship-building between health providers and district or provincial health supervisors.

In terms of payment-for-performance, the approach is largely based on the lessons of agency theory, behavioural economics, contract theory and what are dubbed 'new public management' models. For much of its justification, it relies on the fact that agency theory shows that the objectives of agents (e.g., healthcare providers) do not necessarily coincide with the principal's (Ministry of Health and donors). It also holds that an asymmetry of information to the disadvantage of the principal renders the agents' efforts mostly unobservable, which enables these agents to develop strategies that may not head in the same direction as the objectives targeted by the principal. There is also information asymmetry between doctor and patient, as the patient cannot observe the doctor's efforts or assess the quality of his/her care. The only observation he/she can make is how his/her state of health evolves. This double asymmetry reinforces the relevance of incentives targeting the provider. The RBF also supposes that a theory of change is identified for each project, making it possible to conceptually establish a chain of precise relationships between the payment-related incentive[5] and the expected results.

[5] The situation is more complicated when health care is paid by a third party. This is going to be increasingly common in developing countries with the goal endorsed by the SDGs to accelerate the development of health insurance. In that case, it can be more difficult to align *the whole set* of incentives because the three parties (providers, patients, and insurance) may have different objectives and different levels of information – a classic multiple agent problem. This reinforces the need to have – above the health system itself – a pilot on the plane with a clear vision of the objectives to be achieved and the trajectory to be followed, which is not always the case.

Figure 1: From incentives to results



From Cataldo and Kielmann (2016).

Perakis and Savedoff (2015) distinguish four kinds of channels which should be considered in sound theories of change in order to incorporate the expected effects of RBFs programmes: (1) the offer of a financial incentive will lead to some behavioural change by the recipient; (2) the performance funding makes results visible in a way that improves management; (3) the focus on results will improve accountability to constituents or beneficiaries; and (4) the agreement may give recipients more discretion and autonomy to innovate and adapt their activities. Perakis and Savedoff note that RBF programmes are frequently criticised for relying on the financial incentive to change recipient behaviour when most are also actually designed to work through one of the other three channels.

1.2. From results to performance

The notion of results may be understood in terms of the production of health services or the implementation of processes, of intermediate or end results, or of impacts. That said, the term RBF is sometimes clouded by a semantic ambiguity that is a source of confusion for defining the targeted objectives. In fact, the expression 'performance-based financing' (PBF) is quite often regarded as a close synonym of the umbrella term 'RBF'. Yet, they need to be differentiated as the notion of 'performance' relates to the idea of a gap between what is potentially achievable and what is actually achieved; or even to a notion of effort, taking into account a given environment in

which a basic distinction needs to be made between endogenous elements that fall within the sphere of action of the stakeholders directly involved in the RBF scheme, and the exogenous elements that are independent of the stakeholders' behaviour, at least in the short or medium term. In this sense, a good result achieved with respect to a predefined objective does not necessarily constitute a good performance. If one considers the issue of efficiency to be one of the key issues of healthcare due to the high resource constraints, the question of incentivising *performance* (not only results) by adopting the appropriate financing mechanisms must be given the utmost attention.

Among high-income countries, the United Kingdom is a fine example of this useful distinction. In an effort to improve transparency and performance in primary care, the UK government introduced the Quality and Outcomes Framework (QOF) in 2004 for surveillance of patients with chronic diseases. Performance was measured through indicators relating to successful regular checks and control of medical conditions such as high blood pressure and diabetes. Performance was rewarded with points converted into monetary value for doctors whose practices yielded good results on the indicators. However, studies have shown that some doctors were already achieving the QOF outcome for certain indicators (e.g., blood pressure level) even before the scheme was introduced, meaning that they were rewarded for what they were already doing (Maynard, 2012), not for improving their performance. Moreover, understanding why some doctors had reached the objective before the incentive was introduced, while others did not, has not been elucidated. This possible discrepancy between results and performance creates an ambiguity with respect to the targeted objectives when designing the incentive.

1.3. What types of indicators?

The main indicators targeted by RBF interventions focus more on what are dubbed 'intermediate results', mainly outputs (number of vaccines, improved access to and quality of healthcare, an increase in some types of activity, etc.), rather than on so-called 'end results' or, in other words, health status outcomes[6] (mortality, morbidity) and impacts,[7] or in other words, gains in improved quality-adjusted life-years (QALYs) or disability-adjusted life-years (DALYs) for instance.

The question is under debate, but it must be kept in mind that it is not always relevant to target an incentive based on health status indicators, as these generally depend on multiple factors, including many that are largely beyond the control of the providers involved in RBF schemes – as these schemes are not disconnected from their environment. These factors include those that depend on the healthcare system, which is external to the health facilities participating in RBF

[6] The term refers to the impact healthcare activities have on people – on their symptoms, the ability to do what they want to do and, ultimately, on whether they live or die. Health outcomes include whether a given disease process becomes better or worse, what the costs of care are, and how satisfied patients are with the care they receive. It focusses not on *what is done* for patients but *what results* from what is done.

[7] Measure the quality and quantity of long-term results generated by programme outputs (e.g., measurable change in quality of life, reduced incidence of diseases, etc.).

schemes, and factors that rest on individual behaviours and which impact the demand for healthcare.

This last point highlights the utility of coupling RBF initiatives concerning healthcare supply with schemes targeting healthcare demand (see the example of Cambodia below; cf. also Kandpal, 2016) such as conditional cash transfers (CCT). This coupling creates synergies, as the information asymmetry between providers and patients suggests that what is needed on the supply side to improve health status (e.g., better quality of care with evidence-based protocols) is not necessarily sufficient to trigger an increase in the demand for healthcare. Indeed, several studies have shown that, once again due to information asymmetry, patient behaviour in developing countries is determined more by the perceived than the objective quality of care. The means that it is advisable to align healthcare supply incentives with demand incentives: more patients will then visit health facilities and benefit from adequate care, which should logically lead to improved health status.

2. The main effects: positive results coexist with mixed or no results

2.1. A brief overview

Intermediate and health status indicators. Overall, the literature highlights that RBF schemes quite often have various positive effects/impacts on intermediate results (inputs, outputs and, less often, outcomes), although many mixed results are reported in the literature. *But it is important to bear in mind that not all studies have the same methodological robustness (measurement of effects and consideration of confounding factors).* RBF schemes improve access to and use of health facilities, antenatal and childbirth care dispensed by qualified staff in health facilities, the promotion of family planning activities, the diagnosis and treatment of malaria, HIV and tuberculosis, and also the implementation of protocols to improve the quality of care (cf. Table 1 for a selection of key indicators targeted by RBF programmes; see also Kalk et al., 2010; Van Herck et al., 2010; Basinga et al., 2011; HRTIF, 2013; Miller and Babiarz, 2013; Bonfrer et al., 2014 (a,b); Gertler and Giovagnoli, 2014; Chimhutu et al., 2015; Norad, 2015; Kandpal, 2016; Lannes et al., 2015; Nimpagaritse et al., 2016; Ogundeji et al., 2016 (a,b); Sun et al., 2016, Renmans et al., 2016; Witter et al., 2012).

Yet, it is important to note that, most of the time, these effects are not documented enough. For the most part, they are medium-term effects observed by short-period analyses, which may be heightened by the effects of experience or adversely affected by unexpected and undesirable effects. Regarding the effects of RBF on health indicators (mortality, morbidity, QUALYs-DALYs, etc.), the vast majority of studies conclude that there is no effect. In general, information to clearly understand why the scheme worked somewhat, did not or poorly worked, is often unavailable.

This global picture for developing countries is in line with what Eijkenaar et al. (2013) and Cashin et al. (2014) pointed out for a selection of industrialised countries (along with Argentina and Taiwan), and with a recent OECD study (Srivastava et al., 2016): 'The popularity of P4P schemes in OECD countries continues to grow in primary care, specialist care as well as in hospitals, although there is

still a scarcity of clear evidence on the success – or otherwise – of P4P programmes. Systematic reviews of available evidence tentatively suggest a positive impact on performance, but evidence on the impact of P4P on health outcomes remains inconclusive and limited. While improvements on some indicators in some P4P schemes are found, no clear "breakthrough" in performance improvement can be clearly linked to the introduction of a P4P scheme' (Srivastava et al., 2016, p. 78).

Equity and pro-poor effects. There is little evidence to show whether RBF programmes have an impact on equity and are pro-poor or not. In fact, two questions are involved here, and should not be confused although they are more or less amalgamated in studies on equity. One question is indeed whether or not RBF benefits the poor. Another question is whether it benefits the better-off more than others. For these two questions, available examples show opposite findings and do not point to any compelling conclusion.

In what contexts? There is also limited evidence indicating which conditions are most favourable to RBF schemes, including the mechanisms and channels through which incentives work, cost-effectiveness, comparison with other potential approaches, sustainability and unintended consequences. But as will be seen later, the literature makes it possible to identify a set of benchmarks (or broad guidelines) that indicate the way forward, identify the points where vigilance is required if the best results are to be achieved with RBF programmes, and avoid potential pitfalls, although more research is needed.

Table 1 – Summary of the main targets included in RBF programmes

	Interim results	**Short-term results**
Children (particularly infants and children under 5)	• Number of postnatal visits • Vaccination	• Decline in infant and child mortality rates • Increase in immunisation rates • Immunisation schedule completed
Women (particularly pregnant women)	• Family planning: contraception • Care for pregnant women: tetanus vaccination and antenatal visits • Care for childbirth: caesarean deliveries, childbirth in health facilities, qualified medical staff	• Fertility • Decline in maternal mortality rates • Preterm delivery
HIV/Malaria/Tuberculosis	• Prevention: number of tests carried out, awareness-raising during consultation, insecticides • Diagnosis: number of cases diagnosed • Treatment: number of treatments delivered	• Incidence rates • Prevalence rates • Survival rates

Source: Authors.

We turn now to six case-studies to provide a better understanding of what we summarised above. We present the main features, the goals of the schemes and their main results. Also refer to the annexed table for more details and the characteristics of the methods used.

2.2. Cambodia: the complex effects of a complex architecture during a decade of PBF

The case of Cambodia is particularly interesting insofar as the country has set up diverse PBF systems starting in 1999. These have been followed attentively through monitoring and evaluation. As a result, highly useful elements have been identified, making it possible to assess the effects of different incentive packages and contractual arrangements (Van de Poel et al, 2016). Here, we present the most striking aspects.

The scheme

The pilot phase spans the period 1999–2003. There are two types of schemes. The first (Pilot-Out) handed management authority to NGOs at the operational district (OD) level, and NGOs had complete discretion to determine the nature and level of incentives for health facilities and their staff. The ODs also benefited from incentives, but in the form of fixed-price contracts for fixed performance objectives. The second scheme (Pilot-In) was identical to the first, except that the NGOs work with staff from the Ministry of Health and follow the ministry's procurement procedures.

The PBF programme was expanded in 2004. It combined three types of schemes. The first (2004–2008) dubbed 'IN', took up the features of Pilot-IN and involved 11 ODs. The second scheme (2005–2010) relied on 'internal contracting' in eight ODs. Contrary to Pilot-In, management was not handled by NGOs but by the Ministry of Health structures involved (provincial health department, the OD administration, the reference hospitals and health centres). The NGOs intervened as advisors. The performance contract signed with the Ministry of Health defined the objectives for the different administrative levels mentioned above, but the incentive payment system was set up only at the level of the health facilities. Bonuses were foreseen, backed by a fee-per-case system depending on whether the objective had been reached for childbirth and antenatal visits to the participating health facilities. Incentives were also introduced to encourage the abandonment of illegal practices. This was rounded off by Equity Funds to finance schemes designed to cover user fees and make access to healthcare facilities easier for the poor (Flores et al., 2013). A supplementary scheme is in place in partnership with GAVI as of 2007 in 10 ODs. The GAVI/RBF scheme is directly managed by the Ministry of Health at the level of health facilities. There was no incentive at the district level, but staff received a fee-per-service.

These incentive mechanisms thus had very different forms, which makes evaluating them all the more complex.

The results

Whatever the scheme introduced between 1999 and 2010, and after controlling for different sources of heterogeneity, it appears that the PBFs – all categories combined – overall led to a 6–7.5% increase (depending on how this is calculated) in the probability of a child being born in a public health facility, instead of at home. On the other hand, there was no significant effect on vaccination rates or antenatal care. The PBFs increased the probability of the non-poor giving birth in a health facility by 13.4%, but there was no significant impact for the poorest women (Van de Poel et al., 2016).

The results highlight the synergy between PBF schemes targeting both supply and demand, which seems highly salient for the effectiveness of incentive mechanisms. In the districts where a system of vouchers for delivery in a health facility was introduced in tandem with a PBF scheme, the probability of giving birth in a health facility increased by 26%, whereas the impact was weak when there was no demand-side incentive mechanism.

However, the analyses showed that a substantial share of these overall results (poor and non-poor) is not consistent with a *net* increase of deliveries in health facilities, but comes rather from a shift in demand from private to public providers.

The study also raised the question of whether the effects differed depending on the characteristics of the RBF schemes set up. The analyses showed that the two pilot experiments introduced in 1999 and 2003 (Pilot-In and Pilot-Out) had significant effects on the probability of giving birth in public health facilities. This probability was higher for the Pilot-Out schemes: +16.4% compared to 9.3% for Pilot-In, although this growth was to the detriment of private providers.

As for the effects of the second wave of PBF schemes set up in 2004, it seems that internal contracting had a much greater effect than the IN scheme. The explanation put forward by the studies is that internal contracting offered a substantial payment for each delivery in a health facility, as the IN model provided the districts with an objective based on the number of deliveries to be reached, but it was up to the districts themselves to set the incentives implemented at the level of the health facilities, which implicitly proved to be inadequate.

On the other hand, the IN model was the only one to have a positive impact on antenatal visits, which increased by almost 10% between 2004 and 2008.

The only PBF scheme to have a positive impact on the probability of a woman having two antenatal visits was the one incorporating a contracting mechanism in which district staff comprised government employees, nonetheless supported by NGOs under contract with the Ministry of Health.[8] These very mixed results could be explained by the fact that the marginal cost required to convince pregnant women to visit health facilities for regular check-ups is high compared to the low level of incentives given to staff to achieve this objective.

[8] Which plays a key role in management (the IN model) as pointed out earlier.

Whatever the PBF scheme rolled out by GAVI over the entire period studied, the analyses did not reveal any significant effect on vaccination rates. This is likely due to the fact that, before the introduction of the incentive mechanisms, 80% of mothers in the areas concerned already had at least two antenatal visits and two-thirds of the children were fully vaccinated.

Lastly, the studies revealed no significant effect of PBF schemes on neonatal mortality. On this point, similar results were found for Rwanda (Chari and Okeke, 2014; Lannes et al., 2016).

2.3. Burundi: PBF on a national scale

Burundi, a small land-locked country prone to severe political instability, has some of the world's weakest health indicators. It suffers from particularly strong constraints due to its low level of national human, financial and technical resources. Yet, Burundi is one of the few countries to have gradually set up PBF as the core of its health development[9] under its national health strategy (National Health Development Plan 2006–2010 and 2011–2015) and the achievement of the SDGs (Busogoro and Beith, 2010).

The scheme

Financing health by developing PBF has been part of a proactive national strategy since 2010. After a pilot phase initially launched in three provinces in 2006, Burundi gradually extended PBF until the scheme was scaled up to the whole country in 2010.[10]

Today, PBF involves all public health centres and hospitals, as well as many private non-profit health structures (i.e., a total of 565 health centres, 45 district hospitals and 5 national hospitals, as well as 151 private-sector providers; Renaud, 2013). Funding is mainly provided by the Government and the World Bank. The scheme is managed by the 'PBF technical unit' under the General Directorate of Public Health and the Fight against AIDS. The mechanism exclusively targets healthcare supply and pays subsidies for the achievement of quantitative objectives based on the volume of activity and rate of cover (measured by monthly evaluations), and qualitative objectives based on the quality of care. The quality of care is viewed from a double perspective: a 'technical' view based on whether the health facilities comply with standards (through a quarterly evaluation) and a 'subjective' view based on a patient survey platform (every six months; Renaud, 2013). The patient survey checks three points: i) 'patient existence', which checks that the patient in the health facility's registers really exists; ii) 'health care confirmation' reporting the services actually used by the patient; and iii) 'patient satisfaction' with the health services delivered (Renaud, 2013). The

[9] Note that at the outset, in 2006, Burundi removed the fee-for-services for deliveries and health services for children under five. In 2009, this measure was extended to pregnant women or those who had just given birth. This free care drove an increase in the number of visits to health facilities, which exacerbated supply-side constraints (shortcomings of equipment, availability of drugs, and skills and motivations of the health workers). This situation justified the introduction of supply-side incentives to fund these health services (Busogoro and Beth, 2010; Seheye et al., 2015). These incentives had already been tested since 2006 by three provinces, with funding provided by Dutch NGOs (Renaud, 2013). Burundi thus combines demand-side incentives with free care and supply-side incentives by financing free services through a PBF programme.

[10] More details on the implementation of PBF between 2006 and 2010 can be found in Bugoroso and Beith (2010).

quantitative objectives are remunerated on a monthly basis. The qualitative objectives are remunerated on a quarterly basis: a bonus (calculated on the subsidy paid for achievement of the quantitative objectives – up to 25% of the amount) is awarded if the quality score is above or equal to 70% and a penalty may be imposed if it is below 50% (Bonfrer et al., 2014b; Bugoroso and Beith, 2010; Renaud, 2013). The RBF payment comes on top of the budget of a health facility (direct payment to patients + government subsidies + finance from external sources). It thus accounts for 30–35% of the hospitals' total resources and 80–85% of total resources for health centres. PBF funds first and foremost the expenses and cash requirements of the health facilities. The remaining budget can then be allocated to health workers as bonuses: in 2010 these represented one-third of their base salary (Renaud, 2013).

Initially, the purpose of PBF was to fund a package of minimum services in health centres and additional services in hospitals, with special focus on maternal and infant health in line with the Millennium Development Goals (MDGs) 4 and 5 (Bugoroso and Beith, 2010). Since 2013, the programme has been expanded to include the goals of preventing and treating malnutrition – with a specific focus on children under five (Nimpagaritse et al., 2016). Three tiers of the healthcare pyramid are involved, intervening at different stages to prevent and treat malnutrition. First, the community health workers responsible for prevention, then the health centres that also relay elements of prevention, follow the children, detect signs of malnutrition, treat them (in the early stages to severe cases without complications) or refer them to a hospital in severe cases with complications (Nimpagaritse et al., 2016).

The results

The objective of the pilot phase was to improve maternal and infant health. An initial study by Bonfrer et al. (2014a) involving provinces funded before the scheme was scaled up to national level (9 provinces out of 17 studied) showed that PBF helped to achieve a 36% increase in the share of women receiving delivery assistance in a health facility and a 55% rise in the share of women using modern contraception techniques. The quality score showed a significant 45% increase, although this increase had no repercussion on user perception. No effect on equity of access to healthcare was noted by the authors (Bonfrer et al., 2014a).

A second study by Bonfrer et al. (2014b) on the national rollout of PBF confirmed the trend of these results but also revealed differentiated effects on the poor and non-poor.[11] The authors showed that PBF had no significant effect on the probability of a woman having more than one antenatal care (ANC) visit or the probability that this ANC visit occurs during the first six months of pregnancy. On the other hand, the probability that a pregnant woman will have her blood pressure taken and the probability of her being vaccinated against tetanus did increase. The volume of antenatal consultations was not positively impacted, but according to the authors PBF did enable more detailed check-ups for pregnant women and thus an improved quality of care due to this

[11] Trends that are also visible notably in the works of Falisse et al. (2014) and in the report by Seheye et al. (2015).

more comprehensive approach. These positive effects were considerably higher for non-poor pregnant women than for poor pregnant women. Moreover, the authors showed that PBF only had a significant effect on deliveries in health facilities for non-poor women. As for children, the study showed that PBF had a significant positive impact on the probability of their being vaccinated. This effect was stronger for poor children than non-poor children. Overall, the effects identified in the countrywide study in Burundi (Bonfrer et al., 2014b) were weaker than those found in the nine provinces (Bonfrer et al., 2014a).

These results spur various comments, which are discussed notably in Bonfrer et al. (2014b):

- PBF clearly had an effect that was only significant for the variables that could be directly 'activated' by the providers (this concurs with the study results reported for Afghanistan; cf. below). Such findings are frequent in the literature. Supply-side incentives can only be expected to change providers' behaviour and thus impact the dimensions under their control, such as check-ups and interventions performed during consultation (blood pressure taken, vaccination, etc.). As the authors point out, PBF had no effect on the demand-side variables such as the fact of having the first ANC visit during the first six months of pregnancy: the decision to consult varies in line with demand-side factors and notably with the financial and non-financial barriers to accessing healthcare.

- Secondly, while disparities appeared in favour of non-poor women, the results nonetheless raise the question of the barriers encountered by poor pregnant women in accessing health centres (as in Afghanistan), even though healthcare is free (transport costs for example).

The evaluation of the malnutrition component of PBF is in progress via a country-scale randomised experiment carried out by Nimpagaritse et al. (2016).

Finally, Renaud's (2013) report on performance verification in health facilities participating in PBF schemes led the author to draw several interesting conclusions on the system's functioning and the remuneration paid:

- Between January and August 2012, 69% of cases for health centres and 62% of cases for hospitals showed no discrepancies between declared and verified data at the level of quantitative data. When the data failed to match, they were mostly overestimated: 19% on average for health centres and 13% for hospitals. According to the author, these discrepancies were on the whole not due to fraud, but usually to miscalculations or a failure to match the exact definitions of the indicators used for the PBF.

- Between 2010 and 2012, the so-called 'technical' quality scores reached an average of 70% for health centres, 77% for district hospitals and 74% for national hospitals. While these scores appear to be good and seem to point to the effort made by the health facilities to comply with standards, the author draws attention to potential 'gaming' effects due to the modalities for implementing evaluations and to the definition chosen for technical quality.

- Between 2010 and 2012, the so-called 'subjective' quality scores increased for health centres (rising from under 50% on average to over 75%) and for hospitals (rising from less than 40% on average to over 60%). The detection of 'phantom' patients is very difficult, making it hard to monitor subjective quality.

- Sanctions are few and far between.

2.4. Afghanistan: P4P in a context of conflict and rebuilding the State

The situation in Afghanistan is one of the most precarious in the world. The health inputs (staff, facilities and equipment) are acutely inadequate (particularly in rural areas) and of poor quality. As a result, the health indicators are extremely weak, though they have shown some improvement over the last ten years or so. The armed conflict between the Government and the Taliban have exacerbated the situation, which makes the efforts to reconstruct the health system a particularly tricky task. The PBF programme is part of this effort.

The scheme

Many contracting-based experiments including incentive mechanisms, which we will not go into here, have been run by donors (USAID, World Bank and EU), with support from NGOs and Afghanistan's Ministry of Public Health. These experiments sometimes incorporate performance incentives (for more details see Sondorp et al., 2009; Arur et al., 2010; Alonge et al., 2015). The case study aims to give a more detailed presentation of the P4P experiment implemented in 2010 to improve the quality of maternal and child health (MCH) services. The study draws on an evaluation of a cluster-randomised trial between September 2010 and December 2012, carried out by Engineer et al. (2016).

Focusing on the 'health worker motivation' component to leverage improvement of maternal and child health, the P4P scheme was designed only for health providers. The purpose of the financing was to provide bonuses to health workers – through the NGOs managing the health facilities – in order to incentivise the service providers' extrinsic motivations. The scheme had two types of objectives: quantitative objectives to improve cover, but also qualitative objectives to improve the quality and equity of care. Payments were made on a twofold basis: a first payment was made quarterly to the health facilities depending on their achievement of nine pre-defined objectives related to MCH (especially to MDGs 4 and 5); a second payment was made annually via an additional bonus system based on performance in terms of i) equity (measured by a concentration index of assisted deliveries in health facilities and a concentration index of outpatient visits of children under five, ii) quality of care (measured by a balance scorecard based on 20 indicators proxying multiple aspects of the quality of care), and iii) contraceptive prevalence in the health facility's catchment area. The payment was adjusted through negotiations between the NGOs and the Ministry of Public Health depending on the context in which the health provider was operating (baseline conditions, expected improvements and geographical accessibility). The health facility

managers distributed the P4P payments among the healthcare workers at their own discretion (amounting to 14–28% of their base salary in 2011; Engineer et al., 2016).

The results

The evaluation conducted by Engineer et al. (2016) covered 442 health facilities spread over 11 provinces (out of the country's total 34 provinces). The study used a randomised approach to compare a group of intervention sites benefiting from P4P with a group of comparison sites that did not benefit. The authors then calculated what the impact analysis literature calls the intention-to-treat (ITT) to estimate the impact of P4P.

The evaluation showed that for the target P4P indicators of MCH service coverage (intermediate outputs for antenatal care and skilled birth attendance, family planning and vaccination) and equity there was no statistical difference between the intervention and comparison sites: P4P had no impact. Nevertheless, analysis of the concentration indicator of skilled birth attendance in a health facility yielded a higher probability of non-poor women giving birth in a health facility compared to poor women: a result that does not point to improved equity and which also raises the issue of what barriers are preventing poor women from accessing health facilities for assisted delivery. This is not an isolated effect in the literature, as shown by the studies on Burundi presented in this paper. Moreover, a positive P4P impact on healthcare quality was identified by the authors for three aspects of quality: spending more time with the patients, conducting more complete examinations, and counselling. Here again, the results concur with those found for Burundi: the performance incentives had a positive impact on the completeness of visits, rather than on an increase in the actual number of visits. On the other hand, the other dimensions measured on the balance scorecard (health worker motivation and job satisfaction, client satisfaction, perceived quality of care, community involvement) were not impacted.

The findings of the evaluation are thus not up to what was expected from the programme. The design of the incentive itself was called into question by the authors. The study revealed that there were barriers to fully implementing the role of payment as a driver of change in clients' health-seeking behaviour: delays in setting up payments at the beginning of the programme, shortcomings in communication on the payment and in how it is related to performance, as well as the payment modalities (into a bank account rather than cash), and lastly an amount seen as insufficient. Moreover, the variables that also depend on demand, such as the fact of going to a consultation in a health facility, cannot be changed by a supply-side action alone, as we underlined in the PBF evaluations for Burundi. Demand-side as well as supply-side actions are necessary if the coverage of health services is to expand. By combining supply- and demand-side incentives, headway can be made in changing behaviour regarding utilisation of services, notably access, and in improving patient care, be it terms of the quality of care or the comprehensiveness of the care pathway, and ultimately in impacting the final health outputs, which is also shown by the literature on other countries (cf. Cambodia and Burundi; cf. also Kandpal, 2016). The results of this experiment thus re-position the question of incentives. Are roadblocks only found on the supply

side? Can the roadblocks be removed through financial incentives alone: more specifically, what about the intrinsic motivations of health workers?

2.5. PBF and the poor in Rwanda: 'pro-rich' effects that also benefit the poor

The effects on equity and more particularly on the poor are one of the lesser documented fields in the literature on the effects of PBF-type mechanisms. On this count, the experience of Rwanda is interesting in several respects.

The scheme

Rwanda is one of the few countries in the world to have gradually rolled out a very large-scale PBF policy, beginning with its pilot projects in 2001. The good results of these pilots spurred the Government to extend them countrywide in 2006 and in parallel prepare a rigorous evaluation process. At the same time, the Government engaged in a fast-track policy to develop health insurance, which yielded figures in terms of population coverage relatively comparable to those in China. The coverage rate rose from 7% of the population in 2003 to around 85% in 2008, reflecting a very sharp increase in healthcare access and a decrease in the catastrophic expenditure on health. Additionally, a strong focus was placed on progress in health workers' skill levels, the Ministry of Health's leadership and global governance.

Rwanda's PBF was designed to incentivise health facility staff to increase access to their services, strengthen health worker productivity, and improve service quality. PBF contracting with the health facilities allowed the local health authorities to distribute these supplemental performance funds in line with local priorities; typically, provider bonuses or facility supplies and equipment (Skiles et al., 2015). Special focus was placed on the quality of care, which had to comply with the norms and protocols defined at the national level. Various studies have shown the positive effects of this approach, be it in the area of healthcare access overall or for more specific services such as increased HIV testing and counselling services (de Walque et al., 2015).

A study by Lannes et al. (2016) specifically explored the role of PBF in improving the poor's access to maternal and infant healthcare services in rural areas. The authors used a randomised analysis based on a protocol that makes it possible to control for the observed differences between the district groups where PBF was rolled out and the other groups being attributable solely to the incentive mechanisms implemented. The study covered the period 2006–2008. The authors distinguished two groups of households based on a poverty profile determined by analysing the main poverty components. The households below and above the median are seen as belonging to a lower and upper group ('poor' and 'better-off') respectively.

The results

The results were robust to different methodological specifications. They showed that PBF increased the number of deliveries in health facilities for the upper group but not for the lower group, a result

also found in the cases of Burundi and Afghanistan described above. No effect was observed for antenatal visits. However, PBF did bring about an increase in deliveries in health facilities for women from the 'better-off' *and* the 'poor' groups, provided they benefited from demand-side incentive mechanisms – health insurance in the case of Rwanda (in Burundi, free healthcare) – which is a particularly salient result that confirms a finding now emerging in the literature.

PBF had a positive impact on the use of modern contraception for the upper group, but negative for the lower group. On the other hand, a positive impact was observed for both groups when it comes to the probability that children will benefit from curative and preventive care.

A previous study (Skiles et al., 2013), based on data from the National Institute of Statistics had found that PBF in Rwanda was neutral with respect to healthcare access, meaning it was neither pro-rich nor pro-poor. While the study of Lannes et al. showed that the upper group benefited from the PBF dynamic more than the lower group – which, as its authors underlined, means that it was pro-rich and had a negative impact on equity given that the rich-poor gap in fact widened – *it nonetheless had positive effects on the lower group* regarding deliveries in health facilities and access to curative and preventive care. Yet Skiles et al. (2013) 'consider that the Rwandan PBF could not be "an effective pro-poor strategy" by referring to the lack of resources of health facilities in poorer communities and consequently the lower responsiveness to the needs of poor people and the inability of the Rwandan PBF set-up to respond to this', as pointed out by Renmans et al. (2016, p. 1304). Skiles et al. (2015) found that children living in PBF districts were more likely to receive medication when seeking facility care for illness in 2008 relative to children living in comparison districts, although this finding was not statistically significant. They emphasise that this finding masks the heterogeneous effects of poverty. The poorest children in PBF districts benefitted more compared to the non-poor in PBF districts and to those living in the comparison districts. The estimated effect of PBF on receipt of treatment for poor children was 45 percentage points higher (and significant) compared to the non-poor children seeking care for diarrhoea or fever. They concluded that PBF improved the quality of treatment received by poor children, conditional on the patients seeking care, but it did not impact the propensity to seek care per se. Certainly, PBF effects on poverty are complex and depend on which indicator is being considered. A recent study by Flink et al. (2016) in Cameroon showed that households registered as indigents judged that PBF (the pilot project in the diocese of Maroua in North Cameroon) had facilitated their access to care and also reduced their financial expenditure. On the other hand, the mechanism had no effect on the poor who were not registered as indigents. Moreover, the study gave no information as to whether the RBF project had an impact on equity. Over and above these two examples, the effects of RBF approaches on the poor remain very mixed and insufficiently researched, and the issues of analysing RBF effects on the poor and on equity should not be amalgamated.

2.6. PBF and improving prescribing practices in rural China

Since the early 2000s, China has engaged in sweeping reforms to its healthcare system, particularly regarding insurance, the regulation of health facilities and improvements in the quality of care, while at the same time it has sought to curtail the high increase in costs (Mathonnat et al., 2015). Health providers are traditionally paid through fee-for-service, together with incentives for health facilities and their health workers. Yet, these incentives have perverse effects that encourage overprescribing medication. Several studies have also highlighted, in China as elsewhere, the often poor quality of care, coupled with inappropriate prescribing practices.

The scheme

The national policy for essential drugs rolled out in 2009 prohibits providers from selling drugs for profit. In parallel, in different provinces, experiments have been set up aimed at combining (or replacing) the fee-for-service payment and capitation (payment of a flat fee per patient). The study by Yip et al. (2014) involved a selection of township health centres (THC) in Ningxia province (a poor province), where insurance under the New Cooperative Medical System (NCMS) providing fee-for-service payment was replaced by a capitated budget coupled with pay-for-performance. Capitation had to cover the cost of outpatient services per each NCMS enrolee at each THC and include the costs of the village posts under its supervision (supervising these is one of the THCs' missions). At the beginning of every year, the payer (the NCMS) disburses to each THC 70% of its budget. A score card is established for each THC and the NCMS compares its performance to the average THC score for the county. THCs with under-average scores do not receive the remaining 30% of their budget but, instead, are paid an amount calculated on a scale that factors in its deviation from the average. Contrary to the practice of most countries that have implemented RBF schemes, here the system relies on penalties rather than positive incentives. Burundi is another country that has introduced a 'quality penalty'. The report on the implementation of PBF in 2015 (Burundi, Ministère de la Santé, 2016) underlines that 'at the level of Heath Centres, the quality bonus obtained in 2015 represents only 8.39% of all the funds received through PBF. This is linked to the fact that, during the year 2015, several Health Centres in the townships of Bujumbura and the Provinces of Mwaro and Ngozi were sanctioned by the quality penalty. NGOs have been recruited as from 2015 to carry out evaluations of the quality of Health Centres in 6 Provinces (instead of Health Districts). The rigour of these NGOs has led to a decrease in quality scores' (p. 70).

The results

Yip et al. (2014) and Sun et al. (2016) have analysed the effects that RBF-type schemes have had on the quality of prescribing practices, particularly on antibiotic prescriptions. On this count, the situation had become all the more worrying since the poor use of antibiotics not only reinforces resistance to them, but also drives up the cost of care. Compared to what was observed in the control group, the implementation of RBF led to a reduction in antibiotic prescriptions of around 7% in the THCs and 6% in the village posts. Interestingly, the effects were greater on prescriptions

for injectable antibiotics than on those for oral antibiotics, as in parallel to the RBF mechanism, the fee-for-service (about €0.5) paid for each injection was eliminated. Yip et al. (2014) also integrated into their study patients suffering from a cold: prescribing antibiotics for this pathology is not recommended as they have no curative effect. Here again, the effect of RBF was significant and relatively large (-10%). Lastly, the analysis showed that these results are encouraging with respect to the quality of care, but indicate no change in patient satisfaction.

The study by Sun et al. (2016) found similar results to Yip et al. Their study was conducted on a sample of THCs in two counties in Shandong Province, where income levels per capita are much higher than those in Ningxia Province. The mechanism studied by Sun et al. had similar orientations to that implemented in Ningxia. The authors distinguished two large groups of THCs: the first (group A[12]) was paid through a capitation system with penalties (capitation + PBF) as explained earlier (but with an 80/20 rather than 70/30 share); the second (group B) was paid entirely through capitation but, unlike group A, they were unaware of the conditions that determined payment of the remaining 20% of their budget. Dividing the THCs into two groups made it possible to isolate the separate effects of PBF and of the switch to the capitation system.

The results suggest that the negative incentives embedded in the PBF component to complement the capitation payment (group A) reduced 'inappropriate' drug prescribing for some but not all indicators, compared to the observations for group B (global capitated budget). However, results differed between the two counties studied. The effect was greater in the county where the THCs' initial level was the furthest from the threshold that triggered penalties, and thus more motivated to make efforts. This confirms, as observed in several other countries, that the size of the incentives partly impacts their effects. However, in the same county, 'compensatory' behaviour was observed on the part of the doctors, suggesting that PBF led to a 20% increase in prescription spending despite a reduction in inappropriate prescribing.

While these two studies clearly show the utility of implementing pay-for-performance mechanisms to combat undisclosed motives for inappropriate prescribing, they also show the need to put in place complementary interventions. Lastly, both studies explain that, for different reasons, the authorities opposed the implementation of stricter evaluation protocols than those that were used – and this is an important lesson from the policy viewpoint.

Powel-Jackson et al. (2015) examined the impact of a shift from a fee-for-service payment method to a capitation budget with pay-for-performance amongst primary care providers along with measures aiming to redesign of the rural insurance benefit package in order to reorient patients away from inpatient care towards outpatient care. They found that the insurance intervention alone led to a 47% increase in the use of outpatient care at village clinics, but with an increase in injections. By contrast, they highlighted that the two interventions in combination showed no

[12] Thus termed by the authors.

effect on health care use over and above that generated by the redesign of the insurance benefit package alone.

2.7. Argentina: RBF in an upper-middle-income country

The scheme

In Argentina, RBF is based on Plan Nacer, which has been gradually rolled out since 2004. At the outset, it covered the country's nine poorest provinces, then later was scaled up to the national level. The programme aims to improve MCH for people who have no formal health insurance. A free pre-set package of health services is delivered in Nacer clinics to those eligible: pregnant or lactating women (up to 45 days after birth) and uninsured children under the age of six (World Bank, 2013). RBF payments come on top of government funding. It is paid every four months to the provincial authorities rather than to the health providers directly in the form of a capitation payment, the amount varying in line with performance on health outcomes. In fact, the amount of financing depends on the number of people enrolled in the programme ($5 per eligible person enrolled in the programme; Gertler et al., 2016) on top of this, a payment is made depending on the achievement of ten 'tracer' indicators linked to maternal and child health.[13] These target indicators are subject to annual negotiations between each province and the National Ministry of Health (World Bank, 2013). The province then pays the RBF amount, plus the usual budget payment to each Nacer clinic, on a fee-for-service basis to cover the health services delivered under the scheme (based on the set list of services covered by the healthcare package; Gertler et al., 2016). The way in which the clinic uses RBF is framed by rules delimiting the range of possible utilisations: financing medical equipment, maintenance work, staff recruitment or staff bonuses (up to 50% of RBF; World Bank, 2013; Gertler et al., 2016). Within this scope, the way in which the funds are used is left to the clinic's discretion (Gertler et al., 2016).

The results

The findings of the World Banks's 2013 report on Plan Nacer were positive overall: at the end of 2012, out of the eleven target indicators, seven had been fully achieved or surpassed and three had been almost achieved. The impact of Plan Nacer has been evaluated in more detail by Gertler et al. (2016). Two main types of results were highlighted in the evaluation: the so-called direct programme impact on the target population and the so-called indirect impact on the non-target population (spillover effects). The direct programme impact concerns uninsured pregnant women, new mothers and children under six who are enrolled in the Plan Nacer. Overall, the evaluation conducted by Gertler et al. (2016) highlighted the programme's positive effect. RBF helped to improve the care given to pregnant women as shown by the increase in the number of antenatal care visits (+0.68) and the probability of receiving tetanus vaccine during pregnancy (+5.6%). The

[13] At the outset, payment for achieving the targets was only disbursed if the ten targets had been reached ($3 per person; Gertler et al., 2016). The system was further developed to allow variable amounts depending on the number of targets reached (World Bank, 2013).

conditions for delivery also improved with a reduced number of caesarean births (-21%) and a lower probability of a newborn having a low birth weight. Given the current knowledge on the importance of the 'foetal origins hypothesis' on an individual's life cycle, these results point to probable longer-term effects on the generation that benefited from the programme. The most important effect of RBF was measured by the reduction in neonatal mortality: -74%. According to the authors, this metric appears to encompass the improvements in the care given pregnant women, which in turn helps to raise birth weight, itself a factor that strongly impacts a newborn's chances of survival.

Apart from these direct effects specifically targeted by Plan Nacer, the study also reported on the indirect effects on women and children not included in the programme but who visit Nacer clinics. Overall, the authors showed that the care dispensed to these pregnant women improved, with a rise in the number of antenatal care visits and in the probability of receiving a tetanus vaccine. Similarly, delivery conditions improved with an increase in birth weights and, along with this, a reduction in neonatal mortality: -22%. For the authors, the programme spillovers suggest an overall improvement in the quality of care dispensed in Nacer health facilities, whether or not they participated in the programme (new 'good' healthcare habits seem to have been adopted, not simply among eligible patients). However, the volume of antenatal health visits of non-beneficiary pregnant women was adversely impacted. This negative programme externality seems to indicate, according to the authors, that healthcare supply was re-oriented away from non-eligible patients to eligible ones.[14]

3. Getting the best from RBF and avoiding pitfalls

Five major groups of questions emerge in the literature forming elements that need to be considered to put the possibilities of the RBF approach to best use and avoid the pitfalls that many countries find themselves facing to different degrees.

3.1. One size does not fit everyone: context is crucial

Results strongly depend on the context. In Zimbabwe, for example, where the World Bank conducted a difference-in-difference evaluation, it was crucial to take the differences in context into account to estimate the counterfactual given that, during the first years of RBF, health indicators had improved across the entire population due to the growth that followed the economic crisis at the end of the last decade. Moreover, the situation in those districts selected for the pilot phase was more disadvantaged overall, and health-wise, than the average in rural districts. Renmans et al. (2016) highlighted situations where there are a plethora of weakly coordinated initiatives that may be competing with each other, and contractual arrangements that embed contradictory incentives that conflict with those of other projects.

[14] As mentioned above, a summary of these six case studies can be found in the annexed table with a brief description of the methods used.

In general, if context is not sufficiently taken into consideration, this may lead to incentives being mistargeted. Some of the objectives targeted directly by incentives for service providers very much depend on user behaviour: the incentive is thus – partly – mistargeted. For example, complying with the four antenatal visits recommended by the WHO, completing a child's vaccination schedule or adhering to a course of treatment ultimately depends on the users' decisions, which are certainly influenced by provider behaviour. But it is not solely a matter of improving supply. Mixed-incentive RBF schemes (supply and demand sides) may be worthwhile in many cases and should be considered with the greatest attention.

Two other questions confirming the importance of context for the success of RBF programmes emerge from the literature. *The first* is the adequate size of the incentives. A positive relationship is found between the size of the incentives and the effects on health outputs. It is a function of the relation between the effort required from the service provider and the perceived incentive: 'Is it really worth it or not?'. The effort will be adjusted accordingly. This points to the threshold effects below which the incentives have no effect on health results[15] and do no more than create a deadweight effect due to the information asymmetry between agent and principal. An incentive perceived as too small by the beneficiaries may even encourage them to use strategies to seek spin-off benefits. But more broadly, this raises the question of knowing whether the additional payment is designed to encourage healthcare workers to do what they are already paid for, but which is not done for various reasons (absenteeism, other personal activities, etc.), or whether the financial incentive is indeed payment for an additional workload, over and above what is included in the employee's job description. The literature very rarely makes this explicit. The issue of the size of an incentive and the linearity of its effect is also found in the literature on the developed countries. It also brings up again the question of the providers' extrinsic motivations.

Secondly, the recipients of the incentive should have no uncertainty about obtaining the payment due to delays in payment, ineffective communication or individual assessment tools. Staff who expect to benefit from RBF should have a clear understanding of the scheme. When this is not the case, a survey in Nigeria shows that this undermines confidence in the RBF programme, impairs the motivations of some staff, adversely impacts provider's behaviour and obstructs potential improvements that could benefit health facilities (Ogundeji et al., 2016b). Reducing the risk of non-payment of the incentive thus increases the probability and magnitude of a positive effect of RBF schemes. This point is particularly interesting with respect to the sustainability of RBF schemes, where local financing will ultimately replace external financing.

The profile of the incentive beneficiary (individual or team) seems to have no clear-cut effect on results. However, the literature on high-income countries suggests that, for the same overall amount, payment to the health facility creates less incentive than direct payment to individuals or teams. For the time being, these questions have been insufficiently documented. The same is true

[15] Kandpal (op. cit. p. 14) notes that 'the relatively low power of RBF incentives in relation to guaranteed salary (often on the order of 10% of salary) may limit some of the possible gains from RBF schemes'.

of the comparative effectiveness of incentives based on inputs or outputs, a question on which there is very little work. A study of the merits of performance contracts rewarding input use or outputs in Indian pregnancy and maternity care (Mohanan et al., 2017) found that providers with input and output contracts produced comparable gains in health (20% reductions in post-partum hemorrhage, the leading cause of maternal mortality in India). Interestingly the authors point out that the input contracts cost 25% less than output ones for the same achievement, and contrary to expectation, providers with output contracts did not innovate more than the others.

Furthermore, the literature shows how important it is, depending on the context, to start by addressing simple problems and gradually address the most complex challenges. Kandpal (2016), based on impact evaluation in Argentina, Cameroon, Rwanda, Zimbabwe and Zambia, points out that these studies show 'the importance of continued innovations on ways to intelligently measure and incentivize quality measures of care in maternal and child health, which are more complex than coverage indicators. Related to this, given that quality of care is multidimensional, starting with structural quality indicators and then progressively introducing process measures of clinical care is critical to allow health providers to address less complex quality of care issues first, develop better understanding of RBF and quality of care, and then shift gradually toward more demanding measures of care under the RBF pilot' (p. 14).

3.2. When reported results depend on evaluation methods, the quality of data and information collected

As we have seen, the literature shows that the results of RBF are mixed but quite 'globally positive' with respect to the main indicators targeted in developing countries. Yet, these positive results – although mixed – could well be either more mitigated or, in some cases, better than appears at first sight. This is not only due to a well-known publication bias that often leads to overrepresentation of interventions whose results concur with the expected outcome, but also because many studies are plagued by methodological weaknesses with several flaws that limit their significance in one direction or another. There is thus a huge need to step up *rigorous* evaluations on a large scale.[16] Rigorous does not mean that one should *limit* exclusively on randomised controlled trials since other analysis methods exist which bring in addition very useful information (as focus groups, SWOT analysis for instance) in helping to improve policy.

From this perspective, there is a need to develop *qualitative research* – ex ante when designing the project, and ex post when evaluating the effects – tied into the theories of change that should (normally) underpin RBF-type interventions. This is indispensable to gain deeper insight into the reasons for the success and failure of interventions, and to draw relevant lessons. It is just as vital to understand why an intervention has or has not worked as it is to know whether or not it has produced the expected results. This is a crucial question for policymakers when it comes to

[16] On this point, see for example, the Health Results Innovation Trust Fund (HRITF) created in 2007 to support RBF in view of improving maternal and child health. The fund is administered by the World Bank and supported by Norway and UK (DIFID).

exploring the possible scaling-up of pilot experiments, or grasping what makes an experience in one country transferrable to another country or context. To meet these challenges, qualitative research can be very useful. However, this seems to be not only underdeveloped,[17] but also under-utilised in evaluation processes, even though it could help to identify some elements that could make RBF schemes more effective. Cataldo and Kielmann (2016) point out that in the 17 case studies they reviewed, they found 'no instances in which the qualitative research served to help identify or operationalise the constructs relevant to understanding the impact of RBF schemes on health systems components, and specifically the health workforce. Such research would have helped to understand locally relevant definitions and sources of "motivation", functional as well as more context and culture-specific dimensions of "quality" in performance, but also specific constructs to characterize organizational culture, including, for example, dimensions of management and leadership style – hierarchical, vertical, horizontal – that reflect broader societal norms based on gender, occupation, status and so on' (ibid., p. 12).

A further question arises: knowing to what extent the RBF approach itself leads to *manipulation/gaming* of the data that constitute the yardstick for incentive payments and the evaluation of results. If the raw data are (partly) manipulated, the problem cannot be solved simply by setting up a strictly independent evaluation process. Some examples of manipulation or gaming are reported in the literature (for instance, Kalk et al., 2010; Chimhutu et al., 2014). Additionally, informal feedback from field experiences confirms that the issue needs to be taken seriously. It might be thought that the risk of manipulation is especially high in very corrupt countries and countries where freedom of the press is weak and investigative journalism is discouraged.

Moreover, the RBF approach can offer a favourable ground for reinforcing information asymmetry to the detriment of the principal(s) by favouring what is called 'active information avoidance' behaviours by Golman et al. (2017) if information is expected to be adverse, carrying direct negative utility for (some of) the agents. It can also be the case with interpersonal strategic avoidance where information avoidance is an incidental consequence of wanting others to remain uninformed, because strategic avoidance of information can strengthen the agents' bargaining position (Golman et al., op. cit.) in designing incentives or setting up measures that could arise from the monitoring process.

This raises the issue of the *credibility of the verification mechanism* (cf. the examples given in Renmans et al., 2016) and the principal-agent relationship within the RBF scheme. What interest do the agents in charge of verification have to provide accurate information? In Burundi, Bertone and Meessen (2013) report a conflict of interests in Ngozi province where verification was to be conducted by the District Health Bureaus, which were in turn evaluated on the performance of the health facilities. The authors point out that Bubanza province did not face this problem as verification was performed by the Purchasing Agency. Renaud (2013) underlines that 'At the

[17] In contrast, there is an abundant methodological literature on these questions (Green and Thorogood, 2013).

beginning of the RBF rollout to the whole country, hospital technical quality assessment was done by peers from the same province as the hospitals they were evaluating. It frequently happened that the assessors had themselves been evaluated by those peers from the hospitals they were evaluating. This system could facilitate collusion or, conversely, result in conflicts. For example, in several instances some evaluators gave bad grades in retaliation for a bad grade they had received in the past from the hospital they were assessing. From 2012 on, it was decided that evaluators could come from any hospital in the country that is involved in the RBF scheme. This has not eliminated the problem (only 50 hospitals are involved in the scheme), but it has reduced the scope of the problem. It was decided that a facilitator would be added to the evaluation team. He or she is a third-party observer, commissioned by the CT-FBP, is generally a member of the RBF technical assistance at provincial or central level, and represents the CT-FBP during the peer assessment. The facilitator is to prevent or solve potential conflicts during the assessment. During a quality assessment visit, which the author of this study observed, it was noticed that the role of this facilitator was crucial both for answering technical questions on how to interpret the checklist, and for easing tensions between evaluators and hospital staff' (ibid., p.17).

In their study of RBF verification in Afghanistan, Cashin, Fleisher and Hashemi (2015) consider that the main lesson to be learnt – and one that holds across all countries – is that the verification process must crucially and realistically considers the uneven capabilities of the reporting system to produce reliable data. It is also vital to question the potential value added that more sophisticated approaches could bring. They also show that the limits of what communities can contribute must be taken into account without idealism: 'The community-level verification, while intuitively attractive, has proven to be the most problematic aspect of the process in Afghanistan….There has been an improvement in data quality over the course of the RBF program, and while we do not know the level of gaming that would have occurred in the absence of verification, it is not clear that the cost of the community-level verification - both financial and patient privacy[18] - was justified by the gaming that was potentially avoided' (ibid., p.31). Conversely, Cordaid (2015) shows that, in Africa, if communities are successfully involved in the process of selecting indicators to verify outputs, this not only helps to restore the social contract between citizens and state service providers, but also empowers these communities.

That said, on top of the debate on verification methods, there is a broad consensus that verification needs to be independent. Savedoff (2015) rightly considers that independent verification of results is a cornerstone of RBF programmes 'because these features are most consistent with the way development, innovation and progress typically occur' (p.1).

[18] Also, taking into consideration risk to field teams in this specific context.

3.3. Short and long-term effects, and the sustainability of RBF mechanisms

Several questions are intertwined. In the overwhelming majority of studies, the results cover a short time period, usually one to several years. However, for many reasons, long-term results may differ from short-term results, without prejudging which direction this evolution will take. For example, having no results over three years does not necessarily mean an absence of positive (or negative) results in the following years. The core issue here relates to changes in the behaviour of 'RBF agents' and 'outside RBF agents', (i.e., the service providers), changes in the behaviour of households faced with an evolving or stabilised incentive scheme, and possible developments in the overall environment (socioeconomic context, etc.) and specific environment (changes in the characteristics of healthcare supply, a changing epidemiological profile, the introduction of new interventions, etc.). Werner et al. (2011) provide an interesting and well-documented example regarding the time path of RBF effects in the United States. They examined the effects in 260 hospitals of a pay-for-performance demonstration project carried out by the Centers for Medicare and Medicaid Services in partnership with Premier Inc., a nationwide hospital system. They compared these results to those of a control group of 780 hospitals not in the demonstration project. The performance of the hospitals in the project initially improved more than the performance of the control group: More than half of the pay-for-performance hospitals achieved high performance scores, compared to fewer than a third of the control hospitals. However, after five years, the two groups' scores were virtually identical.

When the effects of RBF erode over time, there is the question of maintaining or removing incentives, which is a very poorly documented one. In the DRC, staff attendance increased during the intervention period but then dropped significantly when the intervention stopped (Norad, 2015). This was analysed as a detrimental effect of RBF on intrinsic motivations.

The temporality of the effects of RBF also relates to the *financing of the programmes*. Most RBF mechanisms are strongly supported by external partners (States, multilateral organisations, trust funds, private foundations, NGOs). In Burundi in 2014, donors (mainly the World Bank) financed 48% of the PBF implementation, including health facilities, regulation schemes and operational costs, and 36% of the amount devoted to the health facilities alone (Burundi, Ministère de la Santé, 2015). What happens and what will become of RBF when external financing is phased out or shrinks? This crucial question still lies in uncharted territory although there is some good news. In Plan Nacer in Argentina, it has been tested whether providing a temporary increase in financial incentives to clinics would encourage providers to initiate care for pregnant women in the first trimester – and whether this would continue even after the additional bonus had been stopped. The evaluation (World Bank, 2015) – which according to its authors was the first of its kind to examine the effects of short-term incentives on long-term performance – found that the boost worked and that better care continued even after the incentives had ended. As stated by the study, 'the results provide valuable insights into the possibility of using temporary incentives to change behavior over the long term' (p.7).

This issue is directly linked with the debate about the so-called '*crowding motivation theory*', holding that monetary incentives in RBFs (one form of extrinsic motivation) would reduce intrinsic non-materialistic motivations, generating adverse effects on areas of importance for health but outside the targeted scope of RBF. This may in the long run prove damaging to the realisation of health policy objectives, etc. The literature is inconclusive on these issues although there are examples of rent-seeking behaviour, shirking, cherry-picking and task trade-offs (Renmans et al., 2016; van Herck et al., 2010; Basinga et al, 2011; Lannes et al., 2016). Bertone and Messen (2013) have observed in Burundi that bonuses have gradually come to be seen by the health workforce as an entitlement or fixed salary supplement, which may have led to an erosion of intrinsic motivation, and completely distorts the true nature of the PBF system. In contrast, a qualitative study in Nigeria, comparing the perceptions of a sample of health workers in facilities receiving PBF finance with other non-participating facilities, showed that the programme had strengthened the non-financial motivations of the workers who were paid bonuses (pride, punctuality, etc.) and facilitated the emergence of an environment more conducive to a more efficient running of the health facilities, despite a heavier work load judged to be difficult (Bhatnagar and George, 2016).

3.4. Externalities, side and net effects

Conceptually, *positive externalities* are expected from the RBF approach regarding its contribution to reforming healthcare systems, driving a dynamic conducive to greater efficiency for healthcare facilities and systems, and improving health indicators through various spillover effects. However, the real contribution of these externalities is insufficiently documented, apart from some specific examples for which we have more information (such as Cambodia, Rwanda, Zimbabwe and the Plan Nacer, among others; other examples can be found in HRTIF, 2013). Renmans et al. (2016) report findings from studies showing that in Burundi, Rwanda and Tanzania, RBF led to more outreach activities and new initiatives to increase performance on the indicators, professionalism and compliance with national norms, reduced absenteeism, and favour more cooperation. Based on impact evaluations in Argentina, Cameroon, Rwanda, Zimbabwe and Zambia, and to a limited extent, from the Haut-Katanga pilot project in RDC, Kandpal notes (p.12) that 'There is also evidence of general health system strengthening in terms of more active supervising and monitoring roles, more quantifiable involvement with communities, and increased health worker satisfaction. These system-level impacts can have knock on effects on population-level health outcomes that may extend well beyond the life of the evaluation period'.

Interestingly, and in the same vein, Cashin et al. (2014) highlight that the case studies in European countries show that most of the 'PBF programmes did contribute to a greater focus on health system objectives, better generation and use of information, more accountability, and in some cases a more productive dialogue between health purchasers and providers. This also can be described as more effective health sector governance and more strategic health purchasing' (p.43). They emphasise that, in line with various reviews, PBF programmes in their entirety may be more powerful than the sum of their parts. Their findings also show that PBF programmes have played

an important role in the evolution of health care systems through 'their reinforcing effects on broader performance improvement initiatives, and their spillover effects, or other health system strengthening that occurs as a by-product of the incentive programmes' (ibid. p.14). They also report that the improved generation and use of data for performance improvement, faster uptake of IT, more quality improvement tools (e.g., guideline-based decision aids), sharper focus on priorities, and better overall governance and accountability,[19] may have been highly important outcomes of RBF programmes and, in some cases, more important than improvements in performance indicators.

One salient question that needs to be asked about the effects of RBF, in the light of the increase in attendance in health facilities, in assisted deliveries or vaccinations that is directly linked to the incentives, is knowing whether this represents *a net increase* in activity, or whether it is simply a transfer of demand coming from the health facilities not concerned by RBF projects. If this is the case, there is no net effect, and at the level of the observed area, the global attendance of health facilities, etc. has not changed. We mentioned earlier a few examples that suggest there is a partial transfer of demand, but the literature makes it impossible to reach a firm conclusion, especially as it would be necessary to be able to simultaneously compare the quality of care in RBF health facilities and the others. If the quality is higher in the RBF facilities, the effect remains positive overall, even if total demand for care remains unchanged.

3.5. Cost-effectiveness

This is obviously a foremost question in a context where governments are facing a lack of funding for their health systems and where many low-income countries are failing to allocate the $60–80 per capita and per year needed to ensure the proper functioning of the first level of the health pyramid (Dieleman et al., 2016). However, knowledge on the cost-effectiveness of interventions funded by RBF-type arrangements is sorely lacking[20]. As pointed out by Eijkenaar et al. (2013), RBF 'can be considered cost-effective when improved quality is achieved with equal or lower costs or when the same quality is achieved with lower costs. Even in case RBF leads to cost increases it may still be viewed as cost-effective, as long as quality improvements are large enough' (p.117). Tanzania is one of the few low-income countries in which a thorough and systematic cost-benefit analysis of RBF has been performed (Borghi et al., 2015). In their study, cost-effectiveness was defined as the incremental economic cost per additional birth in a health facility. The authors found that managing the pay-for-performance programme was the costliest component of ongoing implementation and exceeded the costs of financial incentives by between 1.7 times (in financial costs) and 1.9 times (in economic costs). Yet, the study only provides some rough

[19] Savedoff (2017) shows how RBF can, under a specific design, repair political accountability relations between funders, governments and citizens.

[20] A recent study in France shows, if need be, the crucial importance of this issue. The French mandatory health insurance system had set up in 2012 an incentive scheme to limit the overruns of fees charged by specialists. The analysis shows that in order to avoid 1 € of overruns, the incentive scheme put in place led to an expense of 10 € for the health insurance (Cour des Comptes, Rapport sur l'application des lois de financement de la Sécurité sociale, Paris, 2017).

elements to assess whether the RBF scheme in Tanzanian pay-for-performance was cost-effective or not. It has been compared to alternative maternal and child health interventions, such as demand-side financing, but no cost-effectiveness studies using the same outcomes could be identified in Tanzania. Nonetheless, Borghi et al. (2015) point out that the international literature suggests that the cost-effectiveness of a voucher scheme to promote maternal health through coverage of facility-based deliveries varies from $33 per additional institutional delivery in Uganda to $91 in Bangladesh. In another study, removing user fees for delivery care was estimated to cost $25 per additional delivery. In each of these three studies, the cost per additional delivery is lower than their estimate for RBF in Tanzania ($479 per additional delivery; Borghi et al., 2015). It should be pointed out that although these amounts are not in PPP dollars, the differences are considerable. But, as the authors rightly emphasise, such comparisons should be treated with caution given the differences in data sources, birth rates, cost measurement, etc. This leaves the debate wide open on whether the RBF strategy in Tanzania is cost-effective or not.

Whatever the answer, it is highly likely that RBF will become more cost-effective if it is scaled up to the national level (as in Rwanda and Burundi), as this will ensure that it becomes fully integrated into the health system management over time rather than being implemented through a specific project or programme.

3.6. Community RBF for pushing for better health care: an important but (almost) missing vector in the literature.

Mobilising communities to improve access to better basic care is an old idea, but doing so by introducing RBF-like elements into the desired dynamic is an innovative approach that dates back only a few years. Community RBF (CRBF) involves remunerating community actors on a contractual basis to conduct activities to facilitate access to curative and preventive health services and to information to promote health. The wide interest and the scope of such approach is clear as a complement to the RBF schemes both at the level of the providers and the demand for care. But this issue is very poorly documented at this stage. A recent study by the World Bank (Falisse et al., 2017) helps to begin to fill the knowledge gap in this area by synthesising the first lessons of experiences in five sub-Saharan Africa countries (Cameroon, Benin, DR Congo, The Gambia, Rwanda and the Republic of Congo). It is too early to judge the effects of these initiatives, but the above-mentioned study already reveals some lessons concerning their implementation. It stresses the importance of simple and clear contracts, appropriate communication and careful consideration regarding the amount of payments so that they keep their incentive value. The issue of timely payments is crucial in a context where community actors are often poor. Whether there should be in addition a non-results based payment is still a pending issue, although most schemes have suppressed it because of its lack of incentives. In their study, Falisse et al., also highlight that central to the quality of activities undertaken is the training and monitoring of community actors, 'which is easily undermined by low commitment of district officers and chief nurses'.

4. Conclusion

The literature shows that very extensive grey areas concerning the effects of the RBF approach still exist, and that there is a need to step up efforts to ensure a rigorous analysis of its impact. It is urgent to build up a critical mass of knowledge that could constitute a *global public good* on the topic, which would be of great benefit to all the stakeholders involved in the health SDGs.

This said, despite current limitations, several points emerge in the literature from a policy-oriented perspective, indicating that this approach can very usefully contribute to a 'healthy life…for all at all ages', this being the health objective for SDG 3. Five points hold special importance:

a. ***The RBF approach is a highly promising one, regardless of its mixed results, but it must be very carefully adapted to the context in which it is implemented.*** The literature reveals that the RBF approach has enabled significant headway to be made in multiple domains (regarding mainly outputs, more rarely health status). However, it also shows that these results are by no means systematic as in many cases it highlights an absence of effects and sometimes unexpected and undesirable ones. The approach thus is very promising (and, in our view, it is still underutilised), but it is not the be-all and end-all of approaches that improve the effectiveness and efficiency of health policies or health aid. There are other complementary approaches. On the other hand, there is no reason to rule RBF out on the grounds that several studies have shown the absence of effects in some programmes. Basically, what the literature shows – and this is not particularly original – is that defining and implementing the approach and its appropriateness to the context that it integrates are crucial components that largely determine the results. In other words, one should ask: Was the initial diagnosis of the problem to be dealt with in its context correct? If so, were the measures (incentives) advocated relevant? If so, have they been properly implemented? If so, were they disrupted by unforeseeable exogenous elements?

b. ***It is crucial to define and put in place relevant, context-specific incentives that are grounded in a rigorously argued theory of change.*** A well-designed and effectively implemented RBF scheme needs to align the objectives of the health facilities or the staff concerned with the objectives targeted by health policy: What are the right incentives? Who should be incentivised? As Levy and Peart (2015) remind us, 'Perhaps a policy fails because it fails to align the private goals of acting individuals who administer the policy and those in the collective polity who establish the administrating agencies on the basis of an articulation on public goals….If a policy is designed to address a "public" goal at the expense of the private hopes and desires of those who make up the collective, its failure may be altogether predictable, as those whose hopes and desires conflict with the policy are motivated to undermine the "public" policy goal' (p. 699).

c. **All necessary efforts should be engaged in terms of getting the most of all the components a RBF approach can offer.** RBF is not limited to monetary incentives. This means that questions should be answered, taking into consideration the context, such as: Which components of the RBF 'package' matter most: Financial incentives? Contracting mechanisms? More autonomy in decision-making? Dialogue between the stakeholders? Improved reporting based on better data and information? Well-designed and effectively implemented RBF schemes, whose results are verified by independent organisations to avoid any conflict of interest, and properly evaluated, can generate positive externalities conducive to the reform and strengthening of health systems.

d. **Every practical effort shall be taken to ensure that healthcare supply and demand are targeted jointly by RBF schemes if this proves relevant**. Associating RBF schemes centred on healthcare supply (the focus of this paper) with RBF schemes centred on healthcare demand would help to reduce specific barriers on both the supply side and the demand side, and could create important synergies likely to amplify the specific effects of each type of programme.

e. Finally, to recall what we have already pointed out, vast grey areas persist, opening up the field to **more research** that is indispensable to take full advantage of RBF approaches and creating a valuable global public good that will benefit all potential stakeholders.

Annex. A selection of studies used for the case-studies, with methodologies and main findings

Country	PBF schemes	Study design / methodology	Main findings
Afghanistan (Engineer et al., 2016)	P4P scheme implemented in 2010 Objectives: • Quantitative to improve cover and access for MCH services covered by the scheme • Qualitative to improve quality and equity of care for MCH services Utilisation of payment: • Bonuses to health workers	Cluster-randomised trial covers 442 health facilities spread over 11 provinces between September 2010 and December 2012	Positive effects on: • Probability of non-poor women giving birth in health facility • Healthcare quality: spending more time with the patients, conducting more complete examinations, and counselling No effects on: • P4P indicators of MCH coverage • Equity • Quality measured by a balance scorecard
Argentina (Gertler et al., 2016)	Plan Nacer gradually rolled out since 2004 Objectives: • To deliver a free pre-set package of MCH services for pregnant or lactating women (up to 45 days after birth) and uninsured children under the age of six with no formal health insurance Utilisation of payment: • Medical equipment, • Maintenance work • Staff recruitment • Staff bonuses	Non-random assignment of the scheme. The authors exploit the geographic phasing of the scheme in seven provinces over the period 2004-2008. The authors use a difference-in-difference approach to estimate the intent-to-treat effect, the treatment on the treated (TOT) and spillovers effects.	Direct effects (TOT) of the scheme on: • Number of antenatal care visits: +0.68 visits • Probability of receiving tetanus vaccine during pregnancy: +5.6% • Number of caesarean births: -21% • Probability of newborns having a low birth weight: -19% • Neonatal mortality: -74% Spillovers effects (on the non-targeted population): • Birthweight : negative effect • Tetanus: negative effect • Caesareans : negative effect • Antenatal visits: no effect • Neonatal mortality: -22%

Burundi Bonfrer et al. (2014b)	Proactive national strategy since 2010 Objectives: • Quantitative based on the volume of activity and rate of coverage (initially a minimum package + MDGs 4 and 5 + malnutrition since 2013), paid on a monthly basis • qualitative based on the quality of care, paid on a quarterly basis Utilisation of payment: • expenses and cash requirements of the health facilities • worker bonuses possible with remaining budget	Non-random assignment of the scheme. Difference-in-difference model applied to a nationally representative sample of 4916 women in 2010 to exploit the gradual implementation of the scheme between 2006 and 2010.	Positive effects on: • Share of women receiving delivery assistance in a health facility: + only for non-poor women • Follow-up of pregnant women: + blood pressure, + probability of tetanus vaccination, with higher effect for non-poor pregnant women rather than poor ones • Probability of a child being vaccinated: +, stronger for poor children than for non-poor children No effects on: • Probability of a woman having more than one antenatal care visit • Probability that this antenatal care visit occurs during the first six months of pregnancy • Equity of access to healthcare
Cambodia Van de Poel et al. (2016)	Pilot phase 1999-2003: • Pilot-Out • Pilot-In Roll-out in 2004: • IN (2004-2008) for 11 ODs • Internal contracting (2005-2010) for 8 ODs • GAVI/RBF scheme (2007) for 10 ODs Consolidation since 2009: • Internal contracting	Non-random assignment of the scheme. Difference-in-difference linear probability model applied to a nationally representative sample of women of reproductive age observed in 2000, 2005, 2010, which gives information for the 1995-2010 period.	Positive effects of PBF: • Probability of a child being born in a public health facility: + 6-7.5% • Probability of the non-poor giving birth in a health facility: + 13.4%, but not for the poorest women • Effects are enhanced when supply-side schemes are combined with demand-side ones (i.e., vouchers) • Positive effect seems likely to be a shift of demand rather than a net increase of demand No effects on: • Neonatal mortality • Vaccination rates • Antenatal care Differences between PBF schemes: • Internal contracting had much greater effects than IN scheme • IN model was the only one to have a positive impact on antenatal visits

			• The scheme incorporating a contracting mechanism in which district staff comprised government employees, nonetheless supported by NGOs under contract with the Ministry of Health was the only one to have a positive impact on the probability of a woman having two antenatal visits • PBF schemes rolled out by GAVI: no significant effect on vaccination rates because of high immunisation rates prevailing before
China (Yip et al., 2014; Sun et al., 2016)	Objectives: • Increase of the quality of prescriptions: reduction of drugs expenses and inappropriate drug prescriptions Specificity of the scheme: • Penalty system	Yip et al. (2014) in Ningxia province. Matched-pair cluster randomisation over 2009-2012 Sun et al. (2016) in two counties of Shandong province. Non-random assignation of the scheme. Difference-in-difference over 2011-2012 to exploit the implementation of the scheme.	• Reduction in antibiotic prescriptions of around 7% in township health centres (Yip et al., 2014) • Reduction of inappropriate prescriptions of antibiotics (for patient suffering from a cold): -10% (Yip et al., 2014) • Reduction of inappropriate drug prescribing for some indicators (Sun et al., 2016) • Size of incentives partly impacts effects (Sun et al., 2016) • Compensatory effects highlighted (Sun et al., 2016)
Rwanda (Lannes et al., 2016)	Pilot project in 2001 Roll-out from 2006 Objectives: • Quantitative to improve cover and access • Qualitative Utilisation of payment: • Bonuses to health workers • Facility supplies and equipment	Randomised control trial and difference in difference model over the period 2006-2008 for 166 primary healthcare facilities and 2145 households	Positive effects on: • Number of deliveries in health facilities: for the upper-group but not for the lower • Use of modern contraception: for the upper group • Probability that children will benefit from curative and preventive care: for both upper and lower groups • Deliveries in health facilities and access to curative and preventive care: for the lower group Negative effects on: • Use of modern contraception: for the lower group No effects: • Antenatal visits

References

- Alonge, O., Gupta, S., Engineer, C., Salehi, A., and Peters, D.H. (2015). Assessing the pro-poor effect of different contracting schemes for health services on health facilities in rural Afghanistan Health Policy Plan. *Health Policy and Planning*, 30, 1229-1242.
- Arur, A., Peters, D., Hansen, P., Mashkoor, M.A., Steinhardt, L.C., and Burnham, G. (2010). Contracting for health and curative care use in Afghanistan between 2004 and 2005. *Health Policy and Planning*, 25(2), 135-144.
- Basinga, P., Gertler, P.J., Binagwaho, A., Soucat, A.L., Sturdy, J., and Vermeersch, C.M. (2011). Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: An impact evaluation. *The Lancet*, 377(9775), 1421-1428.
- Bertone, M.P., and Meessen, B. (2013). 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 and Planning*, 28(8), 847-857.
- Bhatnagar, A., and George, A.S. (2016). Motivating health workers up to a limit: Partial effects of performance-based financing on working environments in Nigeria. *Health Policy and Planning*, 31(7), 868-877.
- Bonfrer, I., Soeters, R., Van de Poel, E., Basenya, O., Longin, G, van de Looji, F. et al. (2014a). Introduction of performance-based financing in Burundi was associated with improvements in care and quality. *Health Affairs*, 33(12), 2179-2187.
- Bonfrer, I., Van de Poel, E. and Van Doorslaer, E. (2014b). The effects of performance incentives on the utilization and quality of maternal and child care in Burundi. *Social Science and Medicine*, 123, 96-104.
- Borghi, J., Little, R., Binyaruka, P., Patouillard, E., and Kuwawenaruwa, A. (2015). In Tanzania, the many costs of pay-for-performance leave open to debate whether the strategy is cost-effective. *Health Affairs*, 34(3), 406-414.
- Busogoro, J.F., and Beith, A. (2010). *Pay for performance for improved health in Burundi*. https://www.hfgproject.org/wp-content/uploads/2015/02/Pay-for-Performance-for-Improved-Health-In-Burundi.pdf (accessed March 2, 2017).
- Burundi, Ministère de la santé publique et de la lutte contre le sida (2015). *Rapport de mise en œuvre du financement basé sur la performance et la gratuité des soins pour l'année 2014*, Cellule technique nationale Fbp, 92p.
- Burundi, Ministère de la santé publique et de la lutte contre le sida (2016). *Rapport de mise en œuvre du financement basé sur la performance et la gratuité des soins pour l'année 2015.* Cellule technique nationale Fbp.
- Cadot, O. and de Melo, J., ed. (2014), *What have we learnt? Which way ahead?* FERDI, CEPR Press.
- Cashin, C., Chi, Y.L., Smith, P.C., Borowitz, M., and Thomson, S. (2014). *Paying for performance in healthcare: Implications for health system performance and accountability*. McGraw-Hill Education, Berkshire, 338p.

- Cashin, C., Fleisher, L., and Hashemi, T. (2015). *Verification of performance in results-based financing (RBF): The case of Afghanistan* (No. 103694). The World Bank, 50p.
- Cataldo, F. and Kielmann, K. (2016). Qualitative research to enhance the evaluation of results-based financing programmes: The promises and the reality, *Discussion Paper, HNP*, World Bank.
- Chari A., Okeke E., (2014). Can institutional deliveries reduce newborn mortality, *Working Paper* WR-1072, RAND Labor & Population, 47p.
- Chimhutu, V., Lindkvist, I. and Lange, S. (2014). When incentives work too well: Locally implemented pay for performance (P4P) and adverse sanctions towards home birth in Tanzania - a qualitative study. *BMC Health Services Research*, 14, 1-12.
- Chimhutu, V., Tjomsland, M., Songstad, N.G., Mrisho, M., and Moland, K.M. (2015). Introducing payment for performance in the health sector of Tanzania-the policy process. *Globalization and Health*, 11(1), 38.
- Cordaid. (2015). Results based financing – Engaging communities to strengthen systems in fragile contexts. *Position Paper*, Feb, 4p.
- De Walque, D., Gertler, P.J., Bautista-Arredondo, S., Kwan, A., Vermeersch, C., de Dieu Bizimana, J., and Condo, J. (2015). Using provider performance incentives to increase HIV testing and counseling services in Rwanda. *Journal of Health Economics*, 40, 1-9.
- Dieleman, J. L., T. Templin, N. Sadat, P. Reidy, A. Chapin, K. Foreman, A. Haakenstad, T. Evans, C. JL Murray, and C. Kurowski. 2016. « National spending on health by source for 184 countries between 2013 and 2040 ». *The Lancet*. Vol. 387.2521-2535
- Eijkenaar, F., Emmert, M., Scheppach, M., and Schöffski, O. (2013). Effects of pay for performance in healthcare: A systematic review of systematic reviews. *Health Policy*, 110(2), 115-130.
- Eldridge, C., and Palmer, N. (2009). Performance-based payment: some reflections on the discourse, evidence and unanswered questions. *Health Policy and Planning*, 24(3), 160-166.
- Engineer, C.Y., Dale, E., Agarwal, A., Agarwal, A., Alonge, O., Edward, A. et al. (2016). Effectiveness of a pay-for-performance intervention to improve maternal and child health services in Afghanistan: A cluster-randomized trial. *International Journal of Epidemiology*, 451-459.
- Falisse, J.B., Ndayishimiye, J., Kamenyero, V., and Bossuyt, M. (2014). Performance-based financing in the context of selective free health-care: An evaluation of its effects on the use of primary health-care services in Burundi using routine data. *Health Policy and Planning*, 30(10), 1251-1260.
- Falisse, J.-B., Vergeer, P., Medhin, J., Juquois, M., Akpamoli, A., Robyn, J., Shu, W., Zabiti, M., Hassan, R., Jallow, B., Loum, M., Ndizeye, C., Muvudi, M., Makuma Booto, M., Taptue Fotso, J.-C., Sokegbe, S., Magazi, I. and Shapiro, G. (2017). Community results-based financing in health practice: Reflections on implementation from experiences in six countries, *Knowledge Brief*, HNPGP, World Bank.
- Flink, I.J., Ziebe, R., Vagaï, D., van de Looij, F., and Houweling, T.A. (2016). Targeting the poorest in a performance-based financing programme in northern Cameroon. *Health Policy and Planning*, 31(6), 767-776.

- Flores, G., Ir, P., Men, C.R., O'Donnell, O., and Van Doorslaer, E. (2013). Financial protection of patients through compensation of providers: the impact of Health Equity Funds in Cambodia. *Journal of Health Economics*, 32(6), 1180-1193.
- Fox, S., Witter, S., Wylde, E., Mafuta, E., and Lievens, T. (2014). Paying health workers for performance in a fragmented, fragile state: Reflections from Katanga Province, Democratic Republic of Congo. *Health Policy and Planning*, 29(1), 96-105.
- Gertler, P.J., and Giovagnoli, P.I. (2014). Rewarding provider performance to enable a healthy start to life: evidence from Argentina's Plan Nacer. *World Bank Policy Research Working Paper*, WPS6884.
- Gertler, P., Giovagnoli, P., and Martinez, S. (2016). Rewarding provider performance to enable a healthy start of life. Evidence from Argentina's Plan Nacer. *World Bank Policy Research Working Paper* WPS6884.
- Golman R., Hagmann D., and Loewenstein G. (2017). Information avoidance. *Journal of Economic Literature*, 55(1), 96-135.
- Green, J., and Thorogood, N. (2013). *Qualitative methods for health research*. Sage.
- HRTIF. (2013). Annual progress report. World Bank
- Ireland, M., Paul, E., and Dujardin, B. (2011). Can performance-based financing be used to reform health systems in developing countries?. *Bulletin of the World Health Organization*, 89(9), 695-698.
- Kalk, A., Paul, F.A., and Grabosch, E. (2010). Paying for performance. Rwanda: does it pay off. *Tropical Medicine and International Health*, (15), 182-190.
- Kandpal, E. (2016). Completed impact evaluations and emerging lessons from the Health Results Innovation Trust Fund learning portfolio. *World Bank Group, Washington, DC*.
- Kondo, K.K., Damberg, C.L., Mendelson, A., Motu'apuaka, M., Freeman, M., O'Neil, M., and Kansagara, D. (2016). Implementation processes and pay for performance in healthcare: a systematic review. *Journal of General Internal Medicine*, 31(1), 61-69.
- Lannes, L., Meessen, B., Soucat, A., and Basinga, P. (2015). Can performance-based financing help reaching the poor with maternal and child health services? The experience of rural Rwanda. *The International Journal of Health Planning and Management*. Vol. 31(3), 309-348)
- Lannes, L. (2015). Improving health worker performance: The patient-perspective from a RBF program in Rwanda. *Social Science and Medicine,* 138, 1-11.
- Levy, D.M., and Peart, S.J. (2015). Learning from failure: A review of Peter Schuck's 'Why Government Fails So Often: And How It Can Do Better'. *Journal of Economic Literature*, 53(3), 667-674.
- Mathonnat, J., Audibert, M., Huangfu, X., and Petitfour, L. (2015). Progrès et défis pour trois réformes majeures dans le système de santé chinois. *Mondes en Développement,* 170, 105-121.
- Maynard, A. (2012). The powers and pitfalls of payment for performance. *Health Economics*, 21(1), 3-12.
- Miller, G., and Babiarz, K.S. (2013). Pay-for-performance incentives in low- and middle-income country health programs. National Bureau of Economic Research *Working Paper Series*, No. 18932.
- Mohanan, M., Miller, G., Donato, K., Truskinovsky, Y., and Vera-Hernandez, M. (2017). different strokes for different folks: Experimental evidence on the effectiveness of input and output incentive contracts for health care providers with different levels of skills, Economic Research Initiatives at Duke (ERID) Working Paper No. 245, 90p.

- Nimpagaritse, M., Koracahis, C., Roberfroid, D., Kolsteren, P., El Idrissi, M.D., and Meessen, B. (2016). Measuring and understanding the effects of a performance based financing scheme applied to nutrition services in Burundi – a mixed method impact evaluation design. *International Journal for Equity in Health*, 15(93); DOI 10.1186/s12939-016-0382-0.
- Norad. (2015). Experiences with results-based payments in Norwegian development aid. *Report 4*, Bergen, Norway, 34p.
- Ogundeji, Y.K., Bland, J.M., and Sheldon, T.A. (2016a). The effectiveness of payment for performance in healthcare: A meta-analysis and exploration of variation in outcomes. *Health Policy*, 120(10), 1141-1150.
- Ogundeji, Y.K., Jackson, C., Sheldon, T., Olubajo, O., and Ihebuzor, N. (2016b). Pay for performance in Nigeria: The influence of context and implementation on results. *Health Policy and Planning*, 31(8), 955-963.
- Perakis, R., and Savedoff, W. (2015). Does results-based aid change anything? Pecuniary interests, attention, accountability and discretion in four case studies. *CGD Policy Paper* N°52, 61p.
- Powell-Jackson, T., Yip, W.C, and Han W. (2015). Realigning demand and supply side incentives to improve primary health care seeking in rural China, *Health Economics* 24(6):755-772.
- Renaud, A. (2013). Verification of performance in results-based financing (RBF). The case of Burundi. *Health Nutrition and Population Discussion Paper*. Washington D.C.: The International Bank for Reconstruction and Development / The World Bank, 154p.
- Renmans, D., Holvoet, N., Orach, C.G., and Criel, B. (2016). Opening the 'black box' of performance-based financing in low-and lower middle-income countries: A review of the literature. *Health policy and Planning*, 31(9), 1297-1309.
- Savedoff, W. (2015). Paying for results requires verification and autonomy. CGDev Blog 10/27/15.
- Savedoff, W. (2017). Can payments for results repair political accountability relations? CGDev Blog 1/6/17.
- Seheye, E., Shugugu, R., Ndindurwaha, T., and Kamana, J. (2015). Financement basé sur la performance et gratuité: l'évolution des indicateurs du paquet de services de santé au Burundi entre 2011 et 2013. http://www.fbpsanteburundi.bi/cside/contents/docs/FBPet_Gratuite_:_Evolution_des_Indicateurs_du_paquet_de_Services_de_Sante_au_Burundi_entre_2011_et_2013.pdf. 22p.
- Skiles, M.P., Curtis, S.L., Basinga, P., and Angeles, G. (2013). An equity analysis of performance-based financing in Rwanda: are services reaching the poorest women? *Health Policy and Planning*, 28(8), 825-837.
- Skiles, M.P., Curtis, S.L., Basinga, P., Angeles, G., and Thirumurthy, H. (2015). The effect of performance-based financing on illness, care-seeking and treatment among children: An impact evaluation in Rwanda. *BMC Health Services Research*, 15(1), 375-387.
- Sondorp, E., Palmer, N., Strong, L., and Wali, A. (2009). *Afghanistan: Paying NGOs for performance in a post-conflict setting*. In R. Eichler and R. Levine and the Performance-Based Incentives Working Group (Eds.): Performance Incentives for Global Health: potential and pitfalls, Washington, DC: Center for Global Development, 139-164.
- Srivastava D., Mueller M., and Hewlett E. (2016). Better ways to pay for health care, OECD Health Policy Studies, Paris, 165p..

- Sun, X., Liu, X., Sun, Q., Yip, W., Wagstaff, A., and Meng, Q. (2016). The Impact of a Pay-for-Performance Scheme on Prescription Quality in Rural China. *Health Economics*. 25(6), 706-722.
- Van de Poel, E., Flores, G., Ir, P., and O'donnell, O. (2016). Impact of performance-based financing in a low-resource setting: A decade of experience in Cambodia. *Health Economics*, 25(6), 688-705.
- Van Herck, P., De Smedt, D., Annemans, L., Remmen, R., Rosenthal, M.B., and Sermeus, W. (2010). Systematic review: Effects, design choices, and context of pay-for-performance in healthcare. *BMC Health Services Research*, 10, 247.
- Werner, R.M., Kolstad, J.T., Stuart, E.A., and Polsky, D. (2011). The effect of pay-for-performance in hospitals: Lessons for quality improvement. *Health Affairs*, 30(4), 690-698.
- Witter, S., Fretheim, A., Kessy, F.L., and Lindahl, A.K. (2012). Paying for performance to improve the delivery of health interventions in low-and middle-income countries. *Cochrane Database System Rev*iew 2(2), CD007899.
- World Bank. (2013). Argentina. Provincial maternal-child health investment project in support of the second phase of the provincial maternal-child health program. Implementation completion and results report (IBRD-7409). Report No: ICR2618. June 25, 2013, 108p.
- World Bank. (2015). Argentina: can short term incentives change long term behavior? *From Evidence to Policy,* 100060, 4p. Washington, DC. The World Bank. https://openknowledge.worldbank.org/handle/10986/23343. License: CC BY 3.0 IGO.
- Yip, W., Powell-Jackson, T., Chen, W., Hu, M., Fe, E., Hu, M., and Hsiao, W.C. (2014). Capitation combined with pay-for-performance improves antibiotic prescribing practices in rural China. *Health Affairs*, 33(3), 502-510.

"Sur quoi la fondera-t-il l'économie du monde qu'il veut gouverner? Sera-ce sur le caprice de chaque particulier? Quelle confusion! Sera-ce sur la justice? Il l'ignore."

Pascal

FERDI

Created in 2003 , the **Fondation pour les études et recherches sur le développement international** aims to promote a fuller understanding of international economic development and the factors that influence it.

Contact
www.ferdi.fr
contact@ferdi.fr
+33 (0)4 73 17 75 30